Birch pollen allergy: molecular characterization and hypoallergenic products

Martijn F. Schenk

Promotoren:	***Prof. dr. ir. E. Jacobsen*** *Hoogleraar Plantenveredeling, Wageningen Universiteit*
	Prof. dr. L. J. Frewer *Hoogleraar Voedselveiligheid en consumentengedrag, Wageningen Universiteit*
Co-promotoren:	***Dr. M. J. M. Smulders*** *Senior onderzoeker, Wageningen UR, Plant Research International*
	Dr. L. J. W. J. Gilissen *Senior onderzoeker, Wageningen UR, Plant Research International*
Promotiecommissie:	***Prof. dr. A. E. J. Dubois*** *Universiteit van Groningen*
	Prof. dr. ir. H. F. J. Savelkoul *Wageningen Universiteit*
	Prof. dr. R. Shepherd *University of Surrey, United Kingdom*
	Prof. dr. M. S. M. Sosef *Wageningen Universiteit*

Dit onderzoek is uitgevoerd binnen de onderzoeksschool EPS (Experimental Plant Sciences)

Birch pollen allergy: molecular characterization and hypoallergenic products

Martijn F. Schenk

Proefschrift

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. M. J. Kropff
in het openbaar te verdedigen
op dinsdag 19 februari 2008
des namiddags te vier uur in de Aula

Birch pollen allergy: molecular characterization and hypoallergenic products

Martijn F. Schenk

PhD thesis, Wageningen University, The Netherlands, 2008
With references – with summaries in English and Dutch

ISBN: 978-90-8504-873-2

Contents

List of abbreviations

bp.	base pair(s)
cDNA	complementary DNA (based on mRNA template)
gDNA	genomic DNA
ELISA	Enzyme-Linked ImmunoSorbent Assay
EST	Expressed Sequence Tag
GM	Genetic Modification
GMO	Genetically Modified Organism
IgE	Immunoglobulin E
IL-4, IL-5, *etc*.	InterLeukin-4, InterLeukin-5, *etc.*
LC-MS	Liquid Chromatography-Mass Spectrometry
LC-MS/MS	Liquid Chromatography-tandem MS
$LC\text{-}MS^{E}$	Alternate scanning LC-MS
LTP	Lipid Transfer Protein
mRNA	messenger RNA
NJ	Neighbor Joining
OAS	Oral Allergy Syndrome
pAB	polyclonal AntiBody
PR-10	Pathogenesis-Related class 10
RAST	RadioAllergoSorbent Test
RNAi	RNA interference
REML	REsidual Maximum Likelihood
SD	Standard Deviation
SDS-PAGE	Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrphoresis
SNP	Single Nucleotide Polymorphism
SPT	Skin Prick Test
SIT	Specific ImmunoTherapy
T_H-cells	T helper-cells
Q-TOF	Quadrupole-Time-Of-Flight

CHAPTER 1

General introduction

General context: birch pollen allergy

Birch trees grow in the temperate climate zone of the northern hemisphere, and release large amounts of pollen during spring. Birch pollen is a major cause of Type I allergies. The major birch allergen is a pollen protein from the European white birch *(Betula pendula* syn. *B. verrucosa)* termed Bet v 1 (Breiteneder *et al.* 1989; Jarolim *et al.* 1989). Bet v 1 and other Pathogenesis-Related class 10 (PR-10) proteins constitute the largest group of aeroallergens (Breiteneder *et al.* 2000). Individuals who have been sensitized to birch pollen are particularly prone to develop Oral Allergy Syndrome (OAS) due to an IgE-mediated cross-reaction between Bet v 1 and PR-10 food proteins (Ferreira *et al.* 2004; Wensing *et al.* 2002). As a result, PR-10 proteins are also among the four most common food allergens (Breiteneder *et al.* 2000). Allergic diseases have a negative impact on the patient's quality of life and are associated with high economic costs (Meltzer 2001; UCB 1997). Consequently, allergy to birch and OAS represent relevant targets for prevention. Genetic Modification (GM) can be applied to reduce the allergenicity of the plants involved (Bhalla *et al.* 2004; Le *et al.* 2006). However, developments which apply this technique should take societal concerns about GM into account (Frewer *et al.* 2004; Zechendorf 1994).

Before describing the main aims of this thesis, an overview is presented which summarizes current knowledge regarding the prevalence of hay fever, the genetic background of the Bet v 1 allergens, OAS, potential prevention strategies and societal acceptance of different applications of GM.

Hay Fever

Allergic disorders affect a substantial number of people in westernized countries (Aberg *et al.* 1995; Wuthrich *et al.* 1995). Among the existing allergic disorders, hay fever (medical term: seasonal allergic rhinitis) is the most common disorder, affecting 11-19% of the population in Western European countries (UCB 1997). Symptoms are experienced periodically, during the flowering period of various pollen producing plants. Typical hay fever symptoms are sneezing, rhinorrhea and a blocked nose, while itchy eyes (conjunctivitis), throat and ears are other symptoms that may occur. The diagnosis of hay fever may be difficult, because of similarities with non-atopic complaints. Hay fever can be defined on symptom criteria or pathophysiological characteristics, such as the detection of specific IgE by a positive RAST or Skin Prick Test (SPT). A combination of both symptom criteria and RAST/SPT leads to the most reliable diagnosis (UCB 1997). Allergic diseases have a genetic basis; a history of atopy in the family has been identified as an important risk factor (Jõgi *et al.* 1998; Rönmark *et al.* 2003). The occurrence of allergic diseases also depends on exposure to allergens and environmental or life-style factors.

These life-style factors may explain the strong increase in hay fever which has been observed during the last decades (Wuthrich *et al.* 1995).

To substantiate prevention strategies towards hay fever, knowledge on the basic mechanisms underlying allergic disorders is required. An allergic reaction is preceded by a sensitization phase, in which a T-cell response is directed against a particular protein. Eventually, this leads to the production of IgE that is directed towards this protein, which is then considered an allergen. T-helper (T_H) cells play a central role in this process. T_H cells are subdivided according to the cytokines they produce. T_H1 cells secrete, among others, interferon-gamma when stimulated with viruses or intracellular bacteria. By contrast, T_H2 cells secrete cytokines such as Interleukin-4 (IL-4), IL-5 and IL-13 when stimulated with allergens. These interleukins stimulate, among others, IgE production by B-cells. IgE is then bound to receptors on the surface of mast cells. Mast cells are present in several tissue types, such as the nasal mucosa. Upon renewed allergen contact, the allergen binds to the specific IgE, causing the mast cells to release large amounts of histamine. Immediate clinical allergy symptoms are the result.

The 'hygiene hypothesis' (Strachan 1989) proposes that increased affluence has lead to increased domestic hygiene and reduced incidence of infections. Infections in infants would stimulate the T_H2-directed immune system in newborns to mature and develop into a T_H1-directed immune system. A lack of infections causes an imbalance in the T_H1/T_H2 responses, which would induce an immune system that is skewed towards a T_H2 response and consequently results in a higher risk for developing allergies (Yazdanbakhsh *et al.* 2002). Socio-economic factors that relate to an increased affluence, such as lifestyle (Alm *et al.* 1999), increasing use of antibiotics and vaccination (Alm *et al.* 1997), and decreasing family size (Von Mutius *et al.* 1994) correlate with the increased prevalence of allergies. A childhood lived in a farm environment appears to protect against sensitization (Horak *et al.* 2002; Kilpeläinen *et al.* 2002; Riedler *et al.* 2000). It has recently been suggested that the lack of regulatory T-cells, resulting from a low pathogen and parasite burden, causes a deficit in the regulatory network, and thereby underlies the imbalance in the immune system (Yazdanbakhsh *et al.* 2002). The hygiene hypothesis has undergone numerous modifications since it was first proposed. This has not lead to a unifying concept, although various pieces of the complex interplay between the human immune system, various types of infections and atopy have become apparent (Guarner *et al.* 2006; Prioult *et al.* 2005; Schaub *et al.* 2006).

The occurrence of atopic diseases is age-dependent. Atopic eczema and food allergies have a substantial prevalence shortly after birth. Sensitization to pollen allergens is relatively low among young children (Silvestri *et al.* 1996) and increases with age (Rönmark *et al.* 2003; Silvestri *et al.* 1996). Generally, the prevalence of respiratory allergies (allergic asthma and allergic rhinitis) peaks among young adults (Saarinen *et al.* 1995). Hay fever is sometimes considered to act as a 'gate

opener' to asthma as 13-38% of the allergic rhinitis patients suffer from allergic asthma (Meltzer 2001). Sensitization to pollen allergens is a risk factor for the development of asthma symptoms, although the size of the risk may depend on the source of the pollen (a higher risk for grass than for tree pollen) (Chinn *et al.* 1998; Jõgi *et al.* 1998; Plaschke *et al.* 2000; Raukas-Kivioja *et al.* 2003; Soriano *et al.* 1999). However, others studies indicate that sensitization to pollen allergens is associated to sensitization to indoor allergens, which would represent the actual risk factor for developing asthma (Jaen *et al.* 2002; Kerkhof *et al.* 2003). The suspected induction of asthma involves the formation of small pollen-derived particles that (unlike pollen) are able to reach lower airway tracts and induce asthmatic responses (Schäppi *et al.* 1997).

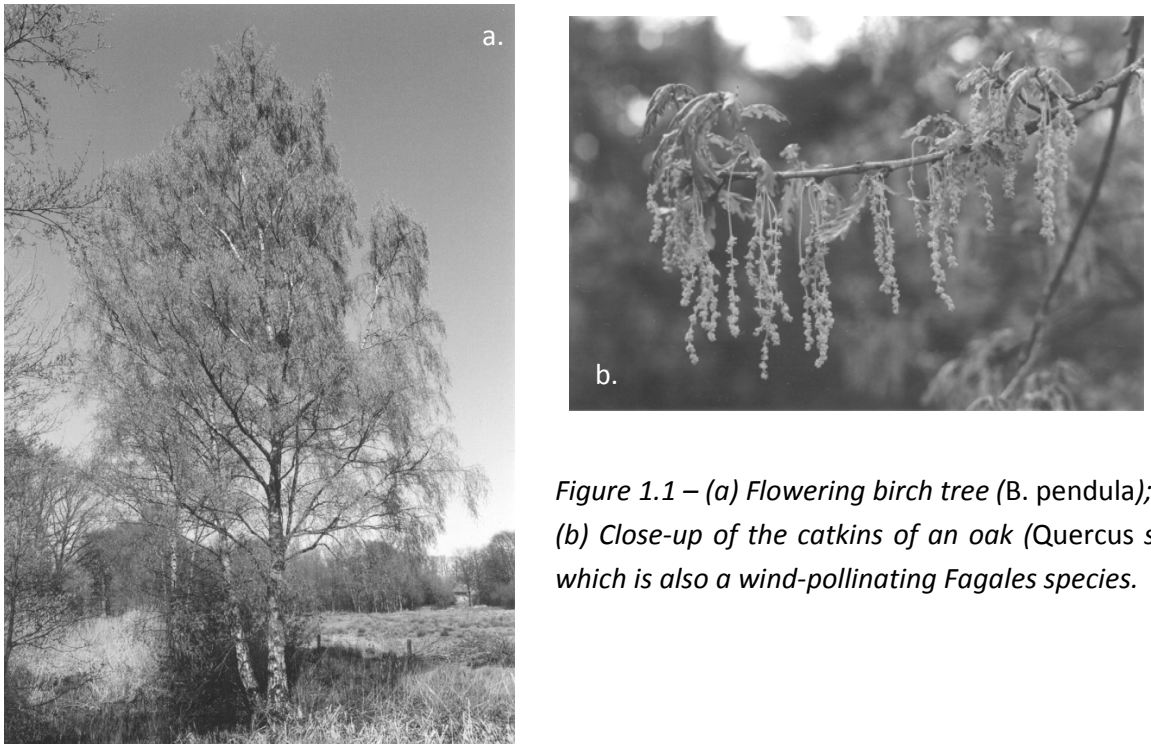


*Figure 1.1 – (a) Flowering birch tree (*B. pendula*); (b) Close-up of the catkins of an oak (*Quercus *sp.), which is also a wind-pollinating Fagales species.*

Exposure to pollen allergens is a key factor in the sensitization process and in the occurrence of allergic reactions after sensitization has taken place. Which pollen allergies occur in different European regions depends on (aero)biological factors (D'Amato *et al.* 1998), such as vegetation composition (Corsico *et al.* 2000) and the degree of urbanization (Crimi *et al.* 1999). Other factors influence the timing and duration of exposure to pollen, like birth season (Björksten *et al.* 1980; Nilsson *et al.* 1997), the exposure to pollen before or shortly after birth (Kihlström *et al.* 2002; Kihlström *et al.* 2003), weather conditions (Latalowa *et al.* 2002), and the fenology of allergenic plants (Emberlin *et al.* 2002). Three groups of wind pollinating species that produce allergenic pollen can be distinguished in Europe: (1) Trees, mainly Fagales species such as birch (*B. pendula*; Fig. 1.1a), alder (*Alnus glutinosa*) and hazel (*Corylus avellana*). (2) Grasses, several species such as

rye grass (*Lolium perenne*), orchard grass (*Dactylis glomerata*) and timothy grass (*Phleum pratense*). (3) Weeds, which form a diverse group of species, including mugwort (*Artemisia vulgaris*), plantain (*Plantago lanceolata*), sorrel (*Rumex sp.*) and nettle (*Urtica sp.*). Alder and hazel flower as early as January-February, while the tree pollen season peaks in April-May when birch trees flower. The main flowering period for grass is June-July, while weed pollen species flower throughout the summer. Wind-pollinated species are abundant pollen producers. For example, birch pollen concentrations may rise over 2,000 pollen grains per m^3 per day when weather conditions are favorable (Latalowa *et al.* 2002).

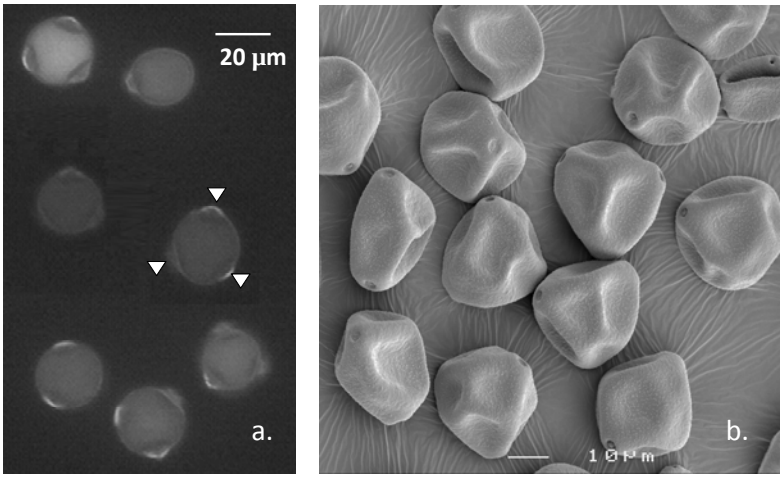


*Figure 1.2 – (a) Microscopic image of hydrated birch pollen (*B. pendula*) observed through a fluorescent filter. Autofluorescence was visible at the three germ pores, which are indicated by arrowheads. (b) Scanning electron microscope picture of hazel (*C. avellana*) pollen (picture by A. van Aelst, WUR).*

Birch pollen: exposure and sensitization

Birch pollen (Fig. 1.2a) is a major cause of hay fever in Europe. *B. pendula* is the most common birch species in Europe and has been widely investigated in relation to hay fever. However, other birch species occur in Europe in addition, namely *B. nana, B. pubescens* and *B. humilis*. *B. nana* has a circumpolar distribution and is primarily found at northern latitudes. *B. pubescens* and *B. pendula* cover large parts of the land and form stable climax forests in Scandinavia (Atkinson 1992). As a result, the birch pollen load is very high in Scandinavia (Rasmussen 2002). *B. pubescens* and *B. pendula* are also present in north-east Europe as is *B. humilis*. Pollen counts in this area are similar to those in Scandinavia (Latalowa *et al.* 2002). Towards the centre of their distribution range, *B. pendula* and *B. pubescens* are colonists in primary or secondary succession (Atkinson 1992). Pollen counts for several West European cities show that birch pollen is abundant and occurs in the same order of magnitude as grass pollen (Spieksma *et al.* 2003). Continuing southwards, birch becomes limited to cooler mountainous areas. Birch is present in the alpine foothills and around the north Italian lakes. Here, *B. pubescens* reaches the southernmost limit of its range, while *B. pendula* is also found in scattered populations throughout the rest of Italy (Atkinson 1992). Birch allergy is also reported in Eastern Asia, Japan, Canada and the USA, where

B. pendula does not occur (Abe *et al.* 1997; Eriksson *et al.* 1998). Other birch species do occur in these areas and are likely candidates for causing allergic complaints. Whether all birch species contribute to hay fever in the same extent is currently unknown. This justifies further investigation, because some *Betula* species may be less allergenic than *B. pendula*.

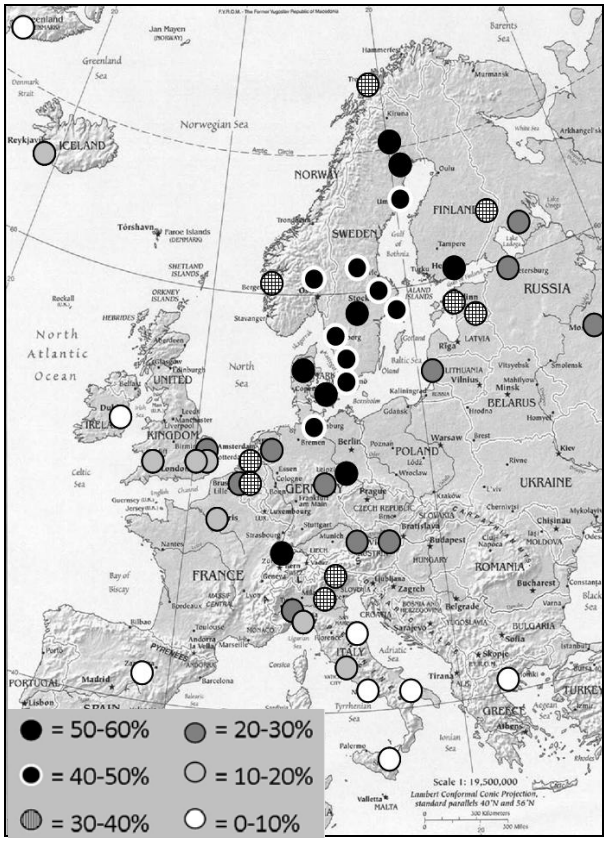


Figure 1.3 - The occurrence of sensitization to birch among atopic patients on several locations throughout Europe based on SPT and RAST data. Note that these studies are from different years and vary according to the age of the subjects, criteria for a positive test, testing method and patient group. Data were combined when multiple studies existed for the same location (Burney et al. 1997; Corsico et al. 2000; Crimi et al. 1999; Eriksson et al. 1998; Gioulekas et al. 2004; Gislason et al. 1999; Horak et al. 2002; Jõgi et al. 1998; Kilpeläinen et al. 2002; Kosunen et al. 2002; Krause et al. 2002; Linneberg et al. 2000; Lødrup Carlsen et al. 2002; Mari et al. 2003; Nowak et al. 1996; Plaschke et al. 2000; Raukas-Kivioja et al. 2003; Riedler et al. 2000; Rönmark et al. 2003; Silvestri et al. 1996; Sunyer et al. 2000; Vartiainen et al. 2002; Von Mutius et al. 1998).

The prevalence of sensitization to birch shows a distinct geographical pattern throughout Europe (Fig. 1.3). There is a good correlation between the degree of exposure throughout the main flowering season of birch and the occurrence of sensitization among atopic patients. Relatively high numbers of birch sensitized patients are found in areas with a strong presence of birch trees and high pollen counts. The 'hot-spot' of sensitization to birch lies in Scandinavia (Fig. 1.3) (Burney *et al.* 1997; Eriksson *et al.* 1998; Gislason *et al.* 1999), while the occurrence of sensitization to birch is also high among atopic patients in Estonia, Lithuania and the Russian Federation (Eriksson *et al.* 1998; Jõgi *et al.* 1998; Raukas-Kivioja *et al.* 2003; Vartiainen *et al.* 2002). In these areas, sensitization to grass occurs at a similar level to birch. The prevalence of sensitization to birch is also high throughout north-west Europe and the Alps (Fig. 1.3), although sensitization to grass has a higher prevalence throughout this region (Burney *et al.* 1997; Horak *et al.* 2002; Nowak *et al.*

1996; Riedler *et al.* 2000; Von Mutius *et al.* 1998). Where exposure levels are intermediate, such as in France or the UK, this leads to intermediate levels of sensitization to birch (Burney *et al.* 1997). Continuing southwards, birch is a relevant allergenic tree in the alpine and Po region in Italy, where over 30% of the atopic patients are sensitized to birch, as opposed to 2% in the most southern regions (Corsico *et al.* 2000). An 18%-20% share of sensitization to birch/hazel was recorded around Genoa (north Italy), while sensitization to olive pollen, grass and *Parietaria* was higher in this area (Crimi *et al.* 1999; Silvestri *et al.* 1996). Birch represents a minor source of allergens in Spain and Greece, compared to grass or olive pollen (Gioulekas *et al.* 2004; Sunyer *et al.* 2000).

Sensitization to birch pollen has increased in recent decades (Kosunen *et al.* 2002; Linneberg *et al.* 2000). It is not clear whether the share of birch pollen in the total allergic burden has changed over time. For example, an increasing prevalence of atopy in Denmark in the 1990s was not associated with a significant increase in the relative occurrence of sensitization to birch (Linneberg *et al.* 2000). A recent study by Stevens *et al.* (2003) have, however, found that sensitization to birch became more prevalent when compared to other allergens, whereas a Swiss study points to the increased prevalence of asthma induced by Fagales pollen relative to asthma induced by grass pollen (Frei *et al.* 2000). Increasing production of birch pollen due to climatic change may explain the increasing prevalence of sensitization. Higher temperatures favor a high pollen production, which, for example, can be observed in the pollen counts of Fagales species in Switzerland (Frei *et al.* 2000) and birch in Denmark (Rasmussen 2002). This trend is, however, not observed at every pollen counting station (Spieksma *et al.* 2003). A second effect of climatic change is a shift towards a slightly longer flowering period (Rasmussen 2002) and an earlier start of the flowering season (Clot 2001; Emberlin *et al.* 2002). This has caused hay fever symptoms to occur earlier during the year.

The major birch pollen allergen Bet v 1

Birch pollen contains several allergens, termed Bet v 1 to Bet v 8 (Table 1.1). There are geographic differences in the occurrence of sensitization to these allergens (Moverare *et al.* 2002). Bet v 1 is the major birch allergen throughout Europe. Over 95% of birch pollen allergic patients display IgE binding to Bet v 1 in Scandinavia, while 60% react exclusively to this allergen (Jarolim *et al.* 1989). Reactivity to birch profilin (Bet v 2) is low in Scandinavia, while becoming more common in central and south Europe. In north Italy, Bet v 1 is still the most common allergen to which sensitization has occurred (Moverare *et al.* 2002; Rossi *et al.* 2003), but in central Italy around Rome sensitization to Bet v 2 is more prevalent (Mari *et al.* 2003). Reactivity to Bet v 4 is low throughout Europe, except in Italy (Mari *et al.* 2003; Moverare *et al.* 2002). The reduced occurrence of sensitization to Bet v 1 is likely to reflect a reduced occurrence of birch trees, whereas the increase in sensitization to Bet v 2 could be due to sensitization to homologues from other sources such as

grass, ragweed and olive pollen. Sensitization to other birch allergens is less prominent, and only 10% of the birch allergic patients in Italy have specific IgE towards allergens other than Bet v 1, 2, and 4 (Rossi *et al.* 2003). Clearly, Bet v 1 is the most relevant target for prevention measures based on the number of patients that are affected.

TABLE 1.1 – Birch pollen allergens

Allergen	Size (kDa)	Type of protein	Uniprot No.	Described by
Bet v 1	17	PR-10 protein	P15494	(Breiteneder *et al.* 1989)
Bet v 2	14	Profilin	P25816	(Valenta *et al.* 1991)
Bet v 3	24	Ca2+ binding protein calmodulin	P43187	(Seiberler *et al.* 1994)
Bet v 4	9	Two EF-hand Ca2+ binding protein	Q39419	(Twardosz *et al.* 1997)
Bet v 6 (= 5)	35	Isoflavone reductase	U65002	(Karamloo *et al.* 2001b)
Bet v 7	18	Cyclophilin	P81531	(Cadot *et al.* 2000)
Bet v 8	65	Pectin esterase	-	(Mahler *et al.* 2001)

The major birch allergen Bet v 1 is a PR-10 protein. PR-10 proteins are reported as a multigene family across a range of phylogenetically distant plant species, including Gymnosperms, Monocots and Dicots (Ekramoddoullah *et al.* 2000; Gao *et al.* 2005a; Huang *et al.* 1997). The classification as PR-proteins (Van Loon and Van Strien, 1999) is based on the induced expression in response to pathogen infections by viruses, bacteria or fungi (McFadden *et al.* 2001; Poupard *et al.* 2003; Pühringer *et al.* 2000; Robert *et al.* 2001; Swoboda *et al.* 1995c), to wounding (Liu *et al.* 2003b; Poupard *et al.* 1998) or to abiotic stress (Moons *et al.* 1997; Srivastava *et al.* 2004; Utriainen *et al.* 1998). Other PR-10 family members are constitutively expressed during plant development (Pinto *et al.* 2005; Walter *et al.* 1996) or in specific tissues (Huang *et al.* 1997; Liu *et al.* 2003a). No general function has been described and PR-10 proteins are thought to act as ribonuclease (Bantignies *et al.* 2000; Moiseyev *et al.* 1997; Park *et al.* 2004), as cytokine-binding proteins (Fujimoto *et al.* 1998), as storage proteins (Flores *et al.* 2002; Richard-Molard *et al.* 2004), as plant steroid carriers (Markovic-Housley *et al.* 2003) or as cryoprotective proteins (Ukaji *et al.* 2004). PR-10 proteins are homogeneous both within the gene family in general, and even more so within species (Wen *et al.* 1997). Homogeneity is suggested to be maintained by concerted evolution. Arrangements of duplicated *PR-10* genes into clusters as found for Mal d 1 genes in apple (Gao *et al.* 2005a), may facilitate this evolutionary process.

The *B. pendula* genome contains multiple *PR-10* genes with varying expression patterns. mRNAs of PR-10 genes are detected in several tissues, including birch pollen (Breiteneder *et al.* 1989; Friedl-Hajek *et al.* 1999; Swoboda *et al.* 1995b), roots (Feugey *et al.* 1999; Koistinen *et al.* 2002; Poupard *et al.* 2001; Poupard *et al.* 1998; Utriainen *et al.* 1998), leaves (Utriainen *et al.* 1998; Valjakka *et al.* 1999) and cells that are grown in a liquid medium in the presence of microbial pathogens

(Swoboda *et al.* 1995c). The true Bet v 1 allergens are the PR-10 proteins that are expressed in pollen. The first Bet v 1 isoform was identified by immunoscreening of a cDNA expression library from pollen with serum of birch allergic patients (Breiteneder *et al.* 1989). Other Bet v 1 alleles have subsequently been sequenced by various authors (Friedl-Hajek *et al.* 1999; Hoffmann-Sommergruber *et al.* 1997; Son *et al.* 1999; Swoboda *et al.* 1995b). Bet v 1 is estimated to encompass 10% of the total protein content of birch pollen (Larsen 1995). A mixture of pollen from multiple birch trees was found to contain multiple Bet v 1 isoforms (Swoboda *et al.* 1995b), but whether multiple isoforms can be found within a single tree has not been investigated.

The number and types of isoforms that can be found within the pollen of a particular tree bear relevance to the allergenicity of that tree, because Bet v 1 isoforms differ in the ability to bind IgE. Based on immunoblot experiments, nine isoforms were grouped into classes of isoforms with high, intermediate and low IgE-reactivity (Ferreira *et al.* 1996). Isoforms from these three groups were tested in SPTs in which isoforms with a low IgE-reactivity (=hypoallergenic isoforms) induced a significantly lower reaction (Ferreira *et al.* 1996; Wagner *et al.* in press). The potency of these isoforms to activate T-cells from birch pollen allergic patients has also been determined. Isoforms that combine a low IgE-reactivity with a high T-cell proliferation score were identified (Ferreira *et al.* 1996). This bears particular relevance to Specific ImmunoTherapy (SIT), since such isoforms would reduce the risk for side-effects without harming the T-cell response. Interestingly, the three dimensional structure of the hypoallergenic isoform Bet v 1l was highly similar to the structure of isoforms with a high IgE-reactivity (Gajhede *et al.* 1996; Markovic-Housley *et al.* 2003). The differences between hypoallergenic and allergenic isoforms formed the basis for using site-directed mutagenesis in order to identify amino acid residues that influence the allergenicity. Several amino acid positions were proven to be crucial for IgE-reactivity (Ferreira *et al.* 1998; Son *et al.* 1999), but so far no IgE-binding epitopes has been identified. Spangfort *et al.* (2003) used murine monoclonal IgE to identify the structure of an allergen-antibody complex. The region containing a phosphate-binding loop with the sequence motive GxGGxGx covered most of the contact residues in this complex. This region is highly conserved throughout all PR-10 proteins. Two major T-cell activating regions that may influence cross-reactivity were also identified (Jahn-Schmid *et al.* 2005). Existing knowledge about the toxicity of Bet v 1 isoforms can also be used to predict the allergenicity of birch cultivars, varieties or species.

Cross-reactivity and oral allergy syndrome

Sensitization to birch pollen is accompanied by the occurrence of cross-reactivity and co-sensitization to pollen of other Fagales species (Corsico *et al.* 2000; Gioulekas *et al.* 2004; Mari *et al.* 2003; Von Mutius *et al.* 1998). The Fagales order includes wind-pollinating species such as hazel, oak, alder, hornbeam, beech and chestnut. Patients that are sensitized to one Fagales species may at the same time display sensitization to other species (Mari *et al.* 2003). The term

cross-reactivity is applied when the source of the sensitization can be identified. If not, the process is referred to as co-sensitization (Ferreira *et al.* 2004). Bet v 1 homologues have been identified in several Fagales species (Table 1.2). The high similarity of these proteins at the amino acid level provides a molecular basis for cross-reactivity (Niederberger *et al.* 1998). Cross-reactivity from birch to non-Fagales pollen also exists, but this involves sensitization to allergens other than Bet v 1 (Ferreira *et al.* 2004; Mari *et al.* 2003).

Sensitization to birch pollen is also linked to food allergy. Birch pollen-related food allergy is a frequent form of food allergy in Europe, and patients that suffer from this type of food allergy generally experience mild local allergic reactions upon consumption of allergenic foods (Ferreira *et al.* 2004; Ortolani *et al.* 1988). Symptoms of oral allergy syndrome (OAS) are mainly limited to the oral cavity and include an itchy or burning sensation of the palate, rarely combined with mucosal swelling or rhinitis symptoms. OAS is generally not life threatening. Other types of food allergy can also affect other organs, such as lungs (rhinitis, atopic asthma), the gastro-intestinal tract (nausea, vomiting and diarrhea), the skin (urticaria, allergic eczema) and the cardiovascular system (anaphylactic shock). The same food may sometime cause different response profiles as can for example be observed for peach, which may induce symptoms of OAS and anaphylactic shock, depending on the allergen involved (Fernández-Rivas *et al.* 2003; Fernández-Rivas *et al.* 1997; Van Ree *et al.* 1995). OAS and other food allergies affect a substantial part of the population. Prevalence estimates among the general community depend heavily on the diagnostic procedures involved (Mills *et al.* 2007; Rona *et al.* 2007) and range from 2 to 5% in SPT and IgE measurements, and from 1 to 11% in food challenge studies (Rona *et al.* 2007).

Up to 70% of the birch allergic sufferers may experience OAS (Ebner *et al.* 1991; Ortolani *et al.* 1988) and have an allergic reaction when consuming various fruits, nuts and vegetables (Bohle *et al.* 2003; Fritsch *et al.* 1998; Karlsson *et al.* 2004; Scheurer *et al.* 1999). The first evidence for cross-reactivity between birch pollen allergy and food allergy has been provided by RAST inhibition assays, which show that IgE-binding to an apple extract in serum from birch pollen allergic patients can be complete inhibited by incubation with a birch pollen extract (Lahti *et al.* 1980). *Vice versa* this is not the case, indicating that birch pollen contains a broader spectrum of allergens or epitopes than apple. The IgE-mediated cross-reactivity is caused by the presence of similar epitopes on Bet v 1 allergens and food proteins (Valenta *et al.* 1996; Wensing *et al.* 2002). Allergenic Bet v 1 homologues have been described for several foods (Table 1.2). For example, Bet v 1 and the homologous apple protein Mal d 1 have IgE epitopes in common (Ebner *et al.* 1991; Fritsch *et al.* 1998). In addition, Bet v 1 and Mal d 1 cross-react at the level of T_H-cells (Fritsch *et al.* 1998). Cross-reactivity has been clearly demonstrated for the celery allergen Api g 1 to which allergic responses are initiated by sensitization to Bet v 1 (Bohle *et al.* 2003). There is evidence suggesting that in some cases, Bet v 1 homologues may be involved in more serious allergic

reactions apart from OAS (Bolhaar *et al.* 2004b; Bolhaar *et al.* 2005b; Kleine-Tebbe *et al.* 2002; Mittag *et al.* 2004).

TABLE 1.2 – PR-10 homologues of Bet v 1 in pollen en food

Allergen	Source	GenBank No.	Described by
	Pollen		
Bet v 1	Birch (*Betula pendula*)	X15877	(Breiteneder *et al.* 1989)
Aln g 1	Alder (*Alnus glutinosa*)	S50892	(Breiteneder *et al.* 1992)
Cor a 1	Hazel (*Corylus avellana*)	X70999	(Breiteneder *et al.* 1993)
Car b 1	Hornbeam (*Carpinus betulus*)	X66932	(Larsen *et al.* 1992)
Cas s 1	Chestnut (*Castanea sativa*)	-	(Ipsen *et al.* 1991)
Que a 1	White oak (*Quercus alba*)	-	(Ipsen *et al.* 1991)
	Fruits		
Mal d 1	Apple (*Malus domestica*)	X83672	(Vanek-Krebitz *et al.* 1995)
Pyr c 1	Pear (*Pyrus communis*)	AF057030	(Karamloo *et al.* 2001a)
Pru av 1	Cherry (*Prunus avium*)	U66076	(Scheurer *et al.* 1999)
Pru ar 1	Apricot (*Prunus armeniaca*)	U93165	(Neudecker *et al.* 2001)
Pru p 1	Peach (*Prunus persica*)	DQ251187	-
Rub l 1	Red raspberry (*Rubus idaeus*)	DQ660361	-
Fra a 1	Strawberry (*Fragaria ananassa*)	AY679601	(Musidlowska-Persson *et al.* 2007)
	Vegetables		
Dau c 1	Carrot (*Daucus carota*)	Z81361	(Hoffmann-Sommergruber *et al.* 1999)
Api g 1	Celery (*Apium graveolens*)	Z48967	(Breiteneder *et al.* 1995)
	Nuts, legumes and seeds		
Cor a 1.04	Hazelnut (*Corylus avellana*)	AF136945	(Lüttkopf *et al.* 2002)
Gly m 4	Soybean (*Glycine max*)	X60043	(Kleine-Tebbe *et al.* 2002)
Ara h 8	Peanut (*Arachis hypogaea*)	AY328088	(Mittag *et al.* 2004)
Vig r 1	Mungbean (*Vigna radiata*)	AY792956	(Mittag *et al.* 2005)

Profilins form another example of cross-reactivity between pollen and food allergens. Profilins are actin-binding proteins that are found in grass, weed and tree pollen (Valenta *et al.* 1992). The birch allergen Bet v 2 is also a profilin. Pollen profilins cross-react to profilins in plant foods, such as tomato, hazelnut, peanut, banana and celery (Miralles *et al.* 2002; Westphal *et al.* 2004). Sensitization to profilin is observed in approximately 20% of the pollen-allergic individuals (Vieths *et al.* 2002). The clinical relevance of sensitization to profilins is under debate (Wensing *et al.* 2002). For example, sensitization to profilins in several foods (cherry, celery, pear) was clinically insignificant in a subgroup of patients with birch pollen allergy (Scheurer *et al.* 2001). Other examples of cross-reactive pollen allergens are a 60 KDa allergen from Mugwort (*Artemisia vulgaris*) which has homologues in apple, kiwi and celery (Heiss *et al.* 1996), and the Bet v 6

allergen which has homologues in various plant foods (Vieths *et al.* 2002). Further evidence is required to determine whether these cross-reactions are clinically significant. Non-specific Lipid Transfer Proteins (LTPs) are another important group of cross-reacting plant allergens, but the cross-reactions are limited to food alone in this case (Asero *et al.* 2007; Ballmer-Weber 2002; Sànchez-Monge *et al.* 1999).

Treatment and prevention strategies

What can be done to reduce or mitigate the occurrence of allergic complaints? There is a clear interest to investigate potential preventive measurements, because hay fever (or more generally, allergic rhinitis) has a negative impact on the patient's health-related quality of life. Fatigue, irritation, loss of sleep, and loss of concentration play a role, in addition to the obvious clinical symptoms (Meltzer 2001). This may lead to long term decreased cognitive functioning and learning impairment in children. Allergies are associated with considerable direct (medication and visiting a physician) and indirect costs (loss of school and working days). In Europe alone, the total annual costs for allergic rhinitis are estimated at 3 billion Euro in 1997 (UCB 1997). OAS has a negative impact on quality of life as well, but a lack of information has limited the quantification of the socio-economic impact for food allergies (Miles *et al.* 2005). Although a lower exposure to pollen may reduce sensitization (Kihlström *et al.* 2002), we focus on measures that can be implemented after sensitization has taken place. Reduction or mitigation of allergic complaints can be achieved by preventing the occurrence of symptoms or by treating the symptoms.

Treatment of hay fever focuses primarily on the mitigation of allergic symptoms through administration of medication. Treatment of symptoms is possible by the administration of anti-histamines, β_2-agonists, corticosteroids and mast cell stabilizing medication. The progression towards more severe hay fever symptoms and development of other allergic disorders may be avoided through proper medication. Long-lasting treatment can be achieved by SIT in some patients. The patient's sensitivity is then decreased by the repetitive and controlled administration of increasing doses of a specific allergen. SIT is directed towards the production of IgG antibodies which intercept allergens before they can bind to IgE on mast cells (Durham *et al.* 1998). At the same time, SIT stimulates IL-10 secreting $CD4^+$ regulatory T-cells which may actively control allergic responses. A downside of SIT is the risk for IgE-mediated side-effects. Allergens with a low IgE-reactivity and high T-cell proliferation score are considered suitable candidates to avoid such risks (Ferreira *et al.* 1996; Wagner *et al.* in press). Another possible option is the use of hypoallergenic fragments of allergens (Gafvelin *et al.* 2005). Whether SIT also reduces the cross-reactions to foods that are associated with pollen allergy is debated, but most recent studies indicated that this is indeed the case (Asero 1998; Bolhaar *et al.* 2004a; Kelso *et al.* 1995).

Hay fever symptoms may be mitigated by avoiding exposure to pollen, which can be achieved by staying indoors and keeping doors and windows closed during the flowering period of the pollen producing plants. Reliable information on the flowering of allergic plants by pollen forecasts is needed to develop effective communication about risk management strategies with patients. Avoidance requires active involvement of patients themselves, and is accompanied by social restrictions as allergy sufferers are restricted in their movements. It is important to note that, for example, during the flowering season of birch, peak pollen concentrations may occur at any time during the day and high concentrations often persist for several consecutive days (Clot 2001). Many patients are polysensitized (Crimi *et al.* 1999; Mari *et al.* 2003) and experience symptoms during the tree-, grass- and weed-pollen seasons, which together may last as much as 9 months a year. The impact of avoidance is also limited because pollen can easily penetrate buildings and may accumulate indoors (Fahlbusch *et al.* 2000).

Another option for mitigation is to reduce the concentration of allergenic pollen by removing particular pollen producing species from the urban environment. In order to do this, it is important to know the origin of the pollen that causes the complaints. This pollen may be of local origin or distributed by long-distance transport. Local flowering causes a strong increase of the pollen concentration in the local air (Clot 2001) and pollen counts are strongly influenced by the nearby presence of birch trees (Spieksma *et al.* 2003). A fraction of the birch pollen may be distributed by long-distance transport, which will influence exposure and allergic reactions as well (Koivikko *et al.* 1986; Skjøth *et al.* 2007; Sofiev *et al.* 2006). A negative side-effect of excluding species like birch from being planted, or removing existing trees, is that many tree species are considered an integral part of the landscape. Their absence can be perceived as an ‘unnatural’, or interpreted as representing an impoverished state of the landscape.

Dietary exclusion of allergenic foods and ingredients is still the primary preventive measure or treatment for food allergy (Ortolani *et al.* 1999). Dietary exclusion requires measures that facilitate consumer choice in the retail environment, such as communication and end-point labeling (Mills *et al.* 2004; Van Putten *et al.* 2006). Poor labeling as well as so-called precautionary (‘may contain’) labeling might lead to unnecessary restrictions in the diet of allergic individuals and their families (Hourihane 2001; McCabe *et al.* 2001). The allergenicity of food products can sometimes be reduced by food processing (Brenna *et al.* 2000; Gruber *et al.* 2004), but increased allergenicity may also result from heating (Bohle *et al.* 2006; Maleki *et al.* 2000). Full dietary exclusion is not necessary when the food is tolerated in its processed form. For example, most PR-10 allergens are disrupted by cooking. It is important to note that avoidance imposes restrictions on social activities, and may have a severe impact on the quality of life experienced by sufferers and their families (Fernandez-Rivas *et al.* 2004).

Preventive measures may focus on the development of hypoallergenic foods or hypoallergenic varieties of pollen producing plants or trees. Variation in the concentration of allergens has been found for foods such as apple, peach, nectarine and peanut, and for pollen from olive trees (Ahrazem *et al.* 2007; Castro *et al.* 2003; Koppelman *et al.* 2001; Marzban *et al.* 2005). Variation in the allergenicity among different protein variants involved in the allergic reaction (Ferreira *et al.* 1996; Wangorsch *et al.* 2007) may also prove to be a source of hypoallergenic products. Plant varieties with a reduced allergenicity can be selected within the natural range of variation. For example, high and low allergenic apple cultivars were identified using SPTs (Bolhaar *et al.* 2005a; Carnés *et al.* 2006). The apple cultivar Santana has been identified as hypoallergenic in these tests, and a clinical trial on this cultivar confirmed the results (Bolhaar *et al.* 2005a; Kootstra *et al.* 2008). Breeding for hypoallergenicity may further reduce the allergenic properties of apples.

The potential selection and breeding of hypoallergenic birch trees requires knowledge on the genetic background of the available birch species to cover the full extent of the variation in the *Betula* genus. The genus *Betula* encompasses 30-60 tree and shrub species. The number of recognized species and the taxonomy of *Betula* is controversial and various divisions into (sub)sections and subgenera have been proposed (De Jong 1993; Furlow 1990; Regel 1865; Winkler 1904). Breeding for hypoallergenic foods or trees is likely to be time-consuming and complicated. Both apple and birch have a generation span of about five years, and multiple generations will be required to develop suitable products. As an alternative, genetic modification (GM) can be applied to develop hypoallergenic foods. Gene silencing has been used to eliminate particular allergens. Examples are transgene-induced silencing of the Mal d 1 allergen in apples (Gilissen *et al.* 2005), Gly m Bd 30 in soybean (Herman *et al.* 2003), Lyc e 1 in tomato (Le *et al.* 2006), and Ara h 2 in peanut (Dodo *et al.* 2007). Ryegrass with knock-outs of the Lol p 5 allergen is an example of a hypoallergenic pollen-producing plant (Bhalla *et al.* 1999; Petrovska *et al.* 2005).

Societal Acceptance

Examples of previous technological innovations in the food chain, such as food irradiation and the introduction of first generation of GMO's in Europe, imply that success or failure of these innovations is contingent on societal acceptance of both products and processes. The incorporation of consumer preferences for product development and product characteristics in the early stages of the development process are critical factors for new product development (Van Kleef *et al.* 2005), which includes development of hypoallergenic food plants or trees. DNA research has been perceived by society as being an unknown, dreaded and uncontrollable risk (Slovic 1987), which is associated with low public acceptance of GM. Many debates about the implementation of GM were instigated from the early 1990s onwards, following the introduction of the first generation of GMO's. Public support for GM food in Europe has declined between 1996 and 1999, slightly increased between 1999 and 2002, and declined again between 2002 and 2005

(Gaskell *et al.* 2006). The attitude towards GM varies between European countries (Gaskell *et al.* 2006), as well as countries outside Europe (Ganiere *et al.* 2006; Ho *et al.* 2006; Nayga Jr *et al.* 2006).

Societal acceptance depends both on public attitude towards GM technology overall and towards specific applications of GM technology. The public is thus inclined to accept or reject different applications of GM on a case-to-case basis (Frewer *et al.* 1997). Case-specific characteristics that play a role are which organism is modified and for what purpose. Very generally, applications that involve plants or micro- organisms are more acceptable than those that involve animals (Frewer *et al.* 1997; Zechendorf 1994), although acceptance varies considerably within these categories. Medical applications tend to be considered more necessary, and thereby more acceptable, than food-related applications (Frewer *et al.* 1997). Applications of GM can be divided into three major categories. Most European consumers are positive about 'red' medical GM, reject the 'green' agricultural GM, and are indifferent to 'white' industrial GM. Demographic variables, such as age, gender, social class, education, and attitude towards the environment are known to influence the attitude towards GM (Siegrist, 1998; Titchener and Sapp, 2002). Apart from differences between categories of GM, consumer recognition of specific and personally relevant benefits, and the perception that consumers, rather than industry, will benefit from a particular application will also influence consumer acceptance of GM products (Miles *et al.* 2001).

Research aim and thesis outline

Birch pollen allergy and related food allergies affect a substantial part of the population. The main goal of this research project was to examine the feasibility of strategies that are directed towards development of hypoallergenic plants or foods in order to alleviate birch-pollen associated complaints.

Strategies focused on the development of hypoallergenic products may involve application of GM. In **Chapter 2**, we examine the attitude and preferences towards GM and non-GM strategies as expressed by the general public. Acceptance of GM is modeled in terms of perceived 'benefits' and perceived 'rejection factors'. Hay fever sufferers perceive greater 'benefits' associated with GM of birch compared to non-sufferers. No differences are observed for perceived 'rejection factors'. This underlines the importance of recognition of specific and personally relevant benefits for consumer acceptance of GM. This observation is validated in **Chapter 3**, in which the introduction of 'personal benefits' into GM foods is compared to the introduction of 'societal benefits'. Chapter 3 also examines the attitude of allergic consumers towards hypoallergenic foods, based on a sales pilot of the apple cultivar Santana. The cultivar Santana is the first example of a hypoallergenic fruit that is marketed as such.

Both previous chapters indicate a societal preference for application of traditional breeding opposed to application of GM. To facilitate strategies for selection and breeding of hypoallergenic birch trees, knowledge on the diversity of Bet v 1 genes and allergenicity of the gene products is required. **Chapter 4** describes a study on the variation of Bet v 1 isoforms in the most common birch species in Europe, *B. pendula*. *PR-10* sequences from three *B. pendula* cultivars are amplified, cloned and sequenced to establish the number of *PR-10/Bet v 1* genes and isoform diversity within a single tree. Multiple Bet v 1 genes are found in *B. pendula* and although we establish some genetic diversity, the sequence of one of the isoforms with a high IgE-reactivity is also present. Therefore, our search is expanded towards other *Betula* species. To select a set of species that would cover the variation within the entire *Betula* genus, knowledge on the phylogenetic relationships is required. In **Chapter 5,** we reconstruct the phylogeny of *Betula* using AFLP markers and identify four subgroups in the genus. **Chapter 6** describes the cloning and sequencing of *PR-10* alleles from eight birch species that cover these four subgroups. Alleles of PR-10 genes are found in all species. Q-TOF LC-MS^E is subsequently applied to identify the pollen-expressed Bet v 1 genes and to determine relative abundances of individual isoforms.

Chapter 7 addresses variation in allergenicity among several birch species. These species are tested in an SDS-PAGE and immunoblot assay, and the amount of Bet v 1 protein is quantified in a fluorescent array using polyclonal antibodies. The abundance of amino acid residues that are associated with high or low IgE-reactivity is determined by applying Q-TOF LC-MS^E. **Chapter 8** is the general discussion in which the results of this thesis are discussed and integrated.

CHAPTER 2

The influence of perceived benefits on acceptance of GM applications for allergy prevention

Martijn Schenk, Arnout Fischer, Lynn Frewer, Luud Gilissen,
Evert Jacobsen and René Smulders

Published as: Schenk MF, Fischer ARF, Frewer L, Gilissen LJWJ, Jacobsen E & Smulders MJM (2008) The influence of perceived benefits on acceptance of GM applications for allergy prevention. Health Risk and Society, *in press*.

Abstract

Allergic diseases, such as hay fever and food allergy, affect an increasing part of the population in westernized countries and have a negative impact on the patient's quality of life. Allergy prevention measures that focus on reducing the allergenic load are currently developed, and these may include the use of genetic modification of allergenic plants. Such developments should take societal concerns about genetic modification into account. We examined the attitude of allergic and non-allergic respondents towards applications of genetic modification for allergy prevention in one food allergy application (apple) and two hay fever applications (birch, grass). Attitude towards genetic modification was described in terms of 'benefits' and 'rejection factors'. We found that respondents suffering from self-reported allergy perceived greater 'benefits' associated with the birch application as compared to non-sufferers. The perceived 'benefits' increased with an increasing impact of allergic complaints on quality of life. No differences were found between sufferers and non-sufferers for the food allergy application. The impact of perceived 'benefits' on acceptance was larger than that of 'rejection factors'. This supports the idea that acceptance of genetic modification is primarily a function of perceived personal benefit. Novel genetically modified products that are perceived to be beneficial by some consumers may consequently experience an increased consumer acceptance.

Introduction

Allergy and societal issues

Allergic diseases affect a substantial proportion of the population in westernized countries (Aberg *et al.* 1995; Strachan 1989; Wuthrich *et al.* 1995). Evidence suggests that the prevalence of allergy has increased during recent decades. Seasonal allergic rhinitis (hay fever) can be caused by various types of pollen. It is the most common allergic disease and affects between 11% and 19% of the population in Western European countries (UCB 1997). It may act as a precursor to allergic asthma (Meltzer 2001). In addition, sufferers that have been sensitized to birch pollen are particularly prone to develop oral allergy syndrome due to IgE-mediated cross-reactivity between pollen allergens and food proteins (Ferreira *et al.* 2004; Wensing *et al.* 2002). Approximately 50-70% of the birch allergic sufferers display symptoms upon the ingestion of fruits and vegetables, including apple, cherry, celery, and carrot (Bohle *et al.* 2003; Ferreira *et al.* 2004; Fritsch *et al.* 1998).

Sufferers experience symptoms such as sneezing, rhinorrhea, and itchy eyes upon exposure to allergenic pollen. Hay fever is accompanied by fatigue, irritated glands, and loss of sleep and of concentration (Meltzer 2001). Oral allergy syndrome leads to an itchy or burning sensation of the palate, rarely combined with mucosal swelling, angioedema, nausea, or rhinitis symptoms. The oral allergy syndrome is generally not life threatening, but does have a negative impact upon quality of life. Quality of life assessments can be used to quantify the impact of hay fever and oral allergy syndrome on daily life from the patient's perspective, and to determine the need for social, emotional, and physical support (Sicherer *et al.* 2001). Annual costs for allergic rhinitis add up to an estimated 3 billion euros in Europe (UCB 1997). A lack of information has limited the quantification of the socio-economic impact for food allergies (Miles *et al.* 2005). Development of effective allergy prevention strategies is likely to have a positive impact on socio-economic factors for both individual sufferers and society in general.

Prevention of hay fever and food allergy

Treatment of hay fever focuses mainly on the mitigation of allergic symptoms through administration of medication (UCB 1997). Long-term treatment for various pollen allergies may be achieved by immunotherapy for some sufferers. The burden for sufferers is relatively high, given that immunotherapy takes a long period and that its eventual success is uncertain. Allergy prevention comprises any intervention that leads to a reduced sensitization to allergens (primary prevention) or to a reduced occurrence of allergic complaints (secondary prevention). Sufferers who have developed hay fever may limit the occurrence of symptoms by avoiding contact with allergenic pollen, for example by minimizing contact with outdoor air during periods with high pollen concentrations.

The primary preventive measure used for oral allergy syndrome is avoidance, which can be achieved by dietary exclusion of allergenic foods and, if necessary, any products derived from it. Full dietary exclusion is not necessary when the food is tolerated in a processed form. Dietary exclusion requires measures that facilitate consumer choice in the retail environment, such as communication and end-point labelling (Mills *et al.* 2004; Van Putten *et al.* 2006). Poor labelling as well as so-called precautionary labelling might lead to unnecessary restrictions in the diet of allergic individuals and their families (Hourihane 2001; McCabe *et al.* 2001). Therefore, avoidance of food allergens imposes restrictions on social activities and potentially has a severe impact on the life of sufferers and that of their family (Fernandez-Rivas *et al.* 2004).

Alternative prevention measures may focus on plants or on products that provoke allergic reactions. In the case of hay fever, varieties with reduced allergen content of the pollen or a reduced amount of pollen may be selected within the natural range of variation. Likewise, apple varieties differ in allergenicity and several cultivars contain low levels of allergens (Bolhaar *et al.* 2005a). Selective breeding may further enhance such characteristics. Selection for such traits is however time-consuming and complicated. Currently, safety assessments are an important tool to ensure the safety of novel genetically modified crops (Taylor 2003). Assessments take into account the unintended introduction of allergens in these crops (Goodman *et al.* 2005), but genetic modification may actually aid in the development of hypoallergenic plants or foods from existing varieties. Examples are hypoallergenic ryegrass (Bhalla *et al.* 1999; Petrovska *et al.* 2005), apple (Gilissen *et al.* 2005), or soybean (Herman *et al.* 2003). Whichever approach is used to reduce allergy problems, the examples of previous technological innovations in the food chain, such as food irradiation and first generation GMO's in Europe, imply that its success or failure is contingent on societal acceptance of both products and processes.

Societal issues in relation to genetic modification

Societal concerns about aspects connected to genetic modification should be taken into account when replacing allergenic products with hypoallergenic counterparts using genetic modification. Such considerations are most useful at the initiation of product development. Public perceptions and attitudes towards new technologies are important factors for successful development and implementation of such technologies. DNA research has generally been perceived as an unknown and moderately dreaded hazard in the past (Slovic 1987), resulting in adverse reactions from the public towards genetic modification. Many debates about the implementation of genetic modification were instigated from the early 1990s onwards, following the introduction of first generation applications of genetic modification. Public support for genetically modified food in Europe has declined between 1996 and 1999, slightly increased between 1999 and 2002, and declined again between 2002 and 2005, although the attitude varied among European countries (Gaskell *et al.* 2006). On the other hand, support for medical applications of modern biotechnology

remains high. Societal acceptance likely depends on both overall public attitudes towards genetic modification technology and on public responses to specific applications. As a result, the public is inclined to accept or reject different applications of this technology on a case-by-case basis (Frewer *et al.* 1997).

Very generally, genetic modification of plants or microorganisms is more acceptable than genetic modification of animals (Frewer *et al.* 1997; Zechendorf 1994), although acceptance also varies considerably within these categories. Applications of genetic modification can be divided into three major categories. The majority of the European consumers are positive about 'red' medical genetic modification, reject the 'green' agricultural genetic modification, and are indifferent to 'white' industrial genetic modification. Medical applications tend to be considered more necessary, and thereby more acceptable, than food-related applications (Frewer *et al.* 1997; Zechendorf 1994). Evidence suggests that consumer acceptance of genetically modified foods is driven by recognition of specific and personally relevant benefits, and by the perception that consumers, rather than industry, will benefit from a particular application (Miles *et al.* 2001). Demographic variables, such as age, gender, social class, education, and attitudes towards the environment are also known to influence the attitude towards genetic modification (Siegrist 1998; Titchener *et al.* 2002).

Research aims

The research presented here examines and models the attitudes of consumers towards the use of genetic modification for allergy prevention. Two hay fever applications (birch, grass) and one food allergy application (apple) were included in the study. A survey methodology was applied to data collection. Given that people's attitude towards genetic modification partly depend on perceptions of personal risk and benefit; it is hypothesized that genetically modified products that result in a reduced occurrence of allergic reactions will be more positively received by allergic sufferers. In turn, acceptance is expected to increase with an increasing severity of allergic complaints experienced by individual sufferers.

All examined applications of genetic modification can be considered medical in nature due to their relation with allergy prevention. In addition, the hay fever applications may be associated with environmental concerns due to the distribution of pollen into the environment, while the apple application is associated with food-related concerns. Given the strong impact of food-related concerns on consumer perceptions (Frewer 2000), we expect the apple application to be rated more negatively than the applications in birch and grass.

Methods

Main survey

The survey investigated attitudes towards the use of genetic modification for allergy prevention for three applications, namely apple, birch and grass. Grass and birch are the main provokers of hay fever in the Netherlands. At the same time, they represent pollen sources of different size and distribution pattern. Apple allergy is connected to birch allergy due to cross-reactions between the major birch allergens Bet v 1 and Bet v 2 and the homologues proteins Mal d 1 and Mal d 4 in apple (Wensing *et al.* 2002). Apple is a major allergenic food. It generally causes relatively mild allergic symptoms, which makes it more comparable to hay fever than, for example, a life-threatening peanut allergy would be. One exception is apple allergy caused by a lipid transfer protein (Mal d 3), which may result in severe symptoms (Pastorello et al., 1999). In recent years, research has been conducted towards development of genetically modified hypoallergenic apples (Gilissen *et al.* 2005), hypoallergenic rye grass (Bhalla *et al.* 1999; Petrovska *et al.* 2005), and birch trees that do not flower (Lemmetyinen *et al.* 2004).

The main survey was carried out in 2005. Increased publicity about allergies is likely during the hay fever period, and hay fever sufferers are actually experiencing symptoms at that time. As the timing of the survey might influence the response, the sample was split into two. Two hundred and eighty two respondents received the questionnaire *prior* to the birch hay fever season (mid-February) and 250 respondents *during* the birch hay fever season (first week of May). The timing of the survey was included as an independent variable in the multivariate analysis.

Data were collected using quota sampling in order to obtain a representative sample. The quota variables were age, gender, and region of the Netherlands where the respondent lived. A professional social research company conducted the survey through an online questionnaire. Respondents were assigned to one of three quota blocks: (A) 178 respondents for the birch and grass application, (B) 179 respondents for the birch and apple application, and (C) 175 respondents for the grass and apple application. No significant differences were observed for demographic and allergy characteristics between the three quota groups (A, B and C). The order of the items was randomized within each block of questions.

Pilot study and survey design

A pilot survey was conducted, comprising two phases. The first phase was used to reduce the number of items and to test their comprehensibility. Survey items were primarily developed from existing literature (see survey design). The pilot questionnaire was made available online, and completed by a convenience sample of 72 Dutch respondents. Items that provided no additional information were removed to reduce questionnaire length. For pragmatic considerations regarding respondent fatigue, each respondent was asked to answer items about two out of three

cases. This further reduced the number of items from 119 to 94 per respondent. For the second phase of the pilot, a social research company selected a sample of 60 Dutch respondents and conducted the pilot through an online questionnaire. In this phase, findings from the first phase were confirmed for the more representative sample and open issues were resolved. Based on these results, an anchored seven-point scale was chosen and the option '*don't know*' was removed. No further adaptations were made.

Attitude items

The survey items regarding attitude towards genetic modification were derived from previous research on acceptance of specific applications of genetic modification (Frewer *et al.* 1997). The original scale included 17 items. Given the health-related nature of our applications, two items on health effects were added. These were '*please indicate to what extent you think that the use of genetic modification will have negative health effects*' and '*please indicate to what extent you think that the use of genetic modification will have positive health effects*'. Principal components analysis was applied to identify underlying factors in the data set, reflecting sets of items that measured the same attitudinal construct. The set of 19 items was reduced to 14 (see Table 2.3) after application of this analysis in the pilot study. Responses were collected on seven-point-scales anchored by '*not at all'* and '*to a large extent*'. The three applications were introduced by a brief explanation, ensuring that all respondents were answering questions after being given the same amount of information about hay fever and food allergy.

Prevention strategy items

Five prevention strategies which could potentially mitigate the occurrence of symptoms due to birch pollen allergy in the urban environment were presented. The strategies might also add to primary prevention, but we will focus on the effects of secondary prevention because of the distinction between sufferers and non-sufferers made throughout this paper. The first strategy was directed towards allergy sufferers and focused on prevention of birch pollen allergy through '*avoiding contact with pollen*'. Three strategies were directed towards replacement of allergenic trees with either '*other tree species*', '*other birch varieties that have been selected from existing varieties for a reduced allergenicity*' or '*genetically modified birch trees that have been altered to have a reduced allergenicity*'. To allow a comparison with the current situation, '*maintaining the current situation*' was included as a fifth option. The same strategies were presented for grass. The first prevention strategy with regard to apple consumption focused on *'avoiding contact with apples*'. The other strategies were directed towards replacement of allergenic apples with either '*other fruits*', or '*other apples that have been selected from existing varieties for their reduced allergenicity*', or '*genetically modified apples that have been altered to have a reduced allergenicity*'. The fifth option was '*maintaining the current situation*'. Respondents were asked to indicate to what extent they found it desirable to implement a particular strategy. Responses were

collected on seven-point-scales anchored by '*highly undesirable*' and '*highly desirable*'. The desirability of implementing the strategy that involved genetic modification was used to measure acceptance of genetic modification for allergy prevention and was used as a dependent variable in the multivariate analysis.

Allergy and demographic items

Respondents were asked to indicate whether they suffered from any allergic disorders. For allergic respondents, information was collected regarding the type of allergic disorder (allergic rhinitis, food allergy, allergic eczema, allergic asthma, and insect sting, medication, and metal allergy), the (recent) use of medication, allergy diagnosis, timing of the symptoms, and substances to which they were allergic. Items on quality of life were included for hay fever sufferers to assess the severity of their complaints. The fourteen quality of life items were developed and validated in previous research on adults suffering from rhino-conjunctivitis (Juniper *et al.* 1991; Juniper *et al.* 2000). Currently, there is no validated method available to measure quality of life for food allergy sufferers. The frequent occurrence of hay fever and food allergy in the Netherlands allowed us to use a random population sample to include a sufficiently large group of allergy sufferers in the study.

Selected demographic characteristics were recorded (age, gender, household composition, education level, income, urbanisation, and employment). Four items measured the self-reported overall health of the respondents. Items that were intended to explain individual differences in general concern over the environment were included. Responses were collected on anchored seven-point-scales. Acceptance of genetically modified hypoallergenic products was modelled in terms of the attitudinal components derived from the survey on attitude towards specific applications of genetic modification, and in terms of the demographic characteristics included in the survey.

Results

Sample demographics

The demographic characteristics of the respondent sample are provided in Table 2.1. There was an equal distribution of missing demographic information between sufferers and non-sufferers. No differences were found between sufferers and non-sufferers with respect to education (χ^2=3.37, *df* 2, *p*=0.19), income (χ^2=6.21, *df* 4, *p*=0.18), and household composition (χ^2=3.52, *df* 4, *p*=0.48). However, the distribution of sufferers differed significantly with respect to age and gender. The average age of sufferers was 41.9 (SD=12.2), which was significantly lower than the average age of non-sufferers (45.7, SD=12.6) (*F*(1, 530)=12.28, *p*<0.001). This is in concordance with allergy prevalence data from cross-sectional studies for the adult population, which have reported a decreasing prevalence of allergic sensitization with increasing age (UCB 1997). The prevalence of

self-reported allergic complaints was higher among women (χ^2=6.80, *df* 1, *p*<0.01) of whom 54% reported allergic complaints, while 41% of the males reported these. Women are known to report higher allergy prevalence when self-reported allergic diseases are concerned (Fagan *et al.* 2001; Knibb *et al.* 1999; Marklund *et al.* 2004). Women often perceive higher risks than men do (Finucane *et al.* 2000), which may partially account for this difference in health reporting behaviour.

TABLE 2.1 – Demographic description of the study population.

Characteristic		Number of respondents
N		532
Gender	Male	266 (50%)
	Female	266 (50%)
Age	Mean (*SD*)	43.9 (12.6)
House hold composition	Single	76 (17%)
	Single with kids	19 (4%)
	Partner	134 (30%)
	Partner with kids	184 (41%)
	Other	40 (9%)
Education	6-10 years	125 (30%)
	11-14 years	176 (42%)
	15-17 years	121 (29%)
Income[1]	Below modal	72 (19%)
	About modal	109 (29%)
	About 1.5 times modal	93 (25%)
	About 2 times modal	59 (16%)
	About 2.5 times modal or more	40 (11%)

1 Modal house hold income in the Netherlands was about €29.000 in 2005

The study population was also characterized according to the allergy background of the respondents (Table 2.2). Forty seven percent of the respondents indicated that they suffered from allergic complaints. This seems higher than the average of 35% of the population commonly reported in research using diagnostic tools to detect elevated levels of specific IgE, such as skin-prick tests or blood tests (UCB 1997). This can partly be explained by allergy sufferers being more willing to respond to a questionnaire about allergies, causing a self-selection bias. In addition, self-report data tend to indicate higher prevalence estimates than those based on diagnosis by health professionals, due to an imperfect distinction between atopic and non-atopic symptoms

(Marklund *et al.* 2004). Please note that in the remainder of this paper, the distinction between allergy sufferers and non-sufferers is based on self-reported allergy. This may lead, of course, to misclassification of individuals who suffer from food intolerances or true (non-allergic) rhinitis symptoms as sufferers. However, as perceived allergy is likely to influence attitude independent on whether a formal allergy diagnosis has been made, this misclassification is unlikely to influence whether or not individuals accept products with a reduced allergenicity.

TABLE 2.2 – Self-reported allergy background of the respondents according to the type of allergy and the allergens to which sufferers are sensitized.

	Number of respondents
Allergy (*N*=532)	
Non-allergic	284 (53%)
Allergic	248 (47%)
Type of allergy among allergy sufferers (*n*=248)	
Only allergic rhinitis	110 (44%)
Only food allergy	21(8%)
Both allergic rhinitis and food allergy	39 (16%)
Other allergies	78 (31%)
Allergen sensitization among allergy sufferers (*n*=248)	
Only pollen allergens	95 (38%)
Only fruit/nut allergens	9 (4%)
Both pollen and fruit/nut allergens	27 (11%)
Other allergens	57 (23%)
Unknown allergens	60 (24%)

The majority of self-reported allergic complaints concerned allergic rhinitis (70%). Food allergy was reported frequently in combination with allergic rhinitis (Table 2.2). Allergy sufferers can also be characterized according to the allergens which they perceive to induce an allergic reaction. We distinguished between pollen, fruit and/or nuts, other, and unknown allergens. Given the large number of sufferers that were pollen allergic without knowing the exact source, and the large overlap between both groups (52% of the sufferers allergic to pollen were allergic to both grass and tree pollen), no distinction was made between tree and grass pollen allergies. Overall, the largest group of respondents were allergic to pollen allergens (49%), while a smaller but still substantial percentage (24%) did not know the source of their allergy. There was a large overlap between pollen and fruit/nut allergic groups (17%), which is expected given the occurrence of cross-reactions between pollen allergens and food allergens.

Attitude subscales

A principal component analysis was applied to determine the number of dimensions of the attitude items. The responses towards the item 'personal objections' did not show a normal distribution. This item was therefore excluded from the analysis. The analysis (Varimax rotated) was first performed separately for the three applications (birch, grass, and apple). The results indicated a two-factor solution for each application, based on Eigenvalues larger than one. The Eigenvalues of the first two components were 6.24 and 1.74 for the birch application. The next component had an Eigenvalue of 0.91. Results were similar for the grass application (Eigenvalues of 6.00, 1.77, and 0.89), and the apple application (Eigenvalues of 6.89, 1.73, and 0.98). The two factors explained a comparable amount of variance (cumulative 61% for birch, 65% for grass, and 66% for apple), while items loaded onto the components in a similar fashion across each application. The combined dataset also indicated a two-factor solution, with the first two components having Eigenvalues of 6.53 and 1.77, while the third component had an Eigenvalue of 0.92. The rotated components respectively accounted for 40%, and 24% of the variance. The dimensions, composed by the items that loaded onto each component, were treated as subscales (Table 2.3). The ambivalent item 'long-term effects' did not load well on either component, and was excluded from further analysis.

TABLE 2.3 – Principal Component Analysis loadings. Bold numbers indicate which variables load on a principal component.[1]

	Component	
	Rejection factors	Benefits
Risky	**0.82**	-0.20
Damaging	**0.80**	-0.24
Personal worries	**0.80**	-0.18
Tampering with nature	**0.80**	-0.27
Unethical	**0.79**	-0.18
Unnatural	**0.79**	-0.16
Negative health effects	**0.76**	-0.26
Long-term effects[2]	0.43	0.42
Progressive	0.03	**0.59**
Positive health effects	-0.40	**0.66**
Beneficial	-0.42	**0.76**
Necessary	-0.34	**0.78**
Important	-0.39	**0.79**

1 the item 'personal objections' was removed because the responses deviated from a normal distribution

2 The ambiguous item 'long term effects' contributed to neither factors and was excluded from further analysis

The internal consistency of the remaining subscales was estimated using Cronbach's alpha. In other words, do all the items contribute to the measurement of the component under

consideration? Alpha was respectively 0.92 and 0.85 for the two components, indicating good internal reliability. However, there is some criticism regarding Cronbach's Alpha as a definitive test for construct reliability as it is, for example, vulnerable in the situation of colinearity. Confirmatory factor analysis (using LISREL) was applied to determine whether the number of factors and the loadings of measured (indicator) variables associated with them were in concordance with the original principal components analysis. The resulting model shows reasonable fit parameters ($\chi^2(53)=357$, RMSEA=0.075, GFI = 0.95, AGFI = 0.92, and CFI = 0.98). Fit measures for a structural equation model: RMSEA<0.08 is an indicator of an acceptable fit. GFI, AGFI, CFI>0.90 are indicators of an acceptable fit; GFI, AGFI, CFI>0.95 are indicators of a good fit. This confirms the reliability of the constructs identified in the original principal components analysis. Based on the content of the items, the two sub-scales were interpreted as 'rejection factors' and 'benefits', and were used as dependent variables in further analyses.

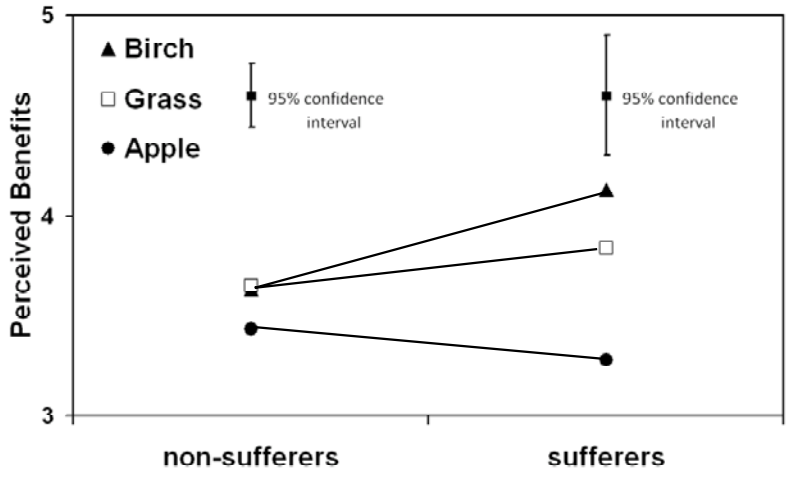


Figure 2.1 - Mean item scores and 95% confidence interval of the attitude component 'benefits' separated for the three applications, and for allergic sufferers and non-sufferers. Allergic sufferers are defined as respondents with a self-diagnosed pollen and/or food allergy.

Differences between food allergy and hay fever applications

The survey covered genetic modification of two hay fever and one food allergy application. A straightforward comparison of the 'application' effect was not possible due to the balanced incomplete block design. We adjusted for block effects by calculating the average score for each quota group and subtracting this score from the individual responses. Sufferers were defined here as being food allergic and/or pollen allergic, while non-sufferers were defined as being neither food allergic nor pollen allergic. This definition allowed a simultaneous comparison of 'sufferer' and 'application' effects. A multivariate analysis of variance (MANOVA) was applied in order to assess group differences across the different dependent variables simultaneously. Specifically, differences between attitudes towards 'benefits' and 'rejection factors' were tested between the different applications as well as between allergy suffers and non-allergy suffers (Pillai's trace $F(4,2116)=5.99$, $p<0.001$). The univariate tests revealed that all differences accrued to the

'benefits' (F(5,1058)=5.93, p<0.001), and not to the 'rejection factors' (F(5,1061)=2.09, p=0.06). The subsequent ANOVA on the 'benefits' was significant for the 'application' variable (F(2,1061)=21.81, p<0.001), the 'sufferer' variable (F(1,1062)=7.30, p<0.05), and there was a significant *'application' x 'sufferer'* interaction (F(2,1061)=7.38, p<0.05). A *post hoc* Tukey test (p<0.05) revealed that the hay fever applications were perceived significantly more beneficial than the apple application, but were not different from each other. The interaction effect shows that the increased perception of benefits for these applications is fully attributable to the larger perceived benefits by allergy sufferers, since no differences between applications were found for non-sufferers (Fig. 2.1).

Attitude and acceptance of allergy sufferers and non-sufferers

The benefits of having genetically modified hypoallergenic products available are not equal for all members of the population. Only people who have an allergic reaction to particular product variants experience a direct benefit from the hypoallergenic counterpart. The distinction between sufferers and non-sufferers was, therefore, case-dependent. Sufferers were categorised as 'pollen allergic respondents' for the hay fever applications (n=122) and as 'food allergic respondents' for the apple application (n=60). Forty-three food-allergic respondents were part of the quota blocks that included items on apple. Given the small group size of food allergic respondents, we randomly selected 43 respondents from the remaining respondents when testing for attitude differences between sufferers and non-sufferers in relation to the apple application.

The attitude towards genetic modification for allergy prevention was influenced by whether the respondent suffered from allergy. There was a significant multivariate effect (MANOVA) of self-reported pollen/food allergy on the variables 'benefits' and 'rejection factors' for the birch application (Pillai's trace F(2,354)=9.46, p<0.001). However, such differences were neither found for grass (Pillai's trace F(2,350)=1.15, p=0.32) nor for apple (Pillai's trace F(2,83)=0.43, p=0.65). A subsequent ANOVA for the birch application revealed a significant difference between sufferers and non-sufferers for the 'benefits' (F(1,355)=15.95, p<0.001), but not for the 'rejection factors' (F(1,355)=0.76, p=0.38). Pollen allergic sufferers perceived greater 'benefits' of the use of genetic modification than non-sufferers (Table 2.4).

We also investigated the attitude of allergic respondents that were neither pollen nor food allergic. No significant attitude differences were found between this group and the non-allergic respondents with regard to the proposed genetically modified plants (birch: Pillai's trace F(2,262)=0.65, p=0.65; grass: Pillai's trace grass F(2,276)=0.84, p=0.43; apple: Pillai's trace F(2,308)=0.17, p=0.85). This confirms the expectation that direct personal benefits play a significant role in acceptance. Therefore, the classification of sufferers as respondents with a self-diagnosed pollen allergy or food allergy was maintained throughout this paper.

TABLE 2.4 – Mean item scores (standard deviation)[1] of the attitude components and prevention strategy ratings separated for allergic sufferers and non-sufferers. The applications are tested separately because the patient group is case-dependent; sufferers are defined as pollen allergic for the hay fever applications, and as food allergic for the apple application.

	Birch		Grass		Apple	
	Non-sufferers n=265	Sufferers n=92	Non-sufferers n=279	Sufferers n=74	Non-sufferers n=311	Sufferers n=43
Attitude characteristics						
Rejection factors	4.17 (1.47)	4.01 (1.62)	4.30 (1.54)	4.12 (1.63)	4.38 (1.55)	4.62 (1.78)
Benefits	**3.61** (1.30)	**4.26** (1.44)	3.64 (1.34)	3.91 (1.49)	3.41 (1.30)	3.27 (1.27)
Prevention strategy[2]						
Avoid contact	**4.78** (1.49) [a]	**5.22** (1.49) [a]	4.67 (1.57) [a]	4.81 (1.59) [a]	4.62 (1.87) [a]	4.95 (1.90) [a]
Replacement with other products	**4.82** (1.51) [a]	**5.37** (1.28) [a]	4.54 (1.54) [a]	4.78 (1.50) [a]	4.81 (1.87) [a]	5.09 (2.01) [a]
Replacement with classical breeding	4.83 (1.42) [a]	5.09 (1.44) [a]	4.78 (1.52) [a]	4.92 (1.44) [a]	4.72 (1.60) [a]	5.09 (1.62) [a]
Replacement with genetic modification	**3.36** (1.73) [b]	**4.29** (1.80) [b]	3.42 (1.78) [c]	3.69 (1.80) [b]	3.11 (1.72) [c]	2.72 (1.67) [c]
Maintaining the current situation	**3.67** (1.47) [b]	**3.23** (1.58) [c]	3.82 (1.56) [b]	3.50 (1.62) [b]	4.11 (1.46) [b]	3.84 (1.62) [b]

1 High numbers indicate high ratings on attitude subscales and a high desirability of implementing the prevention strategy. Numbers in Bold indicate significant differences between sufferers and non-sufferers.

2 A one-way ANOVA with a post-hoc Tukey test was applied to assess significance between the prevention strategies. Prevention strategies with the same subscript were not significantly different at the $p<0.05$ level over other strategies in the same column.

A MANOVA was used to examine differences in acceptance of prevention strategies between sufferers and non-sufferers. There was a significant multivariate effect of self-reported pollen allergy for the birch application (Pillai's trace $F(5,351)=6.06$, $p<0.001$). Significant differences between sufferers and non-sufferers were found for the strategies *avoidance of contact with birch pollen* ($F(1,355)=5.75$, $p<0.05$), *replacement with other tree species* ($F(1,355)=9.97$, $p<0.01$), *replacement with genetically modified birch trees* ($F(1,355)=19.55$, $p<0.001$) and *maintaining the current situation* ($F(1,355)=5.87$, $p<0.05$). Inspection of the means indicated that sufferers rated the use of *genetic modification, avoidance of contact* and *replacement with other tree species* as more desirable than non-sufferers, while rating *maintenance of the current situation* as less desirable (Table 2.4). The finding that sufferers rated the strategy that involves genetic modification higher than non-sufferers is in concordance with the attitudinal differences. We found no significant multivariate effect of self-reported pollen/food allergy for the prevention strategies concerning apple (Pillai's trace $F(5,348)=1.70$, $p=0.14$) and grass (Pillai's trace $F(5,347)=0.71$, $p=0.62$).

Attitude and severity of hay fever complaints

This analysis indicates that suffering from a pollen allergy influences attitude towards genetic modification for allergy prevention, and acceptance of the prevention strategies. To investigate additional influences of the severity of allergic complaints on attitude and acceptance, we divided the sufferers using a median split on their quality of life scores for rhino-conjunctivitis (Juniper *et al.* 1991; Juniper *et al.* 2000). Subsequently, a one-way ANOVA was used to examine differences between sufferers with a low and with a high impact of hay fever on quality of life (Table 2.5). In the birch case, we found significant differences between sufferers with a high and with a low impact of hay fever for the attitude component 'benefits' ($F(1,80)=4.34$, $p<0.05$). In concordance with this finding, significant differences were found for implementation of the strategies *replacement with birch trees developed by genetic modification* ($F(1,80)=5.89$, $p<0.05$) and *replacement with birch varieties developed by classical breeding* ($F(1,80)=4.39$, $p<0.05$). No significant differences were found for the other strategies. Inspection of the means indicated that sufferers with a high impact of hay fever on quality of life perceive higher 'benefits' and a higher desirability of implementing these prevention strategies (Table 2.5). , We also found significant differences for the grass application in the ratings of the strategies *replacement with other grass species* ($F(1,62)=11.25$, $p<0.001$), *replacement with grass varieties developed by classical breeding* ($F(1,62)=20.37$, $p<0.001$), and *maintaining of the current situation* ($F(1,62)=5.63$, $p<0.05$). No significant differences were found for the other strategies or the attitude components. A high impact of hay fever on quality of life lead to higher ratings of the replacement strategies and lower ratings for maintaining of the current situation lower (Table 2.5). A similar comparison could not be made for apple, since there is currently no standardized method available to measure the impact of oral allergy syndrome on quality of life.

TABLE 2.5 - Mean item scores (standard deviation)[1] of the attitude components and prevention strategy ratings for the hay fever applications. Median split was used to separate between hay fever sufferers with a low impact and a high impact of hay fever on quality of life scores.

Impact on quality of life	Birch		Grass	
	Low	High	Low	High
Attitude characteristics				
Rejection factors	3.98 (1.55)	4.11 (1.72)	4.03 (1.61)	4.23 (1.80)
Benefits	**3.84 (1.48)**	**4.52 (1.43)**	3.76 (1.48)	4.26 (1.55)
Prevention strategy				
Avoid contact	4.83 (1.54)	5.37 (1.45)	4.59 (1.70)	5.16 (1.38)
Replacement with other products	5.00 (1.17)	5.48 (1.33)	**4.36 (1.50)**	**5.56 (1.23)**
Replacement with classical breeding	**4.61 (1.59)**	**5.28 (1.31)**	**4.41 (1.37)**	**5.84 (0.99)**
Replacement with genetic modification	**3.67 (1.71)**	**4.63 (1.84)**	3.41 (1.68)	4.04 (2.01)
Maintaining the current situation	3.44 (1.46)	3.09 (1.62)	**3.87 (1.74)**	**2.88 (1.45)**

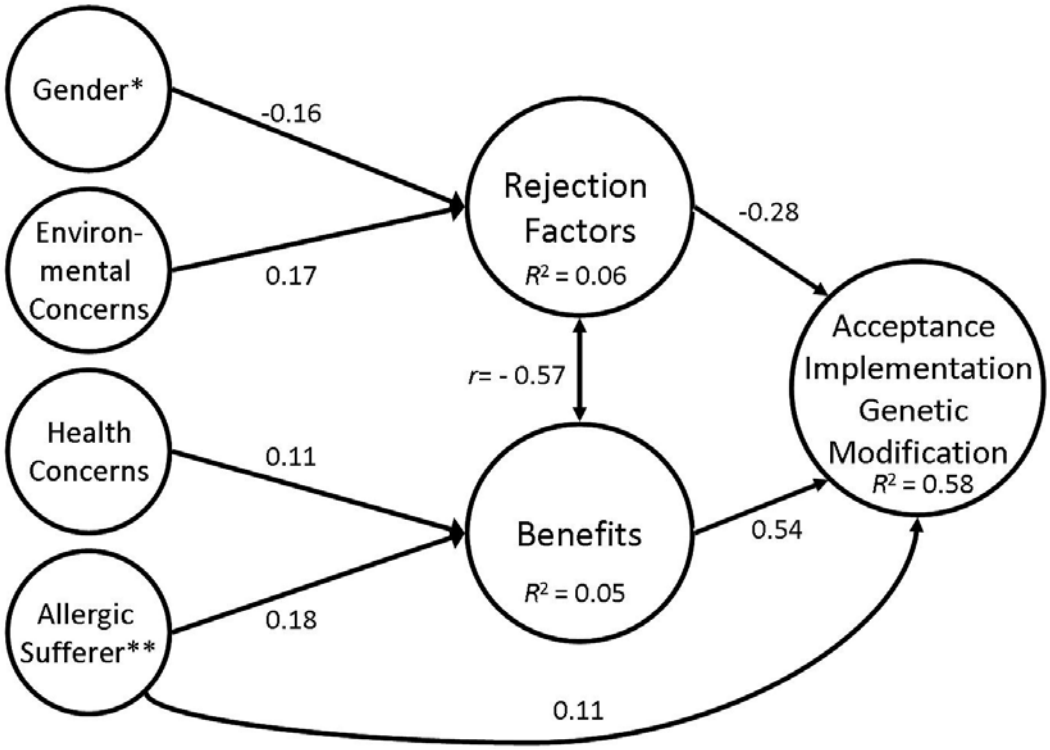
1 High numbers indicate high ratings on attitude subscales and high desirability of implementing the prevention strategy. A one-way ANOVA was applied to assess significance between the groups. Numbers in 'Bold italics' indicate significant differences between 'high and low impact on quality of life' groups at the p<0.05 level.

Self-reported allergy and other demographic variables

To investigate the relevance of the above findings, we performed multiple regression analysis to test simultaneously the effects of several demographic variables (gender, age, worries about the environment, general health, timing of the survey, urbanisation, income, and self-reported allergy), and the effect of the attitudinal components (benefits, rejection factors) on the acceptance of genetically modified hypoallergenic products. The results for the birch application are shown in Fig. 2.2. The attitudinal components: 'rejection factors' and 'benefits' were negatively correlated (r=-0.57, p<0.001). This means that the components were not independent of each other. We compensated for this effect using partial correlations when investigating the relation between attitudinal components and the significant demographic variables, such that the relation between 'worries about the environment' and 'benefits' (r=0.00, p=0.98) was controlled for in the analysis. Thus, it can be concluded that the influence of environmental worries on benefit perception is fully mediated by perceived rejection factors. The variables 'worries about the environment' and gender had significant influences on the attitudinal component 'rejection factors' (F(2,354)=11.38, p<0.001, R^2=0.06). Respondents who indicated more worries about the environment perceived more 'rejection factors' associated with genetic modification (F(1,355)=12.74, p<0.001, R^2=0.03). Female respondents perceived more 'rejection factors' than male respondents did (F(1,355)=11.75, p<0.001, R^2=0.03). The variables 'self-reported pollen allergy' and general health had a direct significant influence on the 'benefits' component

($F(2,354)=10.00$, $p<0.001$, $R^2=0.05$). Respondents who self-reported pollen allergy perceived more benefits associated with genetic modification of birch than those who were not ($F(1,355)=15.95$, $p<0.001$, $R^2=0.04$). In addition, respondents who indicated a lower general health perceived more benefits ($F(1,355)=14.64$, $p<0.01$, $R^2=0.02$).

*Figure 2.2 - Multiple regression analysis for the use of genetic modification in birch for hay fever prevention. The acceptance of genetic modification is the ultimate dependent variable. The attitudinal components 'rejection factors' and 'benefits' were used as mediating variables. Demographic variables were used as independent variables. Significant (*p*<0.05) relations are shown.*



* Male was coded as 1 and female as 0

** Pollen allergic sufferers were coded as 1 and non-sufferers as 0

The implementation of genetic modification for birch was used as dependent variable, while the measured attitudinal components (benefits and rejection factors) were used as independent variables. This analysis indicated that both 'benefits' ($F(1,355)=384.16$, $p<0.001$, $R^2=0.52$) and 'rejection factors' ($F(1,355)=189.97$, $p<0.001$, $R^2=0.35$) had a significant (non-orthogonal) influence on acceptance. There was an additional effect of the demographic variable 'self-reported pollen allergy' on acceptance of genetic modification ($p<0.01$, additional $R^2=0.01$). These non-orthogonal effects added up to a total explained variance of 58%. Mediation analysis showed that the other

demographic effects were fully mediated by either 'benefits' or 'rejection factors', with self-reported pollen allergy being only partially mediated by 'benefits' (Fig. 2.2).

A similar pattern of attitudinal variation for 'rejection factors' emerged for the apple application (F(2,351)=13.86, p<0.001, R^2=0.07) and in grass (F(1,351)=93.67, p<0.001, R^2=0.11). The exception was the variable gender, which did not have a significant influence for the grass application. There were considerable differences for the 'benefits'. For grass, this variable was only slightly influenced by the health variable (F(1,346)=4.52, p<0.05, R^2=0.01), while no significant predictors for 'benefits' were found for apple. There was also no unmediated effect of the variable self-reported food allergy on acceptance of genetic modification. The 'benefits' and 'rejection factors' together did explain a similar amount of variation for all three applications (birch F(3,353)=161.69, p<0.001, R^2=0.58; apple F(2,351)=226.51, p<0.001, R^2=0.56; grass F(2,350)=269.64, p<0.001, R^2=0.60).

Implementation of prevention strategies

To examine which strategies to prevent allergic symptoms are preferred by the respondents, we compared the ratings on desirability of implementing the different strategies. Given the observed differences in attitude between sufferers and non-sufferers, these groups were analysed separately by application of a one-way ANOVA. The prevention strategies had different levels of acceptance for the apple application (sufferers F(4,210)=15.05, p<0.001; non-sufferers F(4,1550)=52.42, p<0.001), birch (sufferers F(4,430)=28.35, p<0.001; non-sufferers F(4,1345)=60.94, p<0.001), and grass (sufferers F(4,335)=13.24, p<0.001; non-sufferers F(4,1420)=39.32, p<0.001). A Post-hoc Tukey test was conducted to examine which strategies differed from each other (Table 2.4). For all three applications, the strategies that focus on *avoidance of contact*, *replacement with other products,* and *replacement with products developed by classical breeding* scored significantly higher than *replacement with products developed by genetic modification* and *maintaining the current situation* (Table 2.4). Both sufferers and non-sufferers rated the use of genetic modification in apple (much) lower than maintaining the current situation. Sufferers rated genetic modification as more acceptable than maintaining the current situation for birch, while these strategies were rated equally acceptable by non-sufferers. Sufferers rated genetic modification of grass equally acceptable to maintaining the current situation, while non-sufferers were more negative about genetic modification (Table 2.4).

Discussion

Attitude towards genetic modification

The results in this study provided partial support for the hypothesis that allergy sufferers perceive genetically modified products that are hypoallergenic more positively than non-allergic people are. To find this confirmation, the attitude towards the use of genetic modification was measured on two subscales that were labelled as 'rejection factors' and 'benefits'. Hay fever sufferers perceived greater 'benefits' associated with genetically modified birch compared to non-sufferers. No differences were observed between allergic and non-allergic groups in terms of perceived 'rejection factors'. Although there was a tendency for sufferers to perceive greater 'benefits' than non-sufferers for genetically modified grass, these differences were not significant. The differences are perhaps less pronounced for grass because this application is perceived as less realistic, at least in the Dutch context where grass is indigenous. This maybe different in, for example, Australia, where the development of transgenic hypoallergenic ryegrass is being evaluated (Petrovska *et al.* 2005). Replacement of birch in the urban environment by other trees or hypoallergenic birch trees certainly appears more feasible. Examining consumer attitudes in a country in which grass is not omnipresent in the environment may clarify this issue.

Besides a generic difference between sufferers and non-suffers of allergic complaints, it is further shown that the more severe the self-reported allergic complaints, the larger the perceived benefits and acceptance of genetically modified birch. As there is currently no method to assess quality of life in relation to the severity of apple allergy these relations could only be investigated for hay fever (i.e. birch and grass).

Consumers perceive fewer 'benefits' associated with the apple application, supporting the hypothesis that food allergy applications of genetic modification are rated more negatively than hay fever applications. This provides support for the idea that food related concerns are larger than many others are. More generally, this research confirms that the attitude towards use of genetic modification depends on the nature of the application (see also: Frewer *et al*. 1997; Zechendorf 1994). Interestingly, all observed differences related to the perceived personal 'benefits'. All three applications are partly medical in nature. In contrast, the apple application also relates to agriculture and food consumption, whereas the hay fever applications relate to environmental issues. This suggests that the apple application follows the broader pattern of societal resistance against food-related GMOs, whereas grass and birch are more in line with societal acceptable medical applications. It should also be noted that the food allergic group was more diverse in its food allergies compared to the hay fever patients, which may reduce the actually perceived personal benefit when specifically focusing on the example of apple. Given the sample, there were not enough respondents allergic to apple to conduct a separate analysis for this group. The lack of personal benefit for people with, for example, shrimp allergy, regarding

their acceptance of genetically modified apples is obviously small, which may reduce acceptability of genetically modified apple In future research we should take care that participants are the recipients of an actual and personally relevant benefit.

Rejection factors, benefit and acceptance of GMOs

Acceptance of genetically modified hypoallergenic products was modelled in terms of the two attitudinal components 'rejection factors' and 'benefits', and the demographic characteristics. Frewer *et al*. (1997) also found a two-dimensional attitudinal structure towards genetic modification. Both attitudinal components are also similar to the main variables in the model of gene technology acceptance proposed by Siegrist (2000), who indicated that the variables 'perceived benefit' and 'perceived risk' were influential in determining consumer acceptance. The perceived 'benefits' of applying genetic modification to birch were mainly influenced by the perception that an individual is suffering from allergy, while for both hay fever applications it was also influenced by a person's perception of his/her health status. The strongest indicator of the 'rejection factors' was a high level of environmental concern. This is in accordance with previous research by Siegrist (1998) who found that high levels of concern about the environment were one of the main predictors of general rejection of gene technology.

'Benefits' and 'rejection factors' have a strong impact on acceptance of specific applications of genetic modification. We observed that differences in attitude both between sufferers and non-sufferers and between different applications all relate to the perception of greater 'benefits' but not to fewer 'rejection factors'. This supports the idea that acceptance of genetically modified products is primarily a function of perceived personal benefit as opposed to personal or environmental risk perceptions *per se*. This is also indicated by the higher regression coefficient between perceived 'benefits' and 'acceptance of genetic modification', compared to the coefficient between 'rejection factors' and 'acceptance of genetic modification' in Fig. 2.2. Frewer (2003) concluded that as long as the risks are not as large as to be completely intolerable, an individual's acceptance would be driven by perceptions of personal benefit.

A limitation of the study is that, although a fairly large fraction of the variance of acceptance of genetic modification could be explained based on general attitudes towards genetic modification, only a modest amount of this attitude can be explained from demographic and allergy characteristics. Future research should aim at identifying the variables that underlie the remaining fraction of the variance.

Food allergy management and genetically modified products

Van Putten *et al*. (2006) concluded that, in the context of novel foods, genetically modified products have the potential to contribute to food allergy management. However, the results of

this study show a clear consumer preference for hypoallergenic apples grown by conventional breeding opposed to breeding by using genetic modification. Miles *et al*. (Miles *et al.* 2005) found that a majority of food allergy sufferers would purchase low-allergen foods produced by genetic modification, but sufferers also expressed a clear preference for low-allergen foods produced by conventional breeding. We observed no differences between the attitudes of food allergy sufferers and non-sufferers towards genetically modified hypoallergenic apples. This may relate to the fact that apple allergy is relatively easy to manage by apple allergy sufferers themselves. The consequences of inadvertent exposure to apples are generally rather minor, and thus tolerable. The response might be different for food products that are harder to avoid or that provoke severe (anaphylactic) allergic reactions, such as peanut or peach. A second explanation might be that people suffering from food allergy or intolerances are more worried about technological food safety issues (Mills *et al.* 2004). A third explanation might be that the introduction of low-allergen products may complicate food allergy management for allergic consumers, who will not be able to discriminate between allergenic and non-allergenic products based on simple heuristic cues (do not eat apples) and on sensory selection criteria like smell or appearance. This would make avoidance of potentially allergenic foods more difficult, requiring an increased level of vigilance to the part of food allergic sufferers (Gowland 2001). The development of specialist niche markets, with explicit labelling protocols for low allergen products may partially solve these problems.

Another difficulty is that hypoallergenic products that are non-allergenic for some sufferers might still be allergenic to others. Most allergenic foods contain a variety of food allergens, each with a different number of patients reacting to them. In a cost-effective development of hypoallergenic products, the focus will likely be on the major allergens. This implies that minor allergens will probably not be reduced in this process. Sufferers will also have individual threshold levels to particular allergens. Allergen reduction levels that suffice for one patient will not necessarily suffice for others, complicating matters further. As a result, some sufferers may still suffer from allergic complaints upon confrontation with acclaimed hypoallergenic foods. To avoid this, sufferers need detailed knowledge about which allergens affect them, and whether the hypoallergenic food still contain these specific allergens, rather than the more general knowledge about which products to avoid. Inadequate addressing of these issues may reduce trust in hypoallergenic foods.

A clear societal preference is shown for prevention strategies involving conventional breeding opposed to those involving genetic modification. In fact, the majority of the respondents prefer maintaining the current situation to the use of genetic modification for the apple and grass applications. The situation is more complicated for birch, since pollen allergic sufferers included in this study prefer the implementation of genetic modification to maintaining the current situation. Nineteen percent of the respondents included in the study preferred reduction of birch pollen

allergens through genetic modification of birch when compared to conventional breeding approaches. In comparison, 47% of respondents preferred conventional breeding to genetic modification when applied to the objectives. Future research on hypoallergenic birch should, therefore, first explore the use of conventional breeding.

There is a trade-off between what the general public demands and what is currently feasible in terms of plant breeding perspectives. The use of conventional breeding techniques would be challenging and may require several decades to yield results, while the expression of the major apple allergen and ryegrass allergens have already been successfully reduced in transgenic plants by RNA interference or gene-silencing (Bhalla *et al.* 1999; Gilissen *et al.* 2005; Petrovska *et al.* 2005). The duration of the developing process was not, as such, included in this survey, but this factor raises the intriguing issue of what people would prefer, a GM solution now or a non-GM solution in several years or decades? Another issue to consider in future research is the attitude of the general public towards recent developments in the research field of plant breeding, such as cisgenic crops, in which no genetic material from foreign species is introduced during transformation (Schouten *et al.* 2006).

Conclusion

The attitude towards genetically modified products that have a reduced allergenicity compared to their conventional counterparts can be modelled in terms of 'benefits' and 'rejection factors'. The research reported shows that hay fever sufferers show higher acceptance levels towards hypoallergenic GM products compared to non-allergy sufferers. No differences were found in rejection factors. In other words, it is the perceived personal benefits associated with such products that account for the difference in an acceptance of novel applications between allergic and non-allergic individuals. Furthermore, the greater the perceived benefit the stronger is its positive effect on acceptance of the application. In the past, the emphasis of discussion about societal acceptance of different applications of genetic modification, whether health-related or otherwise, has tended to focus on the relation between perceived risk and acceptance. The research presented here underlines the importance of perceived personal benefits for consumer acceptance, independent of risk perceptions.

In the context of the introduction of other innovative technologies applied to health, it is possible that the effect reported here may have generic applicability – that is, acceptance or rejection of different technological applications may be primarily dependent on benefit perceptions, rather than on perceptions of risk. This topic merits further research in the future.

CHAPTER 3

Hypoallergenic food products as a novel approach to alleviate mild food allergy

Martijn Schenk, Rien van der Maas, Rene Smulders,
Luud Gilissen, Arnout Fischer, Ivo van der Lans and Lynn Frewer

To be submitted

Abstract

The development of hypoallergenic foods may contribute to food allergy management. The Santana apple is an example of a hypoallergenic food that is suitable for most of the apple allergic patients suffering from mild food allergy. As an alternative, genetically modified hypoallergenic apples are being developed. The present study examines whether the Santana and other (genetically modified) hypoallergenic foods are perceived as beneficial by allergic consumers. A large-scale sales pilot was conducted with the Santana apple. A survey among consumers that bought the Santana measured the self-reported response to the Santana and to a set of apple profiles that varied in the applied breeding method and in allergenicity. The self-reported severity of the apple allergy, the occurrence of other fruit allergies, and age were associated with the occurrence of an allergic reaction to the Santana. Forty-two percent of the apple allergic consumers had no allergic reaction after eating the Santana. Most (96%) consumers who did experience an allergic reaction reported the symptoms as minor. Overall, the Santana was perceived to be beneficial by the majority of apple allergic consumers. The attitude towards similar hypoallergenic products was positive, also when genetic modification was involved in their development. Attitude towards the Santana apple and similar hypoallergenic food products was positive among allergic consumers, because they experience a "personal benefit" associated with such products. Variation among allergic consumers in the response to hypoallergenic products should be taken into account when developing and marketing such products.

Introduction

Food allergies affect a substantial proportion of the population. Prevalence estimates range from 1 to 11% of the population suffering from food allergy complaints (Rona *et al.* 2007). These estimates depend heavily on the diagnostic procedures involved (Mills *et al.* 2007). Dietary exclusion of allergenic foods and ingredients and, if necessary, derived products represents the primary strategy to deal with food allergies (Ortolani *et al.* 1999; Zeiger 2003). Dietary exclusion requires measures that facilitate consumer choice in the retail environment, such as communication and end-point labeling (Cornelisse-Vermaat *et al.* 2007; Mills *et al.* 2004; Van Putten *et al.* 2006). Both poor labeling as well as precautionary ("may contain") labeling might lead to unnecessary restrictions in the diet of allergic individuals and their families (Hourihane 2001; McCabe *et al.* 2001).

Avoidance of allergenic foods imposes restrictions on social activities and may have a severe impact on the life of sufferers and that of their families (Fernandez-Rivas *et al.* 2004). Some of the restrictions may be relieved by the development of hypoallergenic foods, which are foods that are less likely to cause allergic complaints than their traditional counterparts. However, the introduction of hypoallergenic foods may have complications associated with risk management. For example, having to separate allergenic foods and hypoallergenic counterparts requires an increased vigilance from allergic consumers (Gowland 2001). Furthermore, individual consumers may vary with regard to threshold levels and may have different clinical response profiles because they react to different allergens in the same food. As a result, some food allergy sufferers may still experience an allergic reaction upon the consumption of hypoallergenic products.

Different strategies can be applied to develop hypoallergenic foods. The allergenicity of products can be lowered by food processing (Brenna *et al.* 2000; Penás *et al.* 2006; Primavesi *et al.* 2006), although in some cases food processing increases the potential for allergic responses (Bohle *et al.* 2006; Maleki *et al.* 2000). For unprocessed food products, variation in the concentration of allergens (Ahrazem *et al.* 2007; Marzban *et al.* 2005) or the presence of protein variants with a reduced allergenicity (Ferreira *et al.* 1996; Wangorsch *et al.* 2007) may result in hypoallergenic foods. Genetic Modification (GM) may aid in the development of such hypoallergenic foods by application of gene silencing and the consequent elimination of particular allergens. Examples are transgene-induced silencing of Gly m Bd 30 in soy (Herman *et al.* 2003), of Lyc e 1 in tomato (Le *et al.* 2006) and of Ara h 2 in peanut (Dodo *et al.* 2007).

Apple is widely consumed in many parts of the world, and frequently causes allergic reactions in northern Europe. The majority of apple allergic patients experience relatively mild symptoms that characterize Oral Allergy Syndrome (OAS). OAS is caused by an IgE-mediated cross-reactivity between pollen allergens and food proteins (Bohle *et al.* 2003; Fritsch *et al.* 1998; Lüttkopf *et al.*

2002; Wiche *et al.* 2005). OAS in apple is closely linked to birch pollen allergy due to cross-reactions between the birch allergens Bet v 1 and Bet v 2, and their apple homologues Mal d 1 and Mal d 4 (Wensing *et al.* 2002). In southern Europe, a relatively rare type of apple allergy is caused by lipid transfer proteins and may evoke severe symptoms (Pastorello *et al.* 1999). Apple varieties vary in the extent to which they cause allergic reactions (Bolhaar *et al.* 2005a; Carnés *et al.* 2006). The Santana apple has been identified as having a reduced allergenicity by skin prick tests (Bolhaar *et al.* 2005a). The Santana was estimated to alleviate allergic complaints in over 50% of the apple allergic patients and is considered suitable for patients with a mild apple allergy based on a food challenge study (Kootstra *et al.* 2008). Selective breeding may further enhance the hypoallergenic characteristics in progeny of the Santana apple. As an alternative, GM hypoallergenic apples are being developed in which Mal d 1 has been silenced (Gilissen *et al.* 2005).

Incorporating the 'voice of the consumer' into the development process is a critical factor for new products, such as hypoallergenic foods (Van Kleef *et al.* 2005). Consumers in Europe have strong societal concerns about GM foods (Gaskell *et al.* 2006) and these should be taken into account when GM is applied to develop novel hypoallergenic foods. Evidence suggests that consumer acceptance of GM is driven by consumer recognition of specific and personally relevant benefits (Schenk *et al.* in press), and by the perception that consumers, rather than industry, will benefit from a particular application (Miles *et al.* 2001). The attitude towards GM is influenced by case-specific characteristics of the application, such as which type of organism is modified and to what purpose (Frewer *et al.* 1997; Zechendorf 1994). The attitude of allergic and non-allergic consumers towards application of GM for development of hypoallergenic plant foods is to be evaluated.

The present study examines whether hypoallergenic foods are perceived to be beneficial by allergic consumers, by using a sales pilot of the Santana apple as a case study. In this pilot the Santana was sold with the attribute 'hypoallergenic'. We examined whether the occurrence of an allergic reaction to the Santana can be predicted and how its absence or presence influences consumer attitude towards hypoallergenic foods in general. The perceptions of benefits are further explored in a controlled study using "hypothetical apple profiles", in which the attitude towards hypoallergenic foods is examined both for apple developed by GM and by traditional breeding methods. In a wider context, this study examines whether consumer acceptance of GM foods is driven by recognition of specific and personally relevant benefits, as opposed to benefits to society in a more general way.

Methods

Data collection

The survey was carried out in two phases from September 2006 to June 2007. From September 2006 to April 2007, the Santana apple was sold by a large supermarket chain, in several organic

food stores and at several greengrocers in the Netherlands. In the supermarket, the apple was packaged with an explanatory leaflet, which contained information about the hypoallergenic properties of the Santana and apple allergy. The leaflet provided information about an eating protocol for allergic patients, and a request was made to participate in an internet-based survey about the Santana. The vast majority of consumers who responded to this request were suffering from apple allergy. In order to create a large enough control group of non-allergic consumers, further participants were recruited by convenience sampling from an existing panel in the second phase (May-June 2007). The survey took about 8 minutes to complete, and respondents could participate in a lottery for a gift voucher of 50 euro.

Survey design: Santana

The survey included items about location of purchase of the Santana apple, about occurrence of an allergic response to the Santana, and about use of medication after consumption. Respondents were asked to rate the Santana on a seven-point scale anchored by "1. do not like at all" to "7. like very much". Information was collected regarding the self-reported occurrence of allergic disorders, allergy diagnosis, and substances to which allergic respondents reacted (apple, other fruits, other foods, pollen and other substances). Respondents with an apple allergy were asked to report the severity of their allergic reaction to apples in general.

TABLE 3.1 - Conjoint study of apples: attributes and levels

Attribute	Number of levels	Level description
Breeding method	3	- Traditional breeding - Genetic modification using material from other plant species - Genetic modification using material from other apple varieties
Pesticide usage	2	- 5% pesticide reduction during growing - 50% pesticide reduction during growing
Apple allergy	3	- 5% of apple allergic patients without complaints - 66% of apple allergic patients without complaints - 95% of apple allergic patients without complaints

Survey design: apple profiles

Respondents were asked to rate apples with alternative levels of three attributes that were varied in the general description of the apple (Table 3.1). Apples with a combination of these attributes exist, are currently being developed, or can potentially be developed in the future. We will refer to the combinations of alternative levels from the three attributes as "apple profiles" throughout the remainder of this paper. Pesticide and allergy reduction levels were developed in collaboration with plant breeding experts in order to reflect realistically attainable reduction levels. A 5% reduction in pesticide use was used in the descriptions instead of 0% to maintain ecological

validity of the apple descriptors. With respect to breeding method, a distinction was made between GM utilizing material from other plant species and from other apple varieties. This reflects recent developments associated with “cisgenes”, or genes from the crop plants themselves or from crossable species (Jacobsen *et al.* 2007). The apple profile descriptions were tested for comprehension and understanding in a pilot survey that was conducted in 2006. The pilot survey was completed by a convenience sample of 15 Dutch respondents. No adaptations were made on the basis of the pilot results.

A full-factorial within person-design would require the respondents to evaluate a set of (3 x 2 x 3=) 18 apple profiles. Previous research has indicated that respondents have difficulties in rating such a large number of profiles (Jan *et al.* 2007). Therefore, a fractional-factorial design was used in order to reduce the number of profiles presented to each respondent. The design was created with the conjoint module in SPSS 15.0. Nine apple profiles were created and attributed to three clusters (1 to 3) of three profiles each. Respondents were randomly assigned to groups A to C. Each group evaluated two clusters of profiles in a balanced incomplete block design (A: 1+2; B: 1+3; C 2+3). No significant differences were observed for demographic and allergy characteristics between the three groups. Each profile was described on a separate page and respondents evaluated a warm-up apple profile at first. Respondents rated each profile on three items. The first item measured to what extent respondents liked the apple under consideration on a seven-point scale from “1. do not like at all” to “7. like very much”. The second item measured willingness to buy by asking for the agreement with the statement “if this apple is sold in the supermarket, I would buy this apple” on a seven-point scale from “1. completely disagree” to “7. completely agree”. The third item measured what price respondents would be prepared to pay for the apple.

In addition, we included items that were derived from previous research on acceptance of GM (Frewer *et al.* 1997; Schenk *et al.* in press) to allow an interpretation of individual differences in acceptance or rejection of hypoallergenic GM apples in terms of attitudinal variables. Two items measured the self-reported overall health of the respondents. Two items measured the general attitude towards application of GM in food. One item measured the concern over pollution of the environment and one item the preference for organic food products opposed to traditionally grown products. Responses to the health and attitude items were collected on anchored seven-point scales.

Finally, selected demographic characteristics were recorded (age, gender, and education level). Respondents were divided into six age groups (10-20, 21-30, 31-40, 41-50, 51-60 and >61 years).

Statistical analysis

All statistical procedures were performed in SPSS 15.0. The ratings on the apple profiles were subjected to a repeated measures mixed linear model. This procedure can be applied to analyze conjoint data from an incomplete block design (Maas *et al.* 2003). Models were fitted using the residual maximum likelihood (REML) algorithm. Initially, a step down test was performed (significance level at p=0.05) using the deviance values for the covariance matrix to arrive at a well-fitting covariance structure (Maas *et al.* 2003). The "unstructured" covariance matrix fitted better than other models and was therefore retained for the analysis.

The analysis of the conjoint attributes was limited to main effects due to the study design. As demographic variables are known to influence the attitude towards GM (Siegrist 1998; Titchener *et al.* 2002), we corrected for this by introducing age and gender as covariates in the multilevel analysis. To explain individual differences distinctions were made between (1) respondents that were allergic to apple and those that were not, (2) respondents with a high or low preference for organic foods by applying a median-split on the response to the item that measured preference for organic food products, and (3) respondents who perceived a high or low risk for GM by applying a median-split to the item that measured perceived risk of GM in food. Two-way interactions between attributes and these three explanatory variables were included in the model. Higher order interactions were not included because of the increased complexity of interpretation.

Results

Demographic and allergy characteristics

In the first phase of the survey, 341 respondents, who had all eaten the Santana apple, were recruited. One hundred and eleven respondents were recruited in the second phase, ten of whom had eaten the Santana. Forty-eight respondents with multiple missing data in the apple-profile ratings, attitude measurements or in the demographic characteristics were excluded, leaving 404 respondents contributing data to the analysis. The demographic characteristics of the respondents revealed that women and people with a high education level (BSc or MSc) were overrepresented in the sample (Table 3.2). This sampling effect is likely to be due to self-selection bias. Women are more involved with food shopping (Lake *et al.* 2006), while highly educated people are more prone to respond to internet-based surveys (Cook *et al.* 2000).

The respondents were characterized according to their self-reported allergy complaints (Table 3.2). The majority of the respondents reported allergies, which was expected as respondents were recruited through the leaflet from the Santana package, while people who are food allergic may also be more motivated to respond. Most of the allergic respondents reported that their allergy was diagnosed by a general practitioner or an allergologist (Table 3.2). The majority of the apple allergic patients were also allergic to other fruits, indicating a broader range of dietary restrictions.

For these respondents, the introduction of hypoallergenic apples potentially extends to similar applications to other fruits. In fact, many participants indicated this preference in the free comment space provided at the end of the survey.

TABLE 3.2 - Demographic characteristics from participating respondents and their self-reported allergy background (N=404)

Characteristic		Number of respondents (%)
Gender	Male	138 (34%)
	Female	266 (66%)
Age (mean ± *SD*)		39 ± 13.0
Allergy diagnosis	No allergy	63 (16%)
	Allergy, self-diagnosed	83 (21%)
	Allergy, diagnosed by GP	104 (26%)
	Allergy, diagnosed by allergologist	154 (38%)
Allergic complaints (*n*=341)	Apple allergy (and other fruits)	243 (71%)
	Apple allergy (no other fruits)	58 (17%)
	Other allergy	40 (12%)
Severity of apple allergy (*n*=301)	Mild	80 (27%)
	Considerable	183 (61%)
	Severe	31 (10%)
	Very severe	0 (0%)
	Not reported	7 (2%)

Evaluation of the Santana apple

To evaluate the benefits of the hypoallergenic Santana for apple allergic consumers, the experience and purchase details of the respondents that had eaten the Santana were summarized (Table 3.3). Forty-two percent of the apple allergic consumers had no allergic complaints after eating the Santana. The majority of consumers who experienced allergic responses after eating the Santana reported the symptoms as minor. Only 4% of the respondents reported having serious complaints, while 3% took medication after eating the Santana. No increase in the occurrence of allergic reactions to the Santana was observed in relation to the number of months after harvesting in which the Santana was consumed (logistic regression; β=-0.018, p=0.82), even though prolonged storage is known to induce an increase in the allergenicity of apples (Bolhaar *et al.* 2005a).

To what extent the Santana is appreciated, depends highly on both whether consumers experienced an allergic reaction, and if so, on the severity of this reaction (Table 3.3). Many consumers who experienced an allergic reaction, indicated that their complaints were less severe

compared to complaints experienced following consumption of other apple varieties. This is in concordance with the ratings for liking the Santana; allergy sufferers who reported minor allergic symptoms following consumption of the Santana gave a much higher rating than allergy sufferers who experienced more serious complaints. The difference in liking between allergy sufferers with no allergic reaction and those with very minor ones was relatively small (Table 3.3). Several respondents indicated in the space provided for comments that they were satisfied with the product because attention was being paid to their problem, even if they had an allergic reaction. Thus, despite the occasional occurrence of an allergic reaction, the Santana is still perceived to be beneficial by the majority of the apple allergic consumers.

TABLE 3.3 - Purchase details and allergic response to the Santana

Characteristic		Number of respondents (%)
Santana purchased at (*n*=307)	Supermarket chain	198 (64%)
	Organic food store	23 (7%)
	Greengrocer	6 (2%)
	Elsewhere	15 (5%)
	Not reported	65 (21%)
Type Santana purchased (*n*=307)	Organically grown	177 (58%)
	Traditionally grown	51 (17%)
	Don't remember	79 (26%)
Apple allergic consumer that experienced allergic complaints after eating the Santana (*n*=293)	No complaints	124 (42%)
	Yes, minor complaints	157 (54%)
	Yes, heavy complaints	12 (4%)
Liking the Santana (mean ± *SD*): scale 1-7	Apple allergy, no reaction to Santana(*n*=124)	6.24 ± 0.59
	No apple allergy (*n*=14)	5.79 ± 1.22
	Apple allergy, minor reaction to Santana (*n*=157)	4.89 ± 1.48
	Apple allergy, heavy reaction to Santana (*n*=12)	1.42 ± 0.79

The Santana was considered suitable for consumption by patients with a mild apple allergy (Kootstra *et al.* 2008), as was clearly stated in the explanatory leaflet. The self-reported severity of the respondents' apple allergy was associated with the occurrence of an allergic reaction to the Santana (χ^2=7.35, *df* 2, p<0.05). The majority of the respondents in the current study self-reported that they suffered from mild to considerable allergic reactions to apple. Forty-seven percent of the respondents with a mild apple allergy experienced an allergic reaction to the Santana. In the group with "considerable" allergic reactions to apple this increased to 65%. Ten percent of the respondents reported to suffer from a severe apple allergy, but still tried eating the Santana. Consumers with a severe apple allergy were not advised to try the Santana, but among the respondents in this patient group who did try the Santana 38% reported having no allergic reaction. We examined which variables potentially predicted the occurrence of an allergic reaction

following consumption of the Santana. Two other variables were associated with the occurrence of an allergic reaction, namely the occurrence of other fruit allergies (χ^2=6.77, *df* 1, *p*<0.01) and age group (χ^2=11.79, *df* 5, *p*<0.05). The occurrence of an allergic reaction to the Santana increased with an increasing self-reported severity of the apple allergy and with the occurrence of other fruit allergies. The occurrence of an allergic reaction to the Santana decreased with increasing age. Combining these three factors yields the best indication for the occurrence of an allergic reaction.

Apple profiles

Apple profiles with varying characteristics were rated by the respondents to examine consumer preferences. The items that measured "liking" and "willingness to buy" of the apple profiles were highly correlated (*r*=0.92, *p*<0.001) and were averaged to calculate a single attitude scale, labeled as "attitude". The scores on "attitude" were subjected to a repeated measures mixed linear model. We tested the main effects of varying "breeding method", "pesticide reduction" and "allergy reduction", and their interaction with "perceived GM risk", "apple allergy" and "preference for organic products". Neither the health status items, nor the item that measured "environmental concern" had an additional effect on attitude when "apple allergy" and "preference for organic products" were included, and were therefore omitted from further analysis.

TABLE 3.4 - Repeated measure mixed linear model[1] explaining the attitude towards the apple profiles

Fixed parameters	(df_{effect}, df_{error})	F value	*p* value
Main effects			
Breeding method	(2, 370)	96.605	< 0.001
Pesticide reduction	(1, 377)	139.999	< 0.001
Allergy reduction	(2, 371)	208.554	< 0.001
Interaction effects			
Apple allergy x Breeding method	(2, 371)	13.949	< 0.001
Apple allergy x Pesticide reduction	(1, 379)	8.169	< 0.01
Apple allergy x Allergy reduction	(2, 372)	67.546	< 0.001
Organic products x Breeding method	(2, 364)	1.073	0.34
Organic products x Pesticide reduction	(1, 374)	15.406	<0.001
Organic products x Allergy reduction	(2, 373)	2.187	0.11
GM Risk x Breeding method	(2, 365)	35.868	< 0.001
GM Risk x Pesticide reduction	(1, 376)	0.145	0.70
GM Risk x Allergy reduction	(2, 374)	8.569	< 0.001

1 Covariates appearing in the model are evaluated at the following values: Gender 1.66 (male=1; female =2) and age 39.09

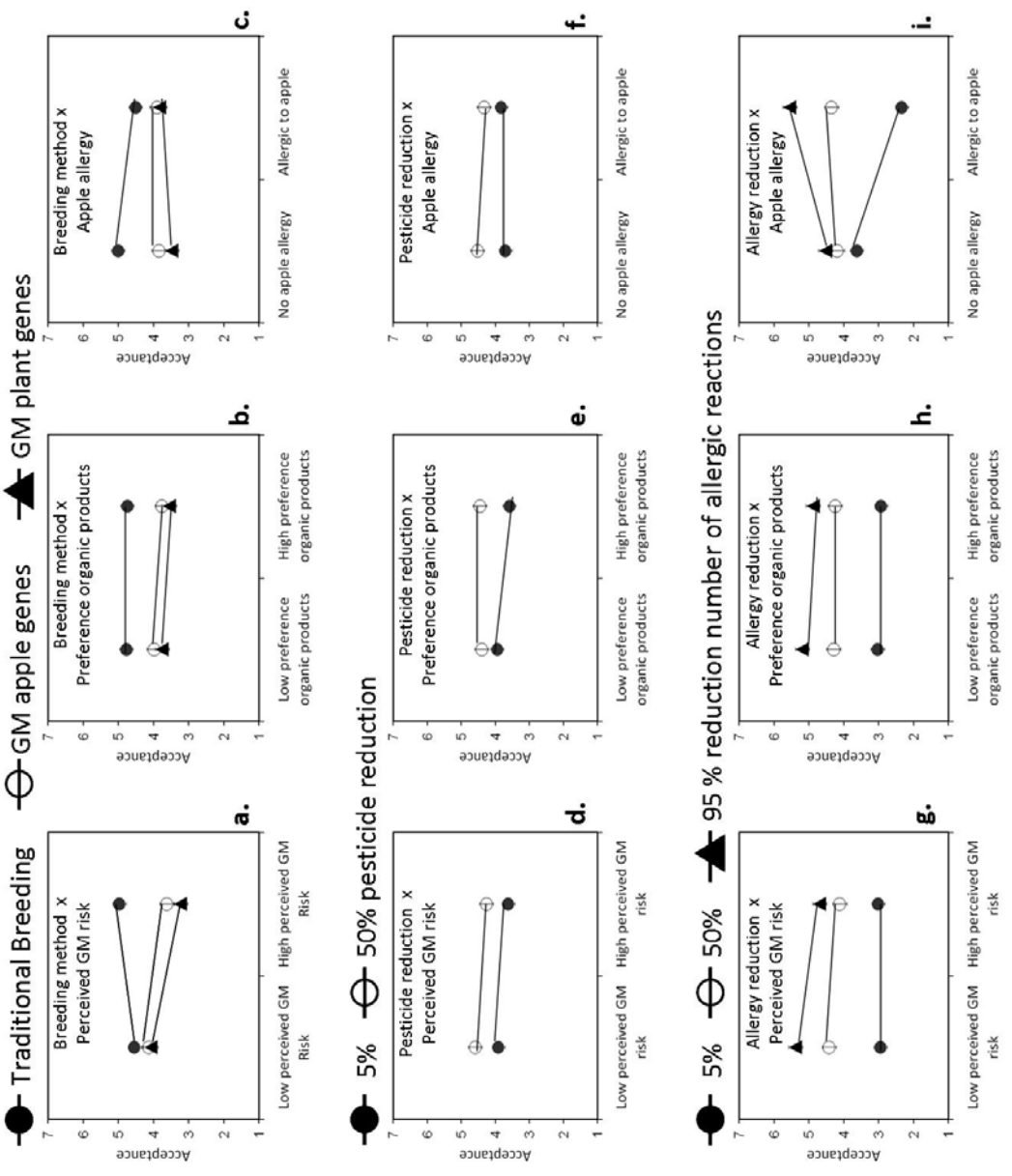


Figure 1a to 1i - Interaction effects between main effects and respondent groups explaining the attitude towards the apple profiles (mean ± confidence interval)

Breeding method

Different breeding methods (traditional breeding and GM) can be applied to develop hypoallergenic apples. The results indicate that the breeding method utilized has a significant effect on consumer attitude (Table 3.4). The mean attitude towards apple profiles differed significantly between all breeding methods (Table 3.5). Traditional breeding was rated more positive than both GM strategies. GM with genes from another apple (cisgenes) was rated more positive than GM with genes from another plant species. The difference between the two GM strategies was significant, but the effect size was about 6 times smaller when compared to the difference between the GM methods and conventional breeding. Hence, it is questionable whether the difference between the two GM methods is relevant as far as consumer acceptance is concerned. Significant interaction effects between breeding methods and both "perceived GM risk" and "apple allergy" were found. Respondents who consider GM to be low in risk differentiated between traditional breeding and the GM strategies, but the differences were much less pronounced than for respondents who consider GM to be high in risk (Fig. 3.1a). Apple allergic respondents were less explicit in their preferences for the different breeding strategies than respondents without an apple allergy, which is to be expected as one of the reasons that GM was applied, was to reduce allergenicity (Fig. 3.1c).

TABLE 3.5 - Mean score (± S.E) of the "attitude" variable for the separate attribute levels

Attribute	Number of levels	Mean	S.E.
Breeding method	Traditional breeding	4.76	± 0.070
	Genetic modification using material from other apple varieties	3.87	± 0.076
	Genetic modification using material from other plant species	3.65	± 0.075
Pesticide usage	5% pesticide reduction during growing	3.76	± 0.062
	50% pesticide reduction during growing	4.43	± 0.070
Apple allergy	5% of apple allergic patients without complaints	2.98	± 0.078
	66% of apple allergic patients without complaints	4.27	± 0.078
	95% of apple allergic patients without complaints	5.04	± 0.078

Covariates appearing in the model are evaluated at the following values: Gender 1.66 (male=1; female =2) and age 39.1. All means were significantly different from each other at $p<0.05$ within each attribute

Pesticide reduction

The results indicate a significant main effect of reduced pesticide usage on attitude (Table 3.4). The respondents accepted 50% pesticide reduction significantly better than 5% pesticide reduction (Table 3.5). Significant interaction effects were detected for "organic products" and "apple allergy" (Table 3.4). Respondents who expressed a high preference for organic products were more negative about the 5% pesticide reduction than respondents with a low preference for organic products (Fig. 3.1d-f). The 50% pesticide reduction was rated equally by both these respondents

groups. No difference was observed in the mean attitude ratings of respondents with or without an apple allergy for the 5% pesticide reduction, but the increase in the ratings of the 50% pesticide reduction was less for apple allergic respondents.

Allergy reduction

The results indicate a significant main effect of the reduction in the occurrence of allergic reactions on attitude (Table 3.4). The 5% reduction in the number of allergic responses following consumption was rated significantly less favorably compared to the 66% reduction, which was in turn rated significantly less favorably than the 95% reduction (Table 3.5). Significant interaction effects were found for “apple allergy” and “perceived GM risk” (Table 3.4). The 66% allergy reduction was rated similarly by apple allergic respondents and respondents without an apple allergy (Fig. 3.1g-i). Allergic respondents were relatively more negative about the 5% reduction, while they were relatively more positive about the 95% reduction than respondents without an apple allergy. Respondents who consider the application of GM to be high in risk were relatively less positive about the 66% and 95% reduction than those who consider GM to be low in risk.

Effect of a negative reaction to Santana on attitude

For apple allergic consumers who had an allergic reaction to the Santana apple, the first experience with a hypoallergenic food was negative. This may influence their attitude towards similar products, because the benefits claimed for the hypoallergenic product have not turned out as real benefits for these consumers. Therefore, we conducted the analysis on the apple profiles for apple allergic consumers alone, while comparing respondents with or without an allergic reaction to the Santana. The attitude scores for the apple profiles were subjected to a repeated measures mixed linear model again, while testing the main effects of “breeding method”, “pesticide reduction” and “allergy reduction”, and the interaction with “perceived GM risk”, “allergic reaction to Santana” and “preference for organic products”. We observed no significant interactions between “allergic reaction to Santana” and the three attributes. The other effects were highly similar to the results reported previously; both with respect to significances and effect sizes (see Table 3.4). Apparently, having an allergic reaction to the Santana did not influence participants’ attitude towards the apple profiles. Even apples that have similar characteristics to the Santana (66% allergy reduction) were rated equally high irrespective of whether consumers had experienced an allergic reaction to the Santana.

TABLE 3.6 - Repeated measure mixed linear model[1] explaining the attitude towards the apple profiles that include the application of GM

Fixed parameters	(df_{effect}, df_{error})	F value	*p* value
Main effects			
Pesticide reduction	(1, 347)	52.432	< 0.001
Allergy reduction	(2, 370)	168.199	< 0.001
Apple allergy	(1, 381)	1.201	0.27
Organic products	(1, 383)	3.701	0.06
GM Risk	(1, 381)	37.574	< 0.001
Interaction effects			
Apple allergy x Pesticide reduction	(1, 350)	7.892	< 0.01
Apple allergy x Allergy reduction	(2, 371)	51.853	< 0.001
Organic products x Pesticide reduction	(1, 362)	5.820	<0.05
Organic products x Allergy reduction	(2, 373)	1.197	0.30
GM Risk x Pesticide reduction	(1, 361)	4.277	<0.05
GM Risk x Allergy reduction	(2, 373)	13.278	< 0.001

1 Covariates appearing in the model are evaluated at the following values: Gender 1.66 (male=1; female =2) and age 39.1

Perceived benefits and GM

The apple profiles with higher benefits (pesticide reduction and allergy reduction) evoked more positive attitudes. However, the described analysis includes both traditional breeding and GM strategies. For GM apples, the differences in perceived benefits may have been overwhelmed by perceived risk (Frewer *et al.* 2002). We therefore examined whether the properties of the apple profiles are also perceived to be beneficial when only a GM method is involved in their development, or whether the effects reported in Fig. 3.1 should be fully attributed to conventionally grown apples. The repeated measures mixed linear model analysis was repeated for apple profiles that included one of the two GM strategies only. For these GM apples, the effects of "perceived GM risk" and "apple allergy" were considered, as these factors had a strong interaction with the breeding method applied (Table 3.4). When doing so, "perceived GM risk" had a strong main effect on attitude towards GM apples. Respondents who perceive GM to be high in risk provided lower attitude ratings. The other results from the analysis on the GM apple profiles (Table 3.6) were highly similar to those from the initial analysis (Table 3.4). The interactions between "pesticide reduction" and "preference for organic products", and between "apple allergy" and "allergy reduction" were still significant for the GM strategies alone. Thus, the interaction effects between benefits of the apple profiles (pesticide and allergy reduction) and explanatory variables were observed, regardless of whether or not GM technology was involved in their development.

Discussion

In this study, we investigated consumer acceptance of hypoallergenic products. The first part of the study evaluated the recent marketing of a hypoallergenic apple variety, called Santana. The benefits claimed for the Santana were compared to the actual benefits experienced by (allergic) consumers. The second part of the study explored consumer acceptance of similar hypoallergenic products and examines whether hypoallergenic GM foods will be accepted by allergic and non-allergic consumers. The introduction of hypoallergenic properties into food products serves as a good example of "personal benefits" for a specific consumer group and can be viewed in the wider context of acceptance and rejection of GM in general, and of the consequences for food allergy management.

The hypoallergenic Santana apple

The Santana apple was not completely new to consumers, but its availability had been very limited in volume and, so far, restricted to organic stores. The Santana is resistant against scab and therefore requires less spraying of pesticides, which makes it an attractive apple for organic growing (CPRO-DLO 1999). The Santana was recently identified as being hypoallergenic (Bolhaar *et al.* 2005a), which was confirmed in a food challenge study (Kootstra *et al.* 2008), which also indicated that the Santana would be suitable for the majority of allergic patients with an apple allergy. Indeed, the results of the current study show that almost half of all apple allergic consumers were able to eat the Santana without experiencing an allergic reaction. It should, however, be kept in mind that the respondents in our study were self-reporting their reactions and were not drawn randomly from the apple allergic population. Although some of these consumers state that their reaction was less severe than reactions experienced following consumption of other apple varieties, the number of consumers reacting to the Santana was higher than expected. A correlation was observed between an increasing occurrence of an allergic reaction to the Santana and an increasing severity of apple allergy in the self-report data. In addition, the occurrence of an allergic reaction to the Santana increased if respondents were younger, or suffered from other fruit allergies. Age is also related to the severity of allergies (Barbee *et al.* 1981), while the occurrence of other fruit allergies may indicate a high degree of cross-reactivity. The factors that determine the occurrence of an allergic reaction require further investigation, since it is currently not possible to provide criteria that guarantee the absence of allergic reactions.

Three potential factors may explain why consumption of the Santana results in allergic responses in some patients and not in others. First, the same food may cause different clinical response profiles depending on the allergens involved. Apple contains multiple allergens (Pastorello *et al.* 1999; Wensing *et al.* 2002) and the Santana may not be hypoallergenic for all allergens involved. Second, variation in the occurrence of allergic reactions may be caused by differences in the amount of allergens between different apple varieties (Marzban *et al.* 2005). Individual consumers

have different thresholds for an allergic reaction; allergen levels that are tolerated by one consumer may trigger an allergic reaction in others (Taylor *et al.* 2002). Third, various protein variants of each apple allergen exist (Gao *et al.* 2005a; Gao *et al.* 2005b). Protein variants which trigger an allergic reaction in one patient may not do so in others, as has been shown for variants of the birch allergen Bet v 1 (Ferreira *et al.* 1997). Apple varieties display variation in the abundance of particular isoforms (Marzban *et al.* 2005), which may also evoke different responses in individual patients.

Overall, the Santana is rated positively by all consumers except those that experienced a strong allergic reaction following consumption. Consumers who did not experience an allergic reaction are now able to consume apples. This contributes to a normal and healthy diet. For consumers who experience a (reduced) allergic reaction following consumption of the Santana, the benefits are small or absent and they still can not eat apples. However, consumers who reported to have minor allergic complaints to the Santana were generally positive about the Santana. These results were validated in the study on the apple profiles which showed that respondents who suffered of an allergic response to the Santana still liked similar hypoallergenic products. Some consumers stated that they were positive about the Santana because attention was being paid to their problem, a well-known psychological effect (Wickstrom *et al.* 2000). This clearly indicates a positive attitude towards hypoallergenic foods, whether or not they are completely successful. It is questionable whether food allergic consumers will continue to be positive about future hypoallergenic products if the latter still do not alleviate the allergic complaints.

Two issues require further attention from the perspective of food allergy management. The Santana was packaged with an explanatory leaflet in the supermarket, but was sold without the explanatory leaflet in organic stores and greengrocers. This could be problematic, since the leaflet contains information about safe consumption by apple allergic consumers. Despite the strong recommendation on the explanatory leaflet that the Santana is only suitable for patients with a mild apple allergy, this did not stop allergic consumers with severe apple allergy from trying the Santana. In the Dutch context, allergic patients with a *very* severe allergic reaction (anaphylactic shock) are unlikely to be found. However, given the potential impact of a severe reaction (severe swelling of the oral mucosa), consumption by such patients is undesirable, and safe consumption relies heavily on the accurate use of the testing protocol that was described on the explanatory leaflet. The step-by-step testing protocol aims at minimizing the consequences of the occurrence of an allergic reaction. How seriously consumers take the protocol, and whether or not they test their reactions to the Santana according to the guidelines provided, is not known. The experience with the Santana may indicate that some allergic consumers will test hypoallergenic food, regardless of whether it is suitable for them. Perhaps more explicit warnings are required to avoid consumers with a severe allergy from trying these products. Given the outcomes of the current

study, we recommend a cautious approach regarding interventions based on hypoallergenic foods, or the use of hypoallergenic products as ingredients, certainly for food products that may cause very severe allergic reactions like peanut.

Hypoallergenic GM foods

The Santana represents a first step in creating hypoallergenic products, and further breeding may enhance hypoallergenic properties. This will take time, as the average time from crossing a new variety to selling it on the market is 20 years. Developments of hypoallergenic foods may, therefore, include the application of GM (Dodo *et al.* 2007; Gilissen *et al.* 2005). Societal concern about GM has proven a major factor delaying the development of GM foods (Gaskell *et al.* 2006). Societal attitudes towards GM is influenced by case-specific characteristics of the application, such as which organism is modified and to what purpose (Frewer *et al.* 1997; Zechendorf 1994). The majority of the European consumers are positive about medical GM, whilst at the same time rejecting agricultural applications. Medical applications are considered more necessary, because of the treatment of illness, and are thereby more acceptable than food-related applications that appeared to benefit the producer, at least for the first generation of GM products (Frewer *et al.* 1997; Zechendorf 1994). Hypoallergenic GM foods are of interest because they are on the boundary between medical- and food-related applications. The results of this study show a clear consumer preference for traditional breeding over breeding by GM in the development of hypoallergenic apples, consistent with previous research (Miles *et al.* 2005; Schenk *et al.* in press). Nevertheless, acceptance of hypoallergenic GM apples is quite high among both apple allergic and non-allergic consumers, which is in line with acceptance of medical applications. Similar results were observed for acceptance of hypoallergenic non-food products, such as hypoallergenic GM birch or grass (Schenk *et al.* in press).

GM acceptance: Personal benefits *vs.* societal benefits

The present study examined the influence of “perceived benefits” on acceptance of GM food products. Previous studies have found a two-dimensional attitudinal structure towards GM in which “perceived benefit” and “perceived risk” were influential in determining consumer acceptance (Frewer *et al.* 1997; Schenk *et al.* in press). Differences in acceptance for GM hypoallergenic products by allergic sufferers were associated with perceived benefits, while the perceived risks remained constant (Schenk *et al.* in press). Frewer (2003) concluded that as long as the risks are not as large as to be completely intolerable, acceptance would be driven by perceptions of personal benefit. Apart from the influence of benefits on GM acceptance *per se*, evidence suggests that the perception that consumers, rather than industry, will benefit from a particular application plays an important role in GM acceptance (Miles *et al.* 2001). In the research presented here, we hypothesized that acceptance of GM products is driven by consumer

recognition of specific and personally relevant benefit and, therefore, can be divided into "personal benefits" and "societal benefits".

Our results indicate that "personal benefits" have a much stronger impact on GM acceptance than "societal benefits (Fig. 3.1). The effect of an increased allergy reduction on acceptance of GM is stronger for apple allergic consumers, who experience a direct "personal benefit" associated with hypoallergenic GM apples, compared to consumers without apple allergy. In contrast, non-allergic consumers may consider hypoallergenic GM products to be beneficial because these products alleviate complaints of other consumers (i.e. as societal benefit). A similar argument can be made for GM apples which require less pesticide during their production. Using less pesticide has an effect on acceptance of GM among consumers with a preference for organic products. These people judge environmental effects as more important; hence, the effects of reduced pesticide usage are expected to have a larger impact on their personal preference. However, in the case of consumers who prefer organic products, GM applications are always less acceptable, even if they are associated with reduced pesticide usage because these consumers associate GM with negative impact on environment (Deliza *et al.* 1999). The results of the study show that societal benefits may also facilitate acceptance of GM products.

Conclusion

Hypoallergenic foods have the potential to contribute to food allergy management. The results from a large-scale sales pilot of the Santana apple show that the Santana alleviates allergic complaints in nearly half of the apple allergic consumers. Consumers that do experience a mild allergic reaction to the Santana are also relatively positive about the Santana. Attitude towards similar hypoallergenic products was positive, also when GM was involved in their development, because allergic consumers experience a "personal benefit" associated with hypoallergenic GM products. This suggests that consumers prioritize their personal needs and beliefs, and accept applications of GM when the benefits are recognized. Still, given the observed differences among individual apple allergic consumers in the response to the Santana, we recommend a cautious approach regarding interventions based on hypoallergenic foods.

CHAPTER 4

Seven different genes encode a diverse mixture of isoforms of Bet v 1, the major birch pollen allergen

Martijn Schenk, Luud Gilissen, Danny Esselink and René Smulders

Published as: Schenk MF, Gilissen LJWJ, Esselink GD & Smulders MJM (2006) Seven different genes encode a diverse mixture of isoforms of Bet v I, the major birch pollen allergen. BMC Genomics 7.

Abstract

Pollen of the European white birch (*Betula pendula*, *syn. B. verrucosa*) is an important cause of hay fever. The main allergen is Bet v 1, member of the pathogenesis-related class 10 (PR-10) multigene family. To establish the number of PR-10/Bet v 1 genes and the isoform diversity within a single tree, PCR amplification, cloning and sequencing of PR-10 genes was performed on two diploid *B. pendula* cultivars and one interspecific tetraploid *Betula* hybrid. Sequences were attributed to putative genes based on sequence identity and intron length. Information on transcription was derived by comparison with homologous cDNA sequences available in GenBank/EMBL/DDJB. PCR-cloning of multigene families is accompanied by a high risk for the occurrence of PCR recombination artifacts. We screened for and excluded these artifacts, and also detected putative artifact sequences among database sequences. Forty-four different *PR-10* sequences were recovered from *B. pendula* and assigned to thirteen putative genes. Sequence homology suggests that three genes were transcribed in somatic tissue and seven genes in pollen. The transcription of three other genes remains unknown. In total, fourteen different Bet v 1-type isoforms were identified in the three cultivars, of which nine isoforms were entirely new. Isoforms with high and low IgE-reactivity are encoded by different genes and one birch pollen grain has the genetic background to produce a mixture of isoforms with varying IgE-reactivity. Allergen diversity is even higher in the interspecific tetraploid hybrid, consistent with the presence of two genomes. Isoforms of the major birch allergen Bet v 1 are encoded by multiple genes, and we propose to name them accordingly. The present characterization of the Bet v 1 genes provides a framework for the screening of specific *Bet v* 1 genes among other *B. pendula* cultivars or *Betula* species, and for future breeding for trees with a reduced allergenicity. Investigations towards sensitization and immunotherapy should anticipate that patients are exposed to a mixture of Bet v 1 isoforms of different IgE-reactivity, even if pollen originates from a single birch tree.

Background

Pathogenesis-related class 10 (PR-10) proteins constitute the largest group of aeroallergens and are among the four most common food allergens (Breiteneder *et al.* 2000). The main allergen is a PR-10 pollen protein from the European white birch *(Betula pendula)* termed Bet v 1 (Breiteneder *et al.* 1989). Birch pollen is a major cause of Type I allergies in the temperate climate zone of the northern hemisphere. Over 95% of the tree pollen-sensitized patients in Scandinavia display IgE binding to Bet v 1, while 60% react exclusively to this allergen (Jarolim *et al.* 1989). Pollen of other Fagales species contains Bet v 1 homologues that share epitopes with Bet v 1 (Niederberger *et al.* 1998).

PR-10 proteins are present as a multigene family across a range of phylogenetically distant species, including Gymnosperms, Monocots, and Dicots (Liu *et al.* 2004; Lüttkopf *et al.* 2002; Wang *et al.* 1999). As a consequence, several foods contain Bet v 1 homologues, including nuts, vegetables, and Rosaceae fruits (Bohle *et al.* 2003; Lüttkopf *et al.* 2002; Neudecker *et al.* 2003). Patients that are sensitized to Bet v 1 may experience mild allergic symptoms upon consumption of these foods due to IgE cross-reactivity. Symptoms of this so called oral allergy syndrome (OAS) are mainly limited to the oral cavity. Cross-reactivity has clearly been demonstrated by allergic responses to the celery protein Api g 1, which is initiated by sensitization to Bet v 1 (Bohle *et al.* 2003).

The *B. pendula* genome contains multiple *PR-10* genes with varying expression patterns. Among these, the Bet v 1 allergens are expressed in pollen. The first Bet v 1 isoform was identified by immunoscreening a pollen cDNA expression library with serum of birch pollen allergic patients (Breiteneder *et al.* 1989). Other Bet v 1 isoforms have been sequenced by various authors since then (Friedl-Hajek *et al.* 1999; Hoffmann-Sommergruber *et al.* 1997; Son *et al.* 1999; Swoboda *et al.* 1995b). Pollen mixtures from multiple trees were found to contain multiple Bet v 1 isoforms (Swoboda *et al.* 1995b). Bet v 1 isoforms differ in the ability to bind IgE and in the T-cell proliferation score (Ferreira *et al.* 1996). Two other types of PR-10 proteins were detected in birch cells that were grown in a liquid medium in the presence of microbial pathogens (Swoboda *et al.* 1995c). These proteins are expressed in roots and leaves under basal conditions or induced under various stress-related conditions (Poupard *et al.* 1998; Swoboda *et al.* 1995c). The *PR-10* genes from *B. pendula* form a homogeneous group, based on sequence similarities. Homogeneity is suggested to be maintained by concerted evolution (Wen *et al.* 1997). Arrangements of *PR-10* genes into clusters, such as found for Mal d 1 genes in apple *(Malus domestica)*, may facilitate concerted evolution (Gao *et al.* 2005a).

Given the prominent role of Bet v 1 allergens in the sensitization to hay fever and OAS, birch is a relevant target for development of allergy prevention strategies. Selection and breeding of

hypoallergenic trees or the application of genetic modification to develop these may potentially reduce the allergenic load caused by birch. Knowledge on the diversity of *PR-10* genes, their expression, and allergenicity of the gene products is required to facilitate these strategies. In the present study, we amplified, cloned and sequenced *PR-10* alleles from three *B. pendula* cultivars to establish the number of *PR-10/Bet v 1* genes and the isoform diversity within a single tree.

Methods

Plant material

The natural distribution range of *B. pendula* Roth *(syn. B. verrucosa)* covers almost the whole of Europe. Several *B. pendula* cultivars have been bred, including interspecific hybrids (also referred to as interspecies hybrids) between *B. pendula* and other *Betula* species. We collected young leaves from three *B. pendula* cultivars in the collection of PPO Boskoop (WUR, the Netherlands), namely 'Long Trunk', 'Schneverdinger Goldbirke', and 'Tristis'. Fresh leaf samples were sent to Plant Cytometry Services (Schijndel, The Netherlands) and screened by flow cytometry to estimate the ploidy level. Diploid *(B. pendula)* and tetraploid *(B. pubescens)* controls were included. The cultivars 'Schneverdinger Goldbirke' and 'Tristis' were diploid, while the cultivar 'Long Trunk' was tetraploid. The latter is likely to be an interspecific hybrid between *B. pendula* and an unknown second *Betula* species. The alleles recovered from 'Long Trunk' were either assigned to *B. pendula* based on sequence and intron similarity or were considered specific for this cultivar. The specific 'Long Trunk' alleles were analyzed separately. DNA was extracted using the DNeasy Plant Mini kit (Qiagen) according to the manufacturer's instructions.

PCR, cloning and sequencing

PR-10 alleles were amplified from birch DNA with primers designed after two cDNA sequences (X15877, X77601). The primers were complementary to the regions around the start and stop codons (shown in bold); BpI-For: 5'-AATCTCTCAGGCCATC**ATG**GGTG-3', BpI-Rev: 5'-**TA**GTTGTAGGCATCGGAGTGTGC-3', BpII-For: 5'-ATCTCAGGTGATCATC**ATG**GGTG-3', and BpII-Rev: 5'-**TA**GTTGTAGGCATTTGGGTGTGC-3'. PCR amplification with both primer pairs was performed with mixtures consisting of 2 µl dNTP (1mM), 2 µl 10x Reaction buffer, 0.8 µl $MgCl_2$ (25mM), 1.2 µl forward primer (10 pmol/µl), 1.2 µl reverse primer (10 pmol/µl), 0.11 µl of a 1:9 mixture of *Pfu* polymerase (Stratagene) and *Taq* polymerase (Goldstar)(5U/µl), and 20-80 ng template DNA. H_2O was added to obtain a total volume of 20 µl. PCR mixtures were subjected to the following conditions: initial heating step at 95˚C for 15 minutes, denaturation at 94˚C for 30 s, annealing at 50˚C for 45 s, and extension at 72˚C for 60s. A final extension step of 10 min at 72˚C was added after 22-30 cycles. Given the observation of recombination among the recovered sequences, we subsequently varied the number of PCR cycles at intervals of 2 cycles. The minimum number of cycles was established by visual inspection of the amplification products on agarose gel at 22 for

the BpI primer pair and at 24 for the BpII primer pair. Originally, 30 PCR cycles had been used. We repeated the experiment at 22-24 cycles to ensure that amplification was in its linear phase.

To obtain the A-tailing that facilitates the ligation procedure, five additional cycles were run on 1-4 µl of PCR product with *Taq* polymerase (Goldstar). PCR conditions were similar as described above. PCR products were purified with the MinElute PCR Purification Kit (Qiagen). Purified samples were ligated into the pGEM-T easy Vector (Promega) and established in *Escherichia coli* XL1 Blue competent cells (Stratagene) according to the manufacturer's instructions. White colonies were picked from agar plates and grown overnight at 37˚C in freeze medium. PCR-based screening was performed with vector-specific M13 primers. PCR products were purified with Sephadex G-50 (Millipore). The DYEnamic™ ET Terminator Cycle Sequencing Kit (Amersham) was used for the sequence reaction. Sequence products were analyzed on a 96-capillary system (ABI 3730xl).

Genomic *B. pendula* sequences have been submitted to GenBank as DQ296566-DQ296598 and DQ325525-DQ325535, and the specific 'Long Trunk' sequences as DQ296599-DQ296610.

GenBank sequences

Sixty-six PR-10 sequences were obtained from GenBank/EMBL/DDJB by searching with MegaBLAST for entries from *B. pendula* with more than 60% sequence identity to Bet v 1a (X15877). The search results included nineteen genomic DNA (gDNA) sequences: Z72429-Z72438 (Hoffmann-Sommergruber *et al.* 1997); AJ001551-AJ001557 (Cvitanich and Larsen, direct submission); AJ289770 and AJ289771 (Pellinen *et al*, direct submission). The accessions Z72435-8 were singletons with a relatively low homology to our and other GenBank *Betula* sequences and were excluded from the analysis. Forty-seven cDNA sequences were found: X15877 (Breiteneder *et al.* 1989); X77200, X77265-X77274, X81972, X82028 (Swoboda *et al.* 1995b); X77599-X77601 (Swoboda *et al.* 1995c); Z80098-Z80106 (Larsen, direct submission); AF124837-AF124839 (Son *et al.* 1999); AJ002106-AJ002110, and AJ006903-AJ006915 (Friedl-Hajek *et al.* 1999).

Phylogenetic analysis

Nucleotide sequences were aligned using CLUSTALW (Thompson *et al.* 1994). Bayesian phylogenetic analysis was performed with MrBayes 3.1.1 (Huelsenbeck *et al.* 2002). The maximum likelihood model employed 6 substitution types, with base frequencies set to the empirically observed values. Rate variation across sites was modeled using a gamma distribution. The Markov chain Monte Carlo search was run twice with 4 chains for 1,000,000 generations. Topology and model parameters were sampled every 100th generation and used to estimate model parameters and to determine the posterior probabilities of clades. The first 100,000 generations were discarded as "burn in". The outgroup was composed of *PR-10* sequences from *Prunus armeniaca* (AF020784), *P. avium* (U66076), *Pyrus communis* (AF057030), and *Malus domestica* (X83672,

Z72425, Z72427). To date, these are the closest non-Fagales relatives of the PR-10 family from *B. pendula*. We confirmed the phylogenetic analysis by constructing a neighbour-joining tree with Kimura two-parameter distances. Bootstrapping was carried out with 1,000 replicates. The results were similar for both analyses; therefore, the results from the neighbor-joining analysis are not shown.

Based on variation among *Mal d 1* alleles of a single locus compared to alleles of different loci (Gao *et al.* 2005a), we predefined a cut-off level of 98% identity in order to identify clusters that encompass alleles of the same gene. A limitation of the applied method is that genes that are homozygous and differ less than 98% from each other remain undetected.

Recombination was detected and visualized with Phylpro 1.0 (Weiller 1998). This program computes the correlation coefficient of pair wise distances between the target sequence and all other sequences on both sides of a sequence position. We used a sliding window of 40 base pairs.

Isoforms and allergenicity

Nucleotide sequences were aligned codon-by-codon. We analyzed general selection patterns at the molecular level using DnaSp 4.00 (Rozas *et al.* 2003). The number of synonymous (K_s) and non-synonymous substitutions (K_a) per site were calculated from pair wise comparisons with incorporation of the Jukes-Cantor correction. Nucleotide data were translated into amino acids. Putative isoforms were analyzed for their potential allergenicity by screening those amino acid positions that have been identified as influencing the IgE binding (Ferreira *et al.* 1996; Son *et al.* 1999) and T-cell activation (Jahn-Schmid *et al.* 2005).

Results

PCR recombination artifacts

When PCR amplification is performed on groups of closely related sequences, such as the PR-10 gene family, accurate sequences are essential to distinguish between members. When we initially determined the relationships among the recovered sequences, most clades in the Bayesian consensus tree had low posterior probabilities. Several sequences occupied intermediate positions between well-defined clusters. This suggested the possibility of recombination. Recombination could have occurred *in vivo* through a crossing-over or gene conversion between prior existing genes, or *in vitro* during the PCR through strand-switching or re-annealing of incompletely amplified fragments.

Evidence supports the view that recombinant sequences were PCR artifacts. Recombination signals were abundant in the sequences obtained after the 30-cycle PCR (Fig. 4.1a) and virtually absent when 22 cycles were employed (Fig. 4.1b). Several GenBank sequences showed clear

evidence for recombination too (Fig. 4.1c). Putative recombinants from our experiments lacked or nearly lacked unique mutations and could be separated into two or three stretches that were identical to other sequences obtained. The GenBank recombinants shared close to 100% sequence identity to combinations of other accessions.

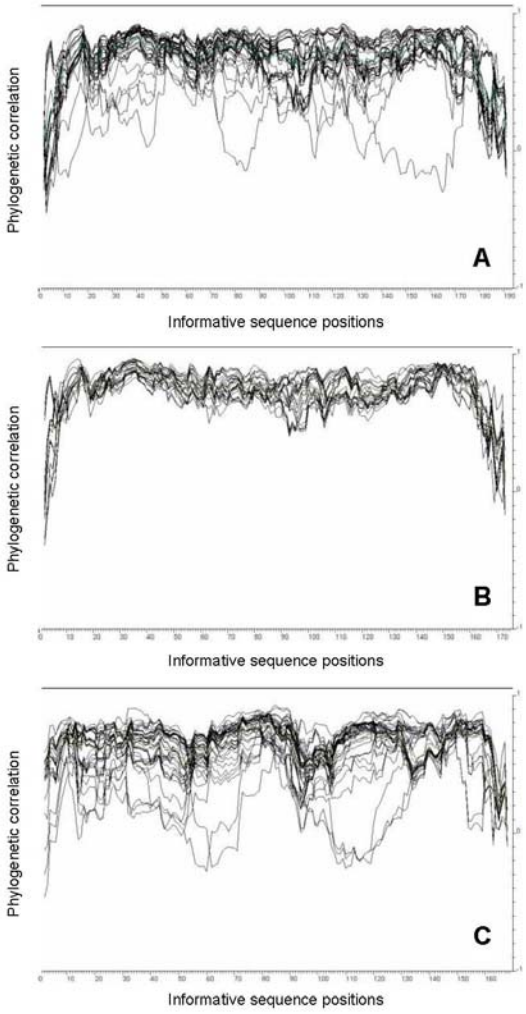


Figure 4.1 - Phylogenetic profiles for detection of recombination. Phylogenetic profile of the sequences from B. pendula *'Long Trunk' obtained after a PCR of (a.) 30 cycles (*n*=72 sequences) and (b.) 20 cycles (*n*=53). (c.) Phylogenetic profile of the GenBank PR-10 sequences from* B. pendula *(*n*=66). The x-axis represents the sequence position (5'-3' including only informative positions). The y-axis indicates the phylogenetic correlation. Low values are indicative for recombination (Weiller 1998). Low values at the edges are artifacts of the employed method.*

The use of independent PCRs resolves which sequences are genuine, since the probability of isolating identical artifacts in independent PCRs is extremely low. Most sequences without recombination signal were confirmed in independent PCRs (Table 4.1), while those with a recombination signal were not. The only exception was the *PR-10.03B02.01* allele from 'Tristis' that was found in three independent PCRs. This allele was an *in vivo* recombination of the first 300 bp, including the intron, from the *PR-10.03D* gene and 183 bp from the original *PR-10.03B* gene as found in 'Schneverdinger Goldbirke' and 'Long Trunk'. Putative recombination artifacts were quite abundant in the 26 and 30-cycle PCR (27-46% of the sequences), but rare in the 22 and 24-cycle PCR (2-11%) (Table 4.1). We conservatively maintained the sequences that were confirmed in

independent PCRs and those with at least three unique mutations for further analysis. A minimum of three unique mutations was chosen to ensure that potential base mis-incorporation artifacts were excluded.

TABLE 4.1 – Cloned and sequenced PR-10 sequences from B. pendula*. Overview of the individual clones of PR-10 sequences from* B. pendula *'Schneverdinger Goldbirke', 'Tristis', and 'Long Trunk'. Different primers (BpI, BpII) were used. De number of cycles varied between 22 and 30. Confirmed sequences are found in multiple independent PCRs. Unique sequences differ at least by three base pairs from any other sequence from the same cultivar. The remaining sequences are either recombination artifacts or presumably result from base mis-incorporations. The number of alleles included for further analysis is also indicated.*

Cultivar 'Long Trunk'							
Primer combination	BpI		BpII		BpI		No. of different
no. of PCR cycles	22 cycles		24 cycles		30 cycles		alleles included
Confirmed sequences	43	(86%)	50	(69%)	24	(45%)	20
Unique sequences	2	(4%)	5	(7%)	5	(9%)	8
Recombination artifacts	1	(2%)	8	(11%)	18[*1]	(34%)	
Base mis-incorporation artifacts	4	(8%)	9	(13%)	6	(11%)	
Total no. of clones	50		72		53		28
Cultivar 'Tristis'							
Primer combination	BpI		BpII		BpI		No. of different
no. of PCR cycles	22 cycles		24 cycles		30 cycles		alleles included
Confirmed sequences	17	(77%)	20	(53%)	26	(54%)	11
Unique sequences	1	(5%)	5	(13%)	0	-	4
Recombination artifacts	2	(9%)	2	(5%)	22[*2]	(46%)	
Base mis-incorporation artifacts	2	(9%)	11	(29%)	0	-	
Total no. of clones	22		38		48		15
Cultivar 'Schneverdinger Goldbirke'							
Primer combination	BpI		BpII		BpI		No. of different
no. of PCR cycles	26 cycles		26 cycles		30 cycles		alleles included
Confirmed sequences	21	(53%)	13	(43%)	21	(43%)	10
Unique sequences	0	-	3	(10%)	3	(6%)	4
Recombination artifacts	12	(30%)	8	(27%)	15	(29%)	
Base mis-incorporation artifacts	7	(18%)	6	(20%)	11	(22%)	
Total no. of clones	40		30		51		14

1 one sequence had both a recombination and a base mis-incorporation artifact
2 two sequences had both a recombination and a base mis-incorporation artifact

Several of the *PR-10* sequences from *B. pendula* that were available from GenBank/EMBL/DDBJ also showed clear evidence for recombination (Fig. 4.1c). However, it was not possible to do a similar check as mentioned above for the GenBank sequences, and we can only hypothesize on the presence of artifact sequences without such direct evidence. Given the regular occurrence of artifact sequences in our experiments, we maintained only those 40 GenBank sequences that were not under suspicion of recombination artifacts for further analysis.

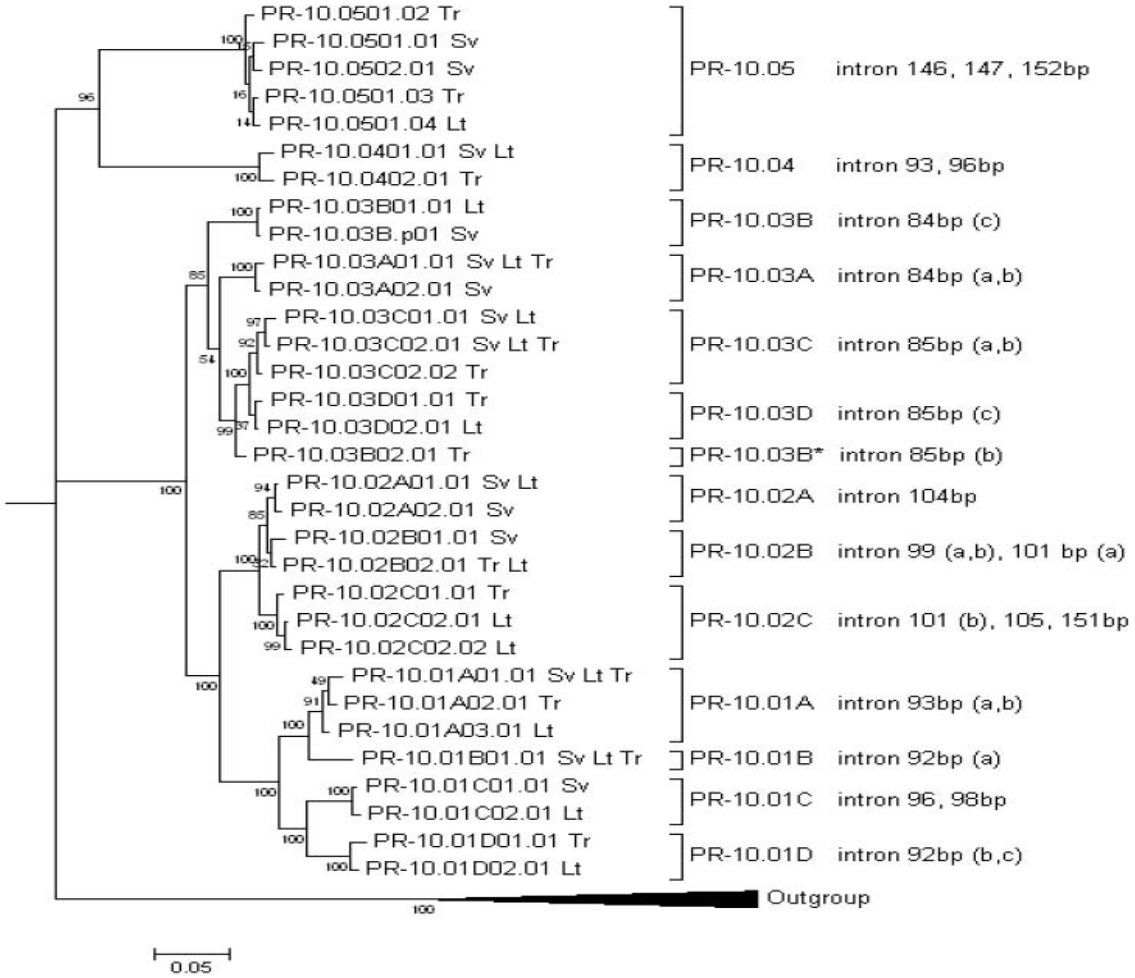
Phylogenetic analysis: newly isolated sequences

We sequenced 404 individual clones in both directions (Table 4.1). Fourteen different sequences were identified in the diploid cultivar 'Schneverdinger Goldbirke', seven of which were unique for this cultivar. Fifteen different sequences (of which 10 were unique) were identified in the diploid cultivar 'Tristis'. Approximately twice as many different sequences, namely 28, were identified in the tetraploid cultivar 'Long Trunk'; of these, 16 sequences were of *B. pendula* origin (eight unique), and 12 sequences were from another *Betula* species (all unique). In total, 32 different sequences were found. The Open Reading Frame (ORF) of the sequences was highly conserved and the alignment was straightforward. All but one ORF contained 483 nucleotides, coding for a putative protein that is 160 amino acids long. One sequence from 'Schneverdinger Goldbirke' required the inclusion of an indel between base 388 and 389 of the consensus. This sequence was denoted as a pseudogene, since the indel introduced a stop codon at seven codons downstream. It cannot be excluded that this pseudogene is expressed, since the stop codon was located near the 3' end. The intron position was identical in all sequences and located at codon 62, being inserted between the first and second nucleotide. Most alleles had 5' splicing sites of AG:GT, with the exception of one allele that had a GG:GT splicing site. The 3' splicing sites were AG:GC or AG:GA. This is in concordance with known motifs for plant introns. The introns were relatively AT-rich (55-65%).

We determined the relationships among the *PR-10* sequences from *B. pendula*. For this, the 'Long Trunk' sequences that were designated to the other parental species were excluded. Excluding primer traces, 171 of the 452 aligned exon positions were variable, while 150 positions were phylogenetically informative. The consensus tree from the Bayesian analysis indicated several well-defined clusters (Fig. 4.2). We implemented a cut-off level of 98% identity and allowed maximally two alleles per cultivar per gene to estimate the number of genes. In this way, we putatively identified ten genes in 'Schneverdinger Goldbirke', eleven in 'Tristis', and thirteen in 'Long Trunk'. Thirteen different genes were distinguished when the information was combined. Each gene was identified as such in at least two birch cultivars and was characterized by a distinct intron, in most cases of a different size between 84 and 152 bp (Fig. 4.2). The similarity between different alleles of one gene ranged from 98.9 to 100% identity in the exons, which corresponds to 0-5 SNPs. *Bet v 1.02A* and *Bet v 1.02B* were not well distinguished in the coding sequences, but

had distinguishable introns. The pseudogene *PR-10.03B-p01* from 'Schneverdinger Goldbirke' was identical to the *PR-10.03B* allele from 'Long Trunk' except for its indel.

Figure 4.2 – Bayesian phylogenetic tree of the PR-10 sequences from B. pendula*. Bayesian phylogenetic tree of the PR-10 sequences from* B. pendula *'Schneverdinger Goldbirke' (Sv), 'Tristis' (Tr), and the* B. pendula *alleles from 'Long Trunk' (Lt). The 'Long Trunk' alleles that belong to the unknown parental species are not included in this figure. Numbers on the branches represent posterior probabilities after running a Markov chain Monte Carlo search for 1,000,000 generations. Sequences of PR-10 genes from* Malus domestica *(apple, X83672, Z72425, Z72427),* Prunus armeniaca *(apricot, AF020784), P.* avium *(cherry, U66076), and* Pyrus communis *(pear, AF057030) were used as outgroup. Each cluster that is identified as a putative gene has maximally two alleles per cultivar. Genes are classified into five major groups. The intron length is indicated on the right. If multiple introns of the same length exist within one group, the different types are shown between brackets.*



* PR-10.03B02.01 from 'Tristis' was an in vivo recombination of the PR-10.03D gene and the original PR-10.03B gene

*TABLE 4.2 – Classification and nomenclature of B. pendula PR-10 sequences from our cultivars and GenBank. Indicated are the subfamily (01 to 05), gene designations (*PR-10.01A *to* PR-10.05*), allergen designation if the genes are known to be pollen expressed (*Bet v 1.01A *to* Bet v 1.02C*), and allele names as defined in Table 4.2. Known isoforms (http://www.allergen.org) are shown, followed by the GenBank accession number (between brackets). The tissue of origin is shown in case of mRNA-derived GenBank sequences (L=leaves, R=roots, P=pollen).*

Sub-family	Gene	Allergen	Allele (GenBank no.)[*1]	Known isoforms (GenBank no.)[*1]	Location of transcription	Sequence identity to reference sequence[*2]
01	PR-10.01A	Bet v 1.01A	Bet v 1.01A01.01[*3] (DQ296566)	Bet v 1a=Bet v 1.0101 (X15877), - (AJ001553)	P, P	100%
			Bet v 1.01A02.01 (DQ296567)		-	99.1 %
			Bet v 1.01A03.01 (DQ296568)		-	99.1 %
				Bet v 1.1501 (Z72429)	-	99.8 %
				Bet v 1.1502 (Z72432)	-	99.8 %
				Bet v 1.0102 (Z80098)	P	99.8 %
				Bet v 1.0103 (Z80099)	P	99.8 %
				Bet v 1.2501 (Z80101)	P	98.8 %
				Bet v 1.2801 (Z80104)	P	99.5 %
				Bet v 1.3001 (Z80106)	P	99.8 %
				- (AF124838)	-	98.6 %
				- (AJ002107)	P	99.8 %
				- (AJ002109)	P	99.3 %
				- (AJ006905)	P	99.8 %[*4]
				- (AJ006911)	P	98.6 %
				- (AJ006913)	P	99.5 %[*4]
	PR-10.01B	Bet v 1.01B	Bet v 1.01B01.01 (DQ296569)	- (AJ001552), - (AJ002106)	P, P	100 %
				Bet v 1d=Bet v 1.0401 (X77266)	P	99.3 %
				Bet v 1h=Bet v 1.0402 (X77270)	P	99.8 %
				- (AJ001551)	P	99.8%
	PR-10.01C	Bet v 1.01C	Bet v 1.01C01.01 (DQ296570)		-	100 %
			Bet v 1.01C02.01 (DQ296571)		-	99.8%
				Bet v 1f=Bet v 1.0601 (X77268)	P	99.8 %
				Bet v 1i=Bet v 1.0602 (X77274)	P	99.8 %
				- (AJ001557)	P	99.5 %
	PR-10.01D	Bet v 1.01D	Bet v 1.01D01.01 (DQ296572)	- (AJ001555)	P	100%
			Bet v 1.01D02.01 (DQ296573)		-	99.1 %
				Bet v 1.1701 (Z72430)	-	98.4%
				- (AF124839)	-	98.4%

02	PR-10.02A	Bet v 1.02A	Bet v 1.02A01.01 (DQ296574)		-	100 %
			Bet v 1.02A02.01 (DQ296575)		-	99.8 %
				- (AJ001554)	P	99.8 %
				- (AJ001556)	P	99.8%
	PR-10.02B	Bet v 1.02B	Bet v 1.02B01.01 (DQ296576)		-	100 %
			Bet v 1.02B02.01a[*5] (DQ296577)		-	99.3 %
			Bet v 1.02B02.01b (DQ296578)		-	99.3 %
				Bet v 1.1801 (Z72431)	-	99.1 %
	PR-10.02C	Bet v 1.02C	Bet v 1.02C02.01 (DQ296580)	Bet v 1k=Bet v 1.0901 (X77272)	P	100 %
			Bet v 1.02C02.02 (DQ296581)		-	99.8 %
			Bet v 1.02C01.01 (DQ296579)		-	99.8 %
				Bet v 1.20101 (Z72434)	-	99.3 %
				Bet v 1c=Bet v 1.0301(X77265)	P	99.8 %
				Bet v 1.1901 (Z72433)	-	98.6 %
				Bet v 1m=Bet v 1.1401 (X81972)	P	99.3 %
				Bet v 1n=Bet v 1.1402 (X82028)	P	99.3 %
03	PR-10.03A	-	PR-10.03A01.01 (DQ296582)		-	100 %
			PR-10.03A02.01 (DQ296583)		-	99.8 %
	PR-10.03B	-	PR-10.03B01.01 (DQ296584)		-	100 %
			PR-10.03B-p01[*5] (DQ296586)		-	100%[*4]
	PR-10.03B[*6]		PR-10.03B02.01 (DQ296585)		-	96.5 %
	PR-10.03C	-	PR-10.03C01.01 (DQ296587)	- (AJ289771)	-	100 %
			PR-10.03C02.01 (DQ296588)		-	99.5 %
			PR-10.03C02.02 (DQ296589)		-	99.3 %
				Bet v 1.1201 (X77600)	L,R	99.5 %
	PR-10.03D	-	PR-10.03D01.01 (DQ296590)		-	100 %
			PR-10.03D02.01 (DQ296591)		-	99.8 %
				Bet v 1.1101 (X77599)	L,R	99.5 %
04	PR-10.04	-	PR-10.0401.01 (DQ296592)		-	100 %
			PR-10.0402.01 (DQ296593)		-	98.9 %
05	PR-10.05	-	PR-10.0501.03 (DQ296596)	Bet v1 1.1301 (X77601), - (AJ289770)	L,R	100 %
			PR-10.0501.02 (DQ296595)		-	99.5 %
			PR-10.0501.01 (DQ296594)		-	99.8 %
			PR-10.0501.04 (DQ296597)		-	99.8 %
			PR-10.0502.01 (DQ296598)		-	99.8 %

*1 The known mRNA-derived GenBank sequences contain no intron, while the new gDNA sequences do, aiding in the gene identification ; *2 The upper most allele was taken as a reference sequence, identities are calculated for an aligned stretch of 425 bp from base 28 to 452 of the consensus; *3 The last two numerals indicate silent mutations (see Results section for further explanation of the nomenclature); *4 These sequences contain an indel and sequence identity is calculated excluding the indel; *5 Pseudogene allele; * 6 I*n vivo* recombination.

When all alleles from 'Long Trunk' were included, 182 variable positions were identified among the 452 aligned exon positions. 154 positions were phylogenetically informative. The topology of the consensus tree from the Bayesian analysis showed that seven specific 'Long Trunk' genes were clustered pair wise to the *B. pendula* genes and these are likely to represent orthologuous genes from the second *Betula* species (not shown). Given the high identity (up to 100%) to PR-10 sequences from *B. ermanii*, this species, or a close relative, is likely to represent the second parental species (unpublished data, Schenk *et al.*). Intron sequences of orthologuous genes mostly showed slight differences in length or base pair composition. Three genes were recovered only from the unknown parent species and five only from *B. pendula*, indicating that *Betula* species do not necessarily have the same (number of) genes.

Phylogenetic analysis: GenBank sequences versus newly obtained sequences

All but two GenBank sequences had an ORF of 483 base pairs that coded for 160 amino acids and was generally interrupted by a position-conserved intron. Two sequences required the inclusion of an indel, following base 354 of the consensus sequence. This resulted in a stop codon directly after the insertion. The GenBank dataset was combined with the newly obtained sequences. Primer traces were discarded, resulting in an aligned stretch of 425 bp from base 28 to 452 of the consensus. Several GenBank sequences are identical at this stretch, reducing the number of unique GenBank sequences in the analysis to thirty-three. 173 out of 425 aligned positions were variable, while 146 were phylogenetically informative. The information from the Bayesian consensus tree was added to Table 4.2 by indicating in which gene cluster the GenBank alleles landed. We used a cut-off of 98% identity to analyse the newly isolated sequences. In the resulting classification, 35 out of 40 alleles that were assigned to a particular gene showed more than 99% identity for the exons. The lowest similarity of an allele that still clustered with a particular gene in the phylogenetic analysis was 98.4% identity with the reference sequence (Table 4.2).

We classified the genes into five subfamilies (01-05) based on identities of the coding regions and the intron. The average identity between alleles within each subfamily was 95-100%. The GenBank sequences were in part derived from RNA extractions from specific tissues. We used this tissue information to predict the transcription of the *PR-10* genes (Table 4.2). For this, alleles of a single gene are assumed to have the same mode and location of transcription. Subfamily 01 and 02 consist of respectively four and three genes and include alleles that are homologous to the pollen-expressed Bet v 1 allergens (Breiteneder *et al.* 1989; Swoboda *et al.* 1995b). For four genes within these subfamilies, we found alleles that were 100% identical to pollen mRNA-derived sequences previously deposited into GenBank. Alleles from two other genes were 99.8% identical (1 SNP difference) to pollen mRNA-derived sequences from GenBank, which we take to predict the location of transcription for these six genes with a very high level of confidence. For only one of the genes in subfamily 02, there was no mRNA-derived homologue in the GenBank database, but

the high homology (97.8-98.2%) to the other genes in this subfamily suggests that this gene will be expressed in pollen as well.

Subfamily 03 consists of four genes, two of which have alleles that are 99.5% identical to homologues of the previously described *ypr10a* and *ypr10b* sequences, which are transcribed in roots and leaves (Swoboda *et al.* 1995c). Transcription of the other two genes in subfamily 03 is unknown, as is the transcription of the single gene in subfamily 04. Subfamily 05 consists of a single gene. One of the recovered alleles was 100% identical to the previously described *ypr10c* sequence that is also expressed in roots and leaves (Swoboda *et al.* 1995c). Given the fact that all sequences in subfamily 03-05 are less than 90% homologous to pollen-derived mRNAs and that there is no evidence of expression in pollen, we consider these as non-allergens.

Nomenclature

All recovered *B. pendula* alleles were added to Table 4.2. This table also lists all known GenBank accessions of Bet v 1 sequences that were not under suspicion of artifacts. If sequences were previously named, we included a cross-reference to the original nomenclature (Bet v 1a-n) as well as to the nomenclature of the allergen nomenclature committee (Bet v 1.0101 to 1.3101) [http://www.allergen.org/]. Gao *et al*. (2005a; 2005b) designed a nomenclature system for the Mal d 1 to Mal d 4 allergens that differentiates between alleles from different genes. We have adopted their system in Table 4.2. Protein sequences that have less than 95% identity are designated according to the current allergen nomenclature. A Latin letter is added to the iso-allergen name for those genes that have more than 95% identity. Two numerals are added for each allelic variant at the protein level. Two additional numerals indicate silent mutations. For example, *Bet v 1.02C02.02* refers to a silent mutation in the second protein variant of the *Bet v 1.02C* gene, which shares more than 95% identity with *Bet v 1.02A* and *B*. The *PR-10* sequences that were not allergens (not pollen-expressed) are named in a similar fashion, except that names started with *PR-10* instead of *Bet v 1*. If future evidence would indicate that a particular PR-10 gene is expressed in pollen after all, the name can easily be modified by replacing the PR-10 tag with Bet v 1.

PR-10 proteins and allergenicity

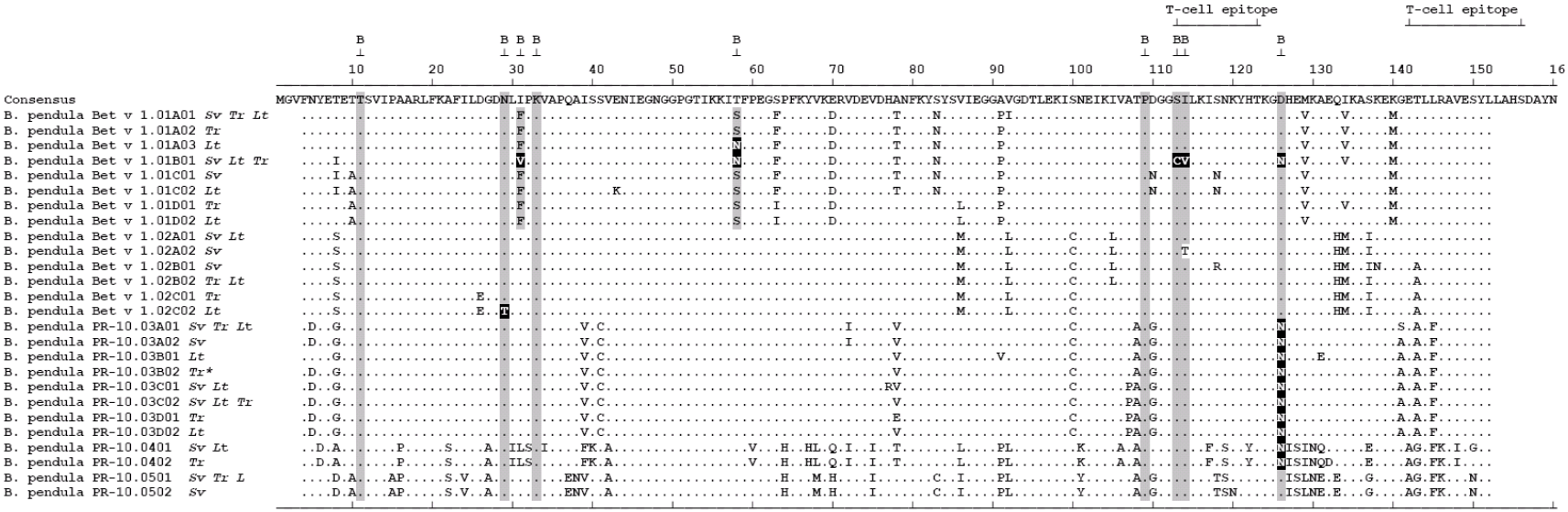
A high similarity between proteins increases the chance that they share epitopes, while on the other hand, a single amino acid change may influence allergenicity drastically. The high homogeneity among the PR-10 genes of *B. pendula* was reflected by a higher allelic variation at the nucleic acid level compared to the protein level. Hence, the 45 different genomic sequences encoded 32 different putative isoforms (Fig. 4.3). This is consistent with the relatively large number of non-synonymous mutations compared to the number of synonymous mutations (K_a/K_s ratios) (Rozas *et al.* 2003). The number of non-synonymous (K_a) and synonymous (K_s) substitutions

per site were calculated from pair wise comparisons of all alleles from the three *B. pendula* varieties. The average value was 0.080 for K_a and 0.247 for K_s, resulting in an average K_a/K_s ratio of 0.33 (*n*=45). Analysis of the occurrence of non-synonymous mutations per codon indicated two interesting regions. The region between codon Asn^{42} and Ile^{56} lacked non-synonymous mutations and is characterized by a phosphate-binding loop with the sequence motive GxGGxGx (Fig. 4.3). Relatively many amino-acid differences were present beyond codon 125, especially between isoforms of the genes *PR-10.04/PR-10.05* and *PR-10.01* to *PR-10.03* (Fig. 4.3).

Previous research has identified isoforms with varying IgE reactivity within mixtures of pollen. In fact, this is also true for pollen from a single tree. Ferreira *et al*. (1996) made a distinction between isoforms with high, intermediate and low IgE-binding activity. The high IgE-binding isoform Bet v 1a (X15877) clustered with alleles of gene *Bet v 1.01A* and was 100% identical at the protein level to the Bet v 1.01A01 isoform. The intermediate IgE binding isoforms Bet v 1c and f (X77265, X77268) clustered with alleles of the genes *Bet v 1.01C* and *Bet v 1.02C* and differed by one amino acid from the Bet v 1.01C01 and the Bet v 1.02C02 isoforms. The low IgE binding isoform Bet v 1d (X77266) clustered with alleles of gene *Bet v 1.01B*, and was 100% identical at the protein level to the Bet v 1.01B01 isoform. Nine amino acids have been identified that affect the allergenicity of the Bet v 1 proteins on the B-cell level (Ferreira *et al.* 1996; Son *et al.* 1999). These are marked in Fig. 4.3. The tetraploid 'Long Trunk' contained several isoforms with unique amino acid substitutions due to its putative hybrid origin. This greatly enlarges the variation in putative IgE-reactivity among isoforms of this cultivar.

Two major T-cell binding epitopes have been identified, which are positioned between amino acids 112-123 and 142-156 (Jahn-Schmid *et al.* 2005). Activation of T lymphocytes with the Bet v $1_{142\text{-}156}$ epitope induced cytokine production (IL-4, IL-5). The major T-cell epitope Bet v $1_{142\text{-}156}$ has an amino acid variation at position Thr^{143}. At this position the genes *Bet v 1.02B* and *Bet v 1.02C* code for Ala^{143} instead. We identified an additional variation in the *Bet v 1.01D* gene from the tetraploid 'Long Trunk', namely the presence of Arg^{150} instead of Ser^{150}. The T-cell epitope Bet v $1_{112\text{-}123}$ is more variable and contains amino acid variations in the genes *Bet v 1.01B*, *Bet v 1.01C*, *Bet v1.02C* and *Bet v 1.02D* (Fig. 4.3).

Figure 4.3 – Amino acid sequences, amino acids that affect IgE-reactivity, and T-cell epitopes of the PR-10 proteins. Amino acid sequences of the PR-10 proteins from B. pendula 'Tristis' (Tr), 'Schneverdinger Goldbirke' (Sv), and the B. pendula alleles from 'Long Trunk' (Lt). Amino acids associated with high allergenicity are marked with grey boxes and those associated with low IgE-reactivity (located within B-cell epitopes) are marked with black boxes (Ferreira et al., 1996; Son et al., 1999). The locations of the two major T-cell activating regions are indicated above the consensus (Jahn-Schmid et al., 2005.



Discussion

PCR artifacts

When PCR amplification is performed on groups of closely related sequences, such as the PR-10 gene family, accurate sequences are essential to distinguish between members. We used *Pfu* polymerase, which has proofreading functionality and reduces the number of base substitution error rates. However, previous research has shown that *Pfu* polymerase may generate more and more complex recombination artifacts than *Taq* polymerase (Whinnett *et al.* 2003; Zylstra *et al.* 1998) through incomplete primer extension and re-annealing to a different template (Judo *et al.* 1998), or strand switching between different templates (Odelberg *et al.* 1995). Reducing the number of cycles (Zylstra *et al.* 1998) was an efficient solution to lower the amount of artifacts.

We identified PCR recombination artifacts in several of our sequences using the computer program Phylpro (Weiller 1998). The comparison of independent PCRs enabled sequence validation and exclusion of both recombination and base substitution artifacts. The high occurrence of recombination artifacts in our experiments (27-46% after 30 PCR cycles) is not uncommon. For example, Wang *et al.* (1996) report 32% recombination artifacts after 30 cycles of PCR amplification for 16S rRNA genes. The occurrence of recombinant sequences within a mixture of approximately 15 different sequences is expected to be high, because almost all recombinations are detected. Half of the recombinations would, for example, remain undetected if only two templates are present. The presence of 13 genes also increases the amount of template compared to a single copy gene.

Using the same analysis to detect recombination in GenBank/EMBL/DDJB accessions, we observed a recombination signal in 22 out of 62 (35%) accessions. These are not necessarily all PCR artifacts, as some recombinations may have occurred *in vivo* during evolution of the genes. However, we found only one true recombinant in our dataset and no evidence of past recombination events in the comparison between sequences from two different species (within the hybrid). This indicates that the occurrence of *in vivo* recombination is probably rare.

Base substitution error rates for mixtures of non-proofreading *Taq* and proofreading *Pfu* are approximately 5.6×10^{-6} under optimal conditions (Cline *et al.* 1996). This error rate is ~2-4 fold higher when only *Taq* is used (Acinas *et al.* 2005; Cline *et al.* 1996). Without confirmation in independent PCRs these errors can not be excluded and it is very likely that Bet v 1 sequences with base substitution errors have been deposited into GenBank/EMBL/DDJB. Therefore, not all published isoforms will be clinically relevant. In addition, the clinical relevance of the isoforms will be influenced by their expression levels. If multiple allergen isoforms exist, there is a risk for selecting a recombinant isoform with low IgE-reactivity as a diagnostic tool, or even selecting isoforms that resulted from PCR artifacts. We therefore strongly suggest the use of primers that

are highly specific for one gene (Gao *et al.* 2005a) or, preferentially, the application of multiple independent PCRs to facilitate sequence validation in future sequencing work on allergens.

PR-10 and Bet v 1 genes

The PR-10 gene family of *Betula pendula* was shown to encompass at least thirteen genes. This is a conservative estimate since we used strict inclusion criteria. The distinction between genes is supported by the presence of a distinct intron. Each gene was identified as such in at least two birch cultivars. We attributed previously described GenBank sequences to these genes. Alleles from ten identified genes had previously been described (Breiteneder *et al.* 1989; Friedl-Hajek *et al.* 1999; Hoffmann-Sommergruber *et al.* 1997; Son *et al.* 1999; Swoboda *et al.* 1995b), while we identified three new genes. The genes are grouped into five subfamilies, based on sequence homologies in the ORF and intron. Differences in transcription coincide with the division between subfamilies. An organization of *PR-10* genes into subfamilies was also reported for *Malus domestica* (Gao *et al.* 2005a) and for *Pinus monticola* (Liu *et al.* 2004).

A striking feature of the *PR-10* isoforms in *B. pendula* is their homogeneity, which may extend to other Fagales species, such as alder *(Alnus glutinosa)* and hazel *(Corylus avellana)*. The intron has conserved 3' and 5' splicing sites and is always located at codon 62, as is *e.g.* reported for *C. avellana* and *M. domestica* (Hoffmann-Sommergruber *et al.* 1997). High homogeneity may result from strong purifying selection or from concerted evolution. The presence of low K_a/K_s ratios among the isoforms suggests the occurrence of purifying selection. Evidence for concerted evolution is present in the overall gene tree of the PR-10 family (Wen *et al.* 1997). Concerted evolution causes genes to evolve as a single unit with members exchanging genetic information through gene conversion and unequal crossing-over. Tandemly arranged genes may have high conversion rates (Ohta 2000), while this is a prerequisite for the occurrence of unequal crossing-over. Most *PR-10* genes in apple map to two loci and are arranged in a duplicated cluster (Gao *et al.* 2005a). This organization may be a common feature for PR-10 genes. However, as pointed out by Nei and Rooney (Nei *et al.* 2005), the molecular mechanism of gene conversion is not well understood, and the model of birth-and-death evolution of genes may also explain the evolution of the PR-10 gene family. The presence of pseudogenes, although at a low frequency, is therefore of particular interest. An analysis of a species at an intermediate evolutionary distance, such as *C. avellana,* would be useful to clarify which mechanisms determine the evolution of PR-10 genes and to investigate a possible recent radiation of PR-10 genes.

The birch genome contains at least seven pollen-expressed genes that encode a mixture of Bet v 1 isoforms with varying IgE-reactivity. Swoboda *et al*. (1995b) found that pollen mixtures from multiple trees contain multiple Bet v 1 isoforms. We identified 14 different Bet v 1 isoforms in the three cultivars, nine of which were entirely new. The IgE-reactivity has been tested for several

isoforms using recombinant proteins (Ferreira *et al.* 1996). The allergenicity of the new isoforms can be examined in the future by expressing the isoforms as recombinant proteins and use these in a SPT or T-cell activation tests. Ferreira *et al.* (1996) divided the Bet v 1 isoforms into three groups according to their IgE-reactivity and confirmed the division between high, moderate, or low IgE-reactivity in a SPT. One high and one low IgE-reactive isoform from their analysis were 100% identical to isoforms that we have obtained from a single tree, while two intermediate IgE-reactive isoforms differed only by one amino acid from the alleles of two other identified genes. This strongly suggests that isoforms of different IgE-reactivity are in fact alleles encoded by different genes. Thus, each examined cultivar has the genetic background to express a mixture of isoforms with a high, moderate, and low IgE-reactivity. We plan to confirm this at the protein level in the near future.

Nomenclature

The nomenclature of Bet v 1 raises several issues. The first isoforms were termed Bet v 1a to Bet v 1n by Swoboda *et al*.(1995b), but these have subsequently been renamed and incorporated into the official database of the allergen nomenclature committee [http://www.allergen.org/]. This database currently lists 37 allergen isoforms that have been termed Bet v 1.0101 to 1.3101. However, as can be seen from the list of known isoforms in Table 4.2, not all published isoforms have been added to this list, even though several of these isoforms were obtained from pollen mRNA. On the other hand, isoforms which have been recovered only from mRNA from roots and leaves are included as allergens (Bet v 1.1101 to Bet v 1.1301). Also, several of the described isoforms are highly suspicious as we observed clear recombination signals. As a result, the list is a random series of alleles that belong to different genes and has no biological basis. Similar problems were described for the Mal d 1 and Mal d 3 allergens by Gao *et al*. (2005a; 2005b). To allow differentiation between alleles from different genes for the Bet v 1 alleles we have adopted their system.

Allergenicity of birch trees

The exact isoform composition of the three cultivars differed due to allelic variations. This may result in differences in allergenicity between cultivars. However, an exact copy of the most allergenic allele, *Bet v 1a,* was present in all three cultivars. Quantity measurements on expression indicate that Bet v 1a is the dominant isoform in pollen (Swoboda *et al.* 1995b). Given that the diversity of *Bet v 1* isoforms within a single tree is larger than the diversity between the examined *B. pendula* cultivars, a characterization of *Bet v 1*-type isoforms should be done in other *Betula* species as well. Investigations towards sensitization and immunotherapy should anticipate that patients are exposed to a mixture of Bet v 1 isoforms of different IgE-reactivity, even if pollen originates from a single birch tree. Differences in allergenicity between birch trees may also result

from variation in allergen content. Variation in allergen content has *e.g.* been shown for apple (Bolhaar *et al.* 2005a) and olive pollen (Castro *et al.* 2003).

Many Betula species and Betula hybrids have higher ploidy levels (tetraploid, hexaploid, and even octaploid) than *B. pendula* and are likely to contain increased numbers of allergen isoforms, as we found in the tetraploid cultivar 'Long Trunk'. For example, the tetraploid *B. pubescens* is dispersed throughout Europe, while other exotic birch species are increasingly introduced as cultivars, contributing to a larger allergen pool. However, interspecific *Betula* hybrids, which have a higher ploidy level, may also pose a potential source of hypoallergenic trees. Especially hybrids between less related species often display a reduced fertility, which may result in a reduced or aborted pollen production.

If breeding for hypoallergenic trees is implemented, approaches should take into account that the Bet v 1/PR-10 genes may be clustered and differences in allergenicity between clusters of genes may be used to guide breeding efforts. Clustering may be determined by mapping studies (Gao *et al.* 2005a) and by screening and partial sequencing of a genomic library. In the genome sequence of *Populus* [http://www.ornl.gov/sci/ipgc] we can indeed observe an organization of PR-10 genes into clusters. Other approaches to generate hypoallergenic trees may include the search for sterile or low pollen producing trees, or the application of RNA interference technology, which proved to be successful to silence the Mal d 1 allergens in apple (Gilissen *et al.* 2005) without phenotypic abnormalities. One issue that needs the be resolved is that although birch PR-10/Bet v 1 proteins have been suggested to act as plant steroid carriers (Markovic-Housley *et al.* 2003), the exact conditions under which transcription is induced are still unknown. Given the abundance of Bet v 1 in birch pollen, silencing may affect pollen viability. However, for breeding of hypoallergenic cultivars that are propagated vegetatively this would be considered a welcome side-effect.

Conclusion

We have shown that the PR-10 gene family of *Betula pendula* encompasses at least thirteen genes that can be grouped into five distinct subfamilies. Differences in expression coincide with the division between subfamilies. Genes from two subfamilies were shown to be transcribed in pollen, based on a high (99.8-100%) homology with cDNA sequences available in GenBank/EMBL/DDJB. The seven genes that belong to these subfamilies encode a mixture of Bet v 1 isoforms of varying IgE-reactivity. The present characterization of the PR-10 family in birch provides a framework for the screening of *Bet v* 1 genes among other *Betula* species or *B. pendula* cultivars and for potential breeding approaches for birch trees with a reduced allergenicity.

CHAPTER 5

Phylogenetic relationships in Betula (Betulaceae) based on AFLP markers

Martijn Schenk, Claire-Noëlle Thienpoint, Wim Koopman,
Luud Gilissen and René Smulders

Submitted to Tree Genetics and Genomes as: MF Schenk, C-N Thienpoint, WJM Koopman, LJWJ Gilissen, MJM Smulders. Phylogenetic relationships in *Betula* (Betulaceae) based on AFLP markers.

Abstract

The genus *Betula* comprises various species in boreal and temperate climate zones of the Northern Hemisphere. The taxonomy of *Betula* is controversial and complicated by parallel evolution of morphological traits, polyploidisation events, and extensive hybridisation and introgression among species. Multilocus molecular data from AFLPs were used to provide phylogenetic information. A large number of polymorphic markers (321 variable bands) were produced in 107 *Betula* accessions from 23 species and 11 hybrids. The AFLP results were largely congruent with the results from previously examined nuclear DNA markers. Four distinct subgenera were identified within the genus *Betula*. These subgenera were partly in disagreement with the traditional (but disputed) division of the genus. In addition, the results indicated several groups of conspecific taxa. The majority of the species fell within subgenus *Betula* and shared a high degree of similarity with *B. pendula*. All hybrids were associated with this group, and the AFLP data contained signal on putative parents for some of the interspecific hybrids. Subgenus *Chamaebetula* and part of the *Neurobetula* species should be merged with *Betula*. The subgenera *Betulenta*, *Betulaster*, and the remaining part of *Neurobetula* are distinct and well-supported. Birch pollen is a major cause of hay fever complaints, and knowledge of the phylogenetic relationships may facilitate the prediction of allergenicity of *Betula* species. The species from subgenera *Betulenta*, *Betulaster* and *Neurobetula* may represent good candidates for screening for lower allergenicity.

Introduction

The genus *Betula* contains trees and shrubs from diverse habitats in boreal and temperate climate zones of the Northern Hemisphere. Estimates of the number of species range from 30 to 60 (De Jong 1993; Furlow 1990). The genus is placed within the Betulaceae family of the order Fagales. In Northern Europe, pollen of birch is a major cause of hay fever complaints, as is pollen of the Fagales species hazel and alder (Breiteneder *et al.* 1992; Breiteneder *et al.* 1989; Lüttkopf *et al.* 2002). The major allergen involved is Bet v 1 of which several variants exist that may differ in their allergenicity (Ferreira *et al.* 1996; Schenk *et al.* 2006). Given the socioeconomic impact of hay fever, birch represents a relevant target for the development of allergy prevention strategies. Selection and breeding of potential hypoallergenic birch trees requires knowledge on the genetic background of the available birch species, as the evolution of allergenic proteins is linked to the evolution of the species in which they are found. Therefore, phylogenetic relationships among *Betula* species may be used to predict allergenicity of birch species.

The taxonomy of *Betula* is controversial and various classifications have been proposed. The first monographer who provided an extensive review of the genus was Regel (1865), who divided the genus into subgenera *Betulaster* and *Eubetula*. Subgenus *Betulaster* contains only one section, the *Acuminatae*. Subgenus *Eubetula* comprises six sections, namely C*ostatae, Lentae*, *Nanae*, *Albae*, *Fruticosae* and *Dahuricae*. Winkler (1904) proposed a slightly different division, lowering the status of the sections to that of subsections, merging the *Fruticosae* and the *Dahuricae* with the subsection *Albae,* and merging the *Lentae* with the *Costatae*. More recently, De Jong (1993) proposed a division into five subgenera, namely *Betulenta*, *Betulaster*, *Neurobetula*, *Betula* and *Chamaebetula*. Subgenus *Betulenta* is considered the most primitive subgenus, followed by *Betulaster* and *Neurobetula*. *Neurobetula* is considered a very heterogeneous and partly artificial group (De Jong 1993). The subgenera *Chamaebetula* and *Betula* are considered to be more derived.

The basic chromosome number of *Betula* is *n* = 14, and the species form a series of polyploids with chromosome numbers of *2n* = 28, 56, 70, 84, 112 and 140 (Furlow 1990). Polyploidy is a common feature among *Betula* species, and its presence within at least four of the five recognised subgenera suggests several independent polyploidisation events. Hybridisation and introgression are common in situations where the natural distributions of birch species overlap, for example among the European birch species *B. pendula*, *B. pubescens* and *B. nana* (Palme *et al.* 2004). Moreover, several of the recognized *Betula* species have a hybrid origin (Nagamitsu *et al.* 2006). Hybrids generally show a morphology intermediate between the parental species, but are not always morphologically distinct as a group (Thórsson *et al.* 2001). This overlap in morphological features complicates species and hybrid identification. Introgression appears to be bidirectional (Williams *et al.* 2001), but asymmetrical (Palme *et al.* 2004). Hybridisation and introgression are

further facilitated by the introduction and distribution of artificially propagated cultivars outside the natural distribution range. The simultaneous occurrence of polyploidisation, extensive hybridisation and introgression, complicates taxonomical studies in the genus *Betula*. In addition, several morphological characters are likely to have evolved independently more than once, or have experienced parallel evolution (Li *et al.* 2005).

Given the difficulties with morphological characters in reconstructing species relationships within the genus *Betula*, alternative markers were explored, *e.g.* flavonoid composition (Keinänen *et al.* 1999), nuclear DNA sequences (Järvinen *et al.* 2004; Li *et al.* 2005; Nagamitsu *et al.* 2006), a microsatellite (Nagamitsu *et al.* 2006), and chloroplast DNA sequences (Järvinen *et al.* 2004). These markers have provided useful information on the evolution of the genus, but relationships between species remain largely inconclusive due to their limited variation. For example, the chloroplast *matK* sequences examined by Järvinen *et al*. (2004) differentiated only three North American species from the other species. Moreover, trees constructed from different nuclear DNA markers display incongruences, *e.g.* nuclear ribosomal Internal Transcribed Spacers (ITS) versus microsatellite sequences (Nagamitsu *et al.* 2006), and ITS versus *ADH* sequences (Järvinen *et al.* 2004; Li *et al.* 2005).

In the present study, we examined the use of AFLP as an alternative for morphological markers, chloroplast DNA sequences and nuclear DNA sequences. AFLP is a DNA fingerprinting technique that generates large numbers of highly reproducible fragment markers with a genome-wide distribution. The technique is relatively fast and cost-efficient, and requires no prior knowledge of the genome (Jones *et al.* 1997; McGregor *et al.* 2000; Russell *et al.* 1997; Vos *et al.* 1995). Relative to morphological markers, AFLPs have the advantage that they are not under direct selection pressure, since most of the fragments represent non-coding parts of the genome (Vos *et al.* 1995). AFLPs have the advantage of being much more variable relative to chloroplast sequences (Koopman *et al.* 2008). Moreover, AFLPs represent both paternal and maternal lineages because they are almost entirely derived from the nuclear genome (Althoff *et al.* 2007). Compared to nuclear DNA sequences such as ITS, AFLPs have the advantage that they are more variable, and that are sampled across the entire genome rather than in a specific location (Koopman 2005). However, AFLPs also have drawbacks that potentially may hamper their use as phylogenetic characters (reviewed in Koopman, 2005), most notably a possible lack of homology between fragments across taxa (Althoff *et al.* 2007). Several studies have shown that homology assignment between AFLP fragments decreases with increasing evolutionary distance between taxa (Althoff *et al.* 2007; Koopman 2005). Koopman (2005) contrasted AFLP variation with ITS sequence divergence in a large number of taxa, and concluded that AFLPs are reliable phylogenetic markers for plant taxa with ITS sequences differing up to 30-35 nucleotides. A GenBank survey for the species in the present study revealed that ITS sequence differences among ingroup species ranged

from 0 to 22 nucleotides, which is well within the range defined by Koopman (2005). Therefore, it is expected that the AFLP marker variation in our data set is a suitable indicator of *Betula* relationships. Arens *et al.* (1998) and Cervera *et al.* (2005) demonstrated in poplar that the AFLP pattern of hybrid offspring contains bands of both parental species. Therefore, comparison of AFLP patterns of taxa may serve to identify hybrids.

The objectives of the present study were (1) to reconstruct the phylogeny of *Betula*, while positioning and identifying hybrid taxa and cultivars, and (2) to evaluate the (sub)sections proposed by Regel (1865) and Winkler (1904), and the subgenera proposed by De Jong (1993). The division of De Jong (1993) will be used as a starting point. Species from all sections and subgenera proposed by the abovementioned authors were included, as were several hybrid taxa.

Material and Methods

Plant material

We collected young leaves from 62 *Betula* accessions in the botanical collections of PPO Boskoop (Boskoop, The Netherlands), the Botanical Garden of Wageningen University (Wageningen, The Netherlands), and the Von Gimborn Arboretum (Doorn, The Netherlands). In addition, leaves were collected from 10 accessions of *B. pendula* in a birch seed orchard in Urk (The Netherlands), and 31 cultivated *Betula* accessions growing as lane trees in Ede (The Netherlands) and Munich (Germany) (Table 5.1). The accessions were originally attributed to 23 species and five interspecific hybrids based on descriptions and names available from the botanical collections. The phenetic analysis revealed that 11 accessions, cultivars mostly, did not group with the expected taxon. Nine suspected misclassified accessions were labelled as hybrids after evaluation of the ploidy levels (Table 5.1). The morphology of the remaining two accessions did not match with the taxon suggested by the original label, and in one of these accessions the ploidy level did not match either. Both accessions were tentatively assigned to the correct species (Table 5.1). All (sub)sections and subgenera proposed by Winkler (1904) and De Jong (1993) were represented by at least two species, except for subgenus *Betulaster* that was represented by a single species. Based on the results of Chen *et al.* (1999) and Li *et al.* (2005), two *Alnus* and two *Corylus* accessions were included as outgroup. Taxonomical names of (sub)sections and subgenera follow De Jong (1993).

Flow cytometry

Fresh leaf samples were sent to Plant Cytometry Services (Schijndel, The Netherlands) to determine the ploidy level. Ploidy levels were estimated by flow cytometry as described in Koopman (2000). Diploid (*B. pendula*) and tetraploid (*B. pubescens*) controls were included.

AFLP genotyping

For DNA extraction, young leaves of approximately 1 cm^2 were collected, immediately frozen in liquid nitrogen, and subsequently freeze-dried for storage. Total genomic DNA was extracted with the DNeasy 96 Plant Kit (Qiagen, Venlo, The Netherlands) from grinded leaf-tissue according to the manufacturer's instructions. The AFLP assay (Vos *et al.* 1995) was performed after digestion/ligation with the 6-bp cutting enzyme EcoRI and the 4-bp cutting enzyme MseI, followed by a two-step PCR amplification protocol (Arens *et al.* 1998) with the modification of using IRD700 fluorescence-labelled primers instead of ^{33}P-labelled primers. We used three selective primer combinations (Bonin *et al.* 2004): *Eco*RI 5'-GACTGCGTACCAATTC**AGT**-3'/*Mse*I 5'-GATGAGTCCTGAG TAA**CTC**-3', *Eco*RI 5'-GACTGCGTACCAATTC**ATG**-3'/*Mse*I 5'-GATGAGTCCTGAGTAA**CAC**-3', and *Eco*RI 5'-GACTGCGTACCAATTC**ATG**-3'/*Mse*I 5'-GATGAGTCCTGAGTAA**CAA**-3'. Amplified fragments were separated on 6.5% denaturing polyacrylamide gels and analyzed on a LI-COR 4300 DNA analyzer (LI-COR Biosciences, Lincoln, NE, USA). Three accessions failed to produce a scorable AFLP pattern due to incomplete digestion. The decaploid species *B. medwediewii* was excluded because it showed an excessive number of bands, which would hamper a reliable analysis.

AFLP data analysis

Li-COR TIFF images were imported into QUANTAR software (Keygene, Wageningen, The Netherlands). Two standard samples were run on each gel to allow automatic positioning of marker bands. Presence (1) or absence (0) of polymorphic AFLP bands was scored for all accessions in the range from 100 to 450 bp. Only intense and well-separated bands were scored. The primer combinations yielded 119, 113 and 89 AFLP markers, respectively (321 in total). Eight duplicate accessions were included as controls.

Several accessions were present in duplicate in our data set. The vast majority of these were identical, but occasionally one band was scored differently. The calculated Dice similarity was, however, always above the 98.5% limit that was indicated by Arens *et al.* (1998). All accessions displaying more than 98.5% similarity potentially represent clones and were removed from further analyses (Table 5.1). The phenetic analyses were conducted on a data set containing 87 unique accessions, including hybrids. Similarity matrices of Jaccard distances and Dice distances were calculated using NTSYSpc 2.10j (Applied Biostatistics). Dendrograms were subsequently constructed using neighbour-joining (NJ) analysis as well as UPGMA clustering. Co-phenetic matrices were calculated from the resulting dendrograms and the product-moment correlation between co-phenetic and similarity matrices was calculated to test the goodness of fit of the cluster analysis. The Mantel test for matrix correspondence was performed with 1,000 permutations. The combination with the best fit (Dice+NJ) was chosen for phenetic analysis. To allow bootstrapping, the analysis was repeated in Paup 4.0b10 (Swofford 2003) using Nei-Li distances (= 1 - Dice similarity) and 1,000 bootstrap replicates.

TABLE 5.1 - Plant Material: the birch accessions analyzed in this AFLP study.

	Accession Code [a]	Species	Subspecies, variety, or cultivar	Ploidy [b]	Subsection (Winkler, 1904)	Subgenus (De Jong, 1993)	Remarks
1	W003	*B. albosinensis*	*var. albosinensis*	4n	*Costatae*	*Neurobetula*	
2[d]	B004	*B. albosinensis*	Fascination	n.d.	-	"	
3	D202	*B. alleghaniensis*		6n	*Costatae*	*Betulenta*	
4	W005	*B. alleghaniensis*		6n	"	"	
5	W009	*B. chichibuensis*		2n	-	*Neurobetula*	
6	D085	*B. costata*		2n	*Costatae*	*Neurobetula*	
7	W012.3	*B. costata*		2n	"	"	
8	W013	*B. davurica*		8n	*Albae*	*Neurobetula*	
9	D092	*B. ermanii*	Blush	4n	*Albae*	*Neurobetula*	
10	W021	*B. ermanii*	*var. ermanii*	4n	"	"	
11[e]	D089	*B. ermanii*		4n	"	"	Original label: *B. papyrifera subsp. humilis*
12	W022	*B. grossa*		4n	*Costatae*	*Betulenta*	
13	W023	*B. humilis*		8n	*Nanae*	*Chamaebetula*	
14	D087	*B. korshinskyi*		4n	*Albae*	*Betula*	
15	D203	*B. lenta*	*subsp. lenta*	n.d.	*Costatae*	*Betulenta*	
16	W027.1	*B. lenta*	*subsp. lenta*	2n	"	"	
17	W029	*B. litwinowii*		4n	*Albae*	*Betula*	
18	D082	*B. maximowicziana*		2n	*Acuminatae*	*Betulaster*	
19	W032	*B. medwediewii*		10n	*Costatae*	*Betulenta*	
20	B003	*B. nana*		2n	*Nanae*	*Chamaebetula*	
21	B002	*B. nigra*		2n	*Costatae*	*Neurobetula*	
22	D093	*B. nigra*		2n	"	"	Identical to B002
23	D097	*B. nigra*		2n	"	"	Identical to B002
24	E021.1	*B. nigra*		2n	"	"	Identical to B002
25	E021.2	*B. nigra*		2n	"	"	Identical to B002
26	E022.2	*B. nigra*		2n	"	"	Identical to B002
27	W037	*B. papyrifera*		6n	*Albae*	*Betula*	
28	W038	*B. papyrifera*		6n	"	"	
29	W044	*B. papyrifera*	*subsp. cordifolia*	6n	"	"	
30	W047	*B. papyrifera*	*var. commutata*	4n	"	"	
31	M040	*B. papyrifera*		n.d.	"	"	
32[e]	W075	*B. papyrifera*		n.d.	"	"	Original label: *B. populifolia*
33	B006	*B. pendula*	Laciniata	2n	*Albae*	*Betula*	
34	D094	*B. pendula*	Fastigiata	2n	"	"	
35	D095	*B. pendula*	Tristis	2n	"	"	
36	W058	*B. pendula*	Youngii	2n	"	"	
37	W059	*B. pendula*	Youngii	2n	"	"	Identical to W059
38	D096	*B. pendula*	Obelisk	2n	"	"	
39	W051	*B. pendula*	Dalecarlica	2n	"	"	
40	W057.2	*B. pendula*	Tristis	2n	"	"	
41	E015.1	*B. pendula*	Youngii	2n	"	"	Identical to W059
42[c]	E016	*B. pendula*		2n	"	"	
43	E017	*B. pendula*		2n	"	"	
44	E019	*B. pendula*		2n	"	"	
45	E020	*B. pendula*		2n	"	"	
46	E023	*B. pendula*		2n	"	"	
47	U001	*B. pendula*		n.d.	"	"	
48	U002	*B. pendula*		n.d.	"	"	
49	U003	*B. pendula*		n.d.	"	"	
50	U004	*B. pendula*		n.d.	"	"	
51	U005	*B. pendula*		n.d.	"	"	
52	U006	*B. pendula*		n.d.	"	"	
53	U007	*B. pendula*		n.d.	"	"	
54	U008	*B. pendula*		n.d.	"	"	
55[c]	U028	*B. pendula*		n.d.	"	"	
56[c]	U029	*B. pendula*		n.d.	"	"	
57	M001	*B. pendula*		n.d.	"	"	
58	M003	*B. pendula*		n.d.	"	"	
59	M004	*B. pendula*		n.d.	"	"	

60	M005	*B. pendula*		n.d.	"	"	
61[c]	M006	*B. pendula*		n.d.	"	"	
62	M007	*B. pendula*		n.d.	"	"	
63	M008	*B. pendula*		n.d.	"	"	
64[c]	M028	*B. pendula*		n.d.	"	"	
65[c]	M029	*B. pendula*		n.d.	"	"	
66	D084	*B. platyphylla*	*subsp. mandshurica*	2n	*Albae*	*Betula*	
67[c]	W060	*B. platyphylla*		2n	"	"	
68	W068	*B. platyphylla*	*subsp. szechuanica*	2n	"	"	
69	D086	*B. populifolia*		2n	*Albae*	*Betula*	
70	W080.2	*B. pubescens*	*subsp. tortuosa*	4n	*Albae*	*Betula*	
71[c]	E015.2	*B. pubescens*		4n	"	"	
72	M022	*B. pubescens*		n.d.	"	"	
73	M031	*B. pubescens*		n.d.	"	"	
74	M023	*B. pubescens*		n.d.	"	"	
75	M220	*B. pubescens*		n.d.	"	"	
76	M014	*B. pubescens*		n.d.	"	"	
77	D090	*B. pumila*	*subsp. pumila*	8n	*Nanae*	*Chamaebetula*	
78	W085	*B. schmidtii*		2n	*Costatae*	*Neurobetula*	
79	D091	*B. utilis*	*subsp. utilis*	4n	*Costatae*	*Neurobetula*	
80	D100	*B. utilis*	Doorenbos	4n	"	"	
81	E010	*B. utilis*	*subsp. jacquemontii*	4n	"	"	Identical to D100
82	E011	*B. utilis*	*subsp. jacquemontii*	4n	"	"	Identical to D100
83	M039	*B. utilis*	Doorenbos	4n	"	"	Identical to D100
84[d]	B012	*B. utilis*		n.d.	"	"	
85[d]	E009	*B. utilis*	*subsp. jacquemontii*	n.d.	"	"	
86	D098	*B. x caerulea*		n.d.	-	-	
87	W001	*B. x* 'Edinburgh'		4n	-	-	
88	D201	*B. x fetisowii*		n.d.	-	-	
89	W024	*B. x koehnei*		4n	-	-	
90	W036	*B. x obscura*		4n	-	-	
91[e]	B007	Unknown hybrid		4n	-	-	Original label: *B. ermanii*
92[e]	E013	Unknown hybrid		4n	-	-	Original label: *B. ermanii*
93[e]	E001	Unknown hybrid		4n	-	-	Identical to B007 Original label: *B. ermanii*
94[e]	W077	Unknown hybrid		4n	-	-	Identical to B007 Original label: *B. pubescens*
95[e]	D099	Unknown hybrid		4n	-	-	Original label: *B. papyrifera subsp. papyrifera*
96	W052	B. x 'Elegans pendula'		3n	-	-	Original label: *B. pendula* 'Elegans pendula'
97	W020	B. x 'Holland'		4n	-	-	Original label: *B. ermanii* 'Holland'
98	D088	Unknown hybrid		2n	-	-	Original label: *B. albosinensis*
99	B011	B. x 'Long Trunk'		4n	-	-	Original label: *B. pendula* 'Long Trunk'
100	W204	Alnus avellana	Aurea	-		*Corylus*	
101	W203	Alnus colurna		-		"	
102	W202	Corylus incana		-		*Alnus*	
103	W201	Corylus rubra	Oberon	-		"	

a Accession Number; W=Botanical garden Wageningen (the Netherlands (=NL)), B=Botanical garden Boskoop (NL), D=Botanical garden Doorn (the Netherlands), U=Seed orchard Urk (the Netherlands), E=Lane tree Ede (the Netherlands), M=Lane tree Munich (Germany).

b Ploidy level; n. d. = not determined

c Samples from these accessions were run in duplicate

d Accessions excluded because of incomplete digestion

e Accessions in which the morphology conflicted with the original label

The relationships among species classified in subgenus *Betula* (cluster IV, see Results) were examined in detail with a Principle COordinate analysis (PCO). The PCO was based on Dice distances and carried out in NTSYSpc. A further classification of these accessions was made in a Bayesian analysis using STRUCTURE 2.2 (Falush *et al.* 2003; Pritchard *et al.* 2000). The objective of this analysis was to test whether species form separate clusters or species groups, and whether hybrids are classified within or between these groups. STRUCTURE was developed for studies on populations, in which individual samples are assumed to be able to exchange genetic material. This is clearly not the case for the genus *Betula* as a whole, but may be realistic for the subgenus *Betula*, in which hybridization and introgression are common features. Ploidy levels in this subgenus vary from diploid to octoploid. As STRUCTURE does not contain models that can deal with this situation, all accessions were treated as diploid (explained below). Dominant AFLP marker data were entered by coding both alleles as '1' when a band was present and as '0' when a band was absent, while specifying '0' as a recessive allele for all markers. Estimates were obtained under the admixture model using the correlated allele frequencies option. Version 2.2 accommodates genotypic uncertainty in dominant marker data by sampling present bands as homozygotes or heterozygotes according to their posterior probabilities (Falush *et al.* 2007). This does not fully account for the fact that, given the dominant nature of AFLP markers, higher ploidy levels would contain a higher level of genotypic uncertainty. This may distort the absolute genetic distances, with polyploids ending up genetically more similar to one of the parents, but this will not interfere with the goal of identifying hybrid accessions. The number of inferred groups was evaluated at values of K ranging from 1 to 17, in which the maximum of $K = 17$ corresponds to the number of sampled species in subgenus *Betula* (excluding hybrids). Three replicate runs were performed for each value of K. A burn-in of 50,000 cycles and data collection for 100,000 cycles was used. The admixture model estimates the proportion of each accession's genome that descended from each of the K inferred groups.

Phylogenetic analyses were conducted on two data sets. The first set contained a maximum of four accessions per species because otherwise some species would be overrepresented in the data set. Accessions that were initially misclassified were preferentially excluded, followed by accessions that had the most missing values. Hybrid cultivars were also excluded. The second set was a subset of the first set, containing only the diploid species. Phylogenetic signal in the data set was quantified with the g_1 statistic (Hillis *et al.* 1992). Parsimony analyses were conducted in PAUP as heuristic searches with 100,000 random additions (holding one tree at each step), TBR branch swapping, multrees switched off, and ACCTRAN for character optimization. The initial search was followed by additional branch swapping on the Most Parsimonious Trees (MPTs) with the multrees option switched on. Branch support was assessed by bootstrap analysis comprising 10,000 replicates consisting of 10 random addition sequences with TBR branch swapping.

Results

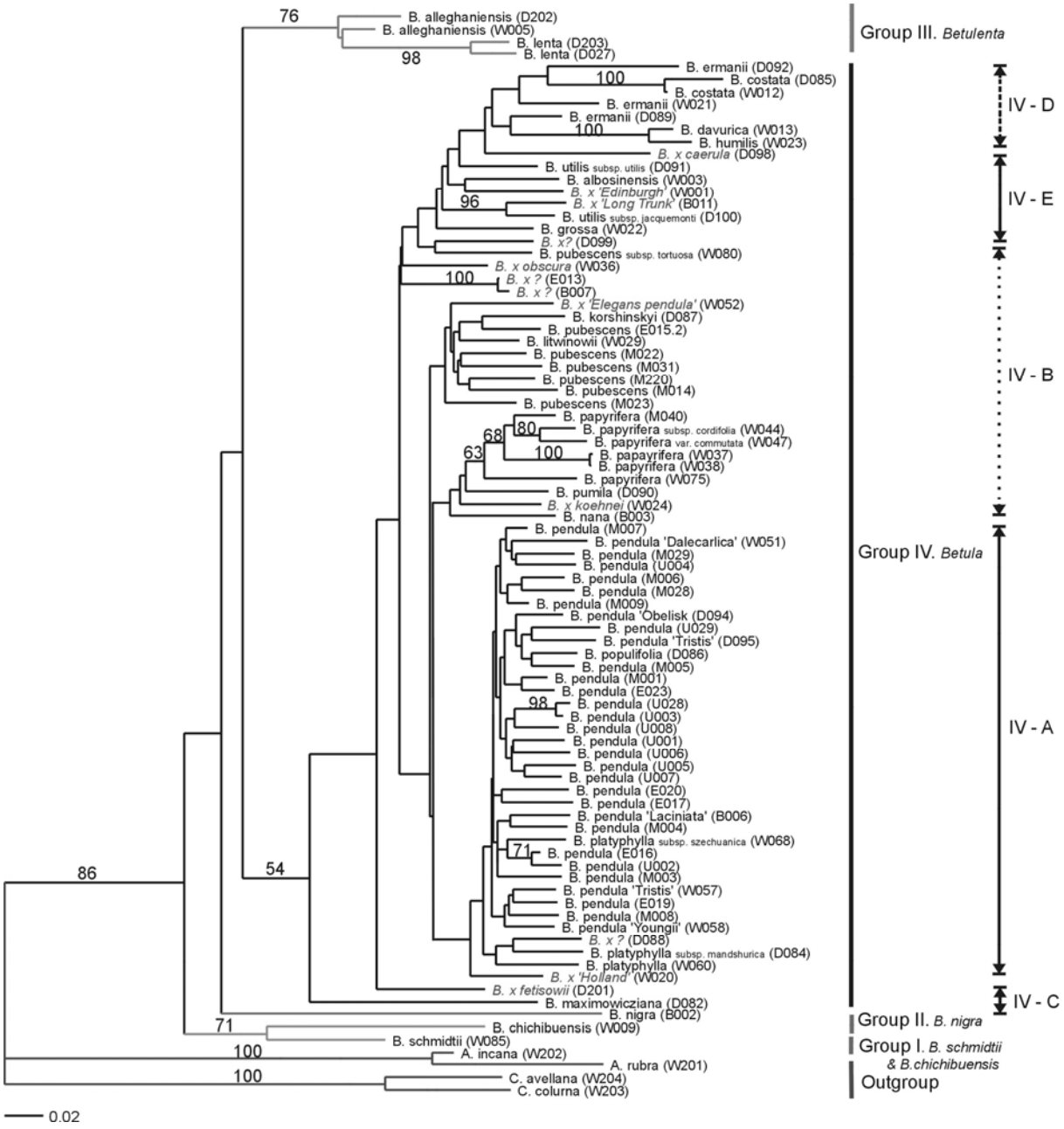
AFLP Similarities

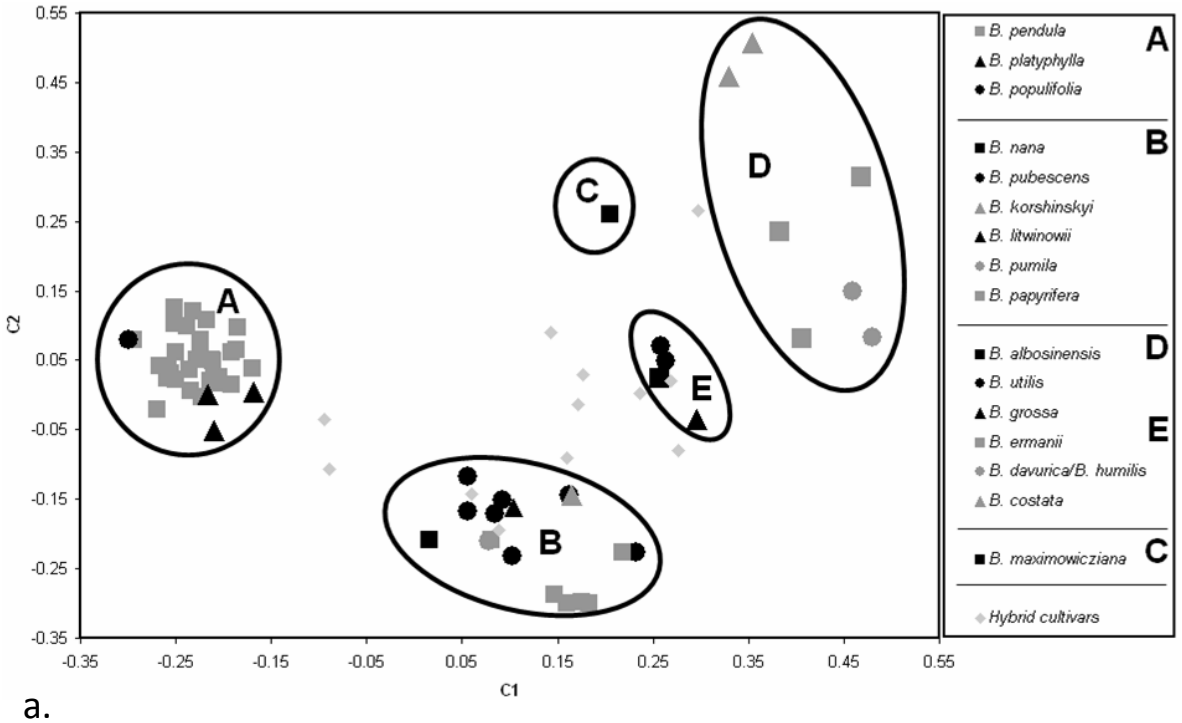
The three AFLP primer combinations produced 321 variable bands in 99 *Betula* and four outgroup accessions. Dendrograms obtained by UPGMA and NJ using Jaccard and Dice distances were highly similar (data not shown). Correlations between Dice and Jaccard similarity matrices and co-phenetic matrices from NJ trees and UPGMA dendrograms were high (0.96-0.98), with a one-tailed probability of 0.001 at 1,000 permutations. This indicates that the dendrograms provided a good fit to the similarity matrices. The highest correlation was found for Dice distances in combination with NJ (0.98). Similarity values between outgroup (*Corylus* and *Alnus*) and ingroup (*Betula*) ranged from 0.14 to 0.33, while ingroup similarity values all exceeded 0.32. The NJ tree shows a clear structure, although the support values for most branches were quite modest (Fig. 5.1). Four major clusters were present: (I) the *B. schmidtii/B. chichibuensis* cluster, (II) the *B. nigra* cluster, (III) the subgenus *Betulenta* cluster, and (IV) the subgenus *Betula* cluster. The latter contained all accessions from subgenus *Betula* and additional accessions from other subgenera (discussed below). Group IV was by far the largest group and contained several supported subgroups, such as the cluster with both *B. costata* accessions, the cluster with *B. davurica* and *B. humilis*, the cluster with all *B. papyrifera* accessions, and the cluster with *B. utilis subsp. jacquemontii* and *B.* x 'Long Trunk'.

Relationships within the subgenus *Betula*

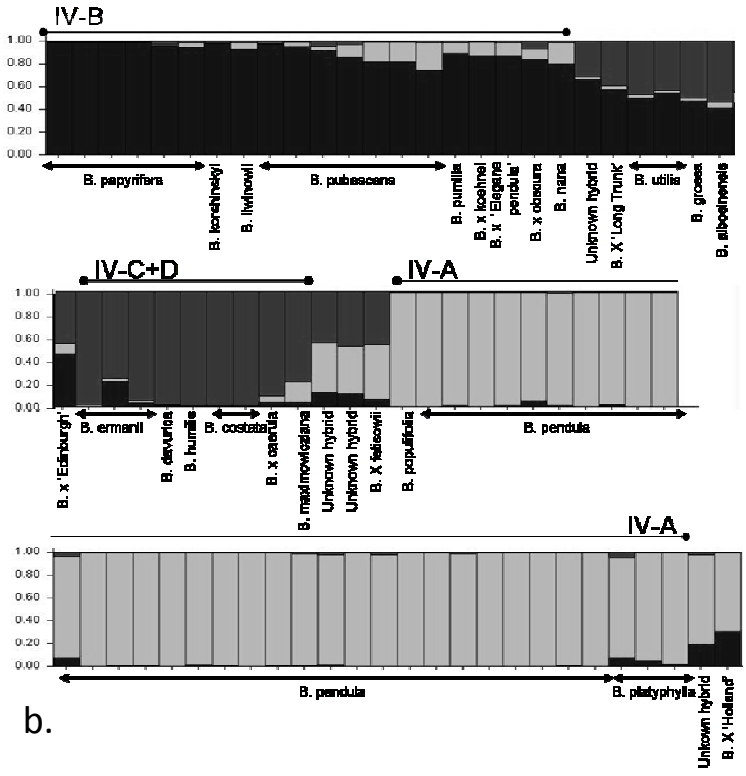
To allow a detailed analysis of the relationships within the *Betula* cluster (IV) we performed a PCO on the accessions within this group. The first three components had Eigenvalues of over 1.0 and explained 29.6% of the variation (16.8, 7.4 and 5.3, respectively). The first two components of the PCO are plotted in Fig. 5.2a, in which five groups are distinguished. Group A is represented by a large number of accessions comprising the species *B. pendula*, *B. plathyphylla* and *B. populifolia*. These species are not separated from each other on the first three components. Group B in the PCO plot comprised six species: *B. pubescens*, *B. litwinowii*, *B. korshinskyi, B. papyrifera, B. pumila* and *B. nana*. With the exception of *B. nana*, these species are hardly separated on the first two components. The third component did separate *B. papyrifera* and *B. pubescens,* and to a lesser extent *B. pumila* and *B. pubescens* (not shown). *B. maximowicziana* is placed in group C. Although the PCO puts *B. maximowicziana* close to group D, *B. maximowicziana* has the most basal position in cluster IV and a relatively low similarity to the other accessions in subgenus *Betula* (Fig. 5.1). The species *B. costata, B. davurica*, *B. humilis* and *B. ermanii* made up group C. These species branch off sequentially at more derived positions in cluster IV (Fig. 5.1) as do the species in group E, comprising *B. grossa*, *B. utilis* and *B. albosinensis.*

Figure 5.1 - Neighbour Joining tree of Dice similarities among 22 Betula *species and 10 interspecific hybrids (shown in italics) based on 321 AFLP markers. Bootstrap percentages ≥ 50 are shown on the branches. The groups within subgenus* Betula *as found in the PCO (Fig. 5.2) are shown on the right.*





a.



b.

Figure 5.2 - (a.) Principal coordinates plot of the Betula *accessions in subgenus* Betula *(group IV) (see Fig. 5.1) for the first two principal components estimated with 234 AFLP markers. (b.) STRUCTURE analysis of the* Betula *accessions in subgenus* Betula *(group IV) inferred from AFLP markers. In this figure, each accession is represented by a vertical bar partitioned into K=3 coloured segments. The K=3 analysis separated three groups (light gray, black and dark gray). The corresponding groups (IV-A, B and C+D) from Fig. 5.2a are displayed above the bars.*

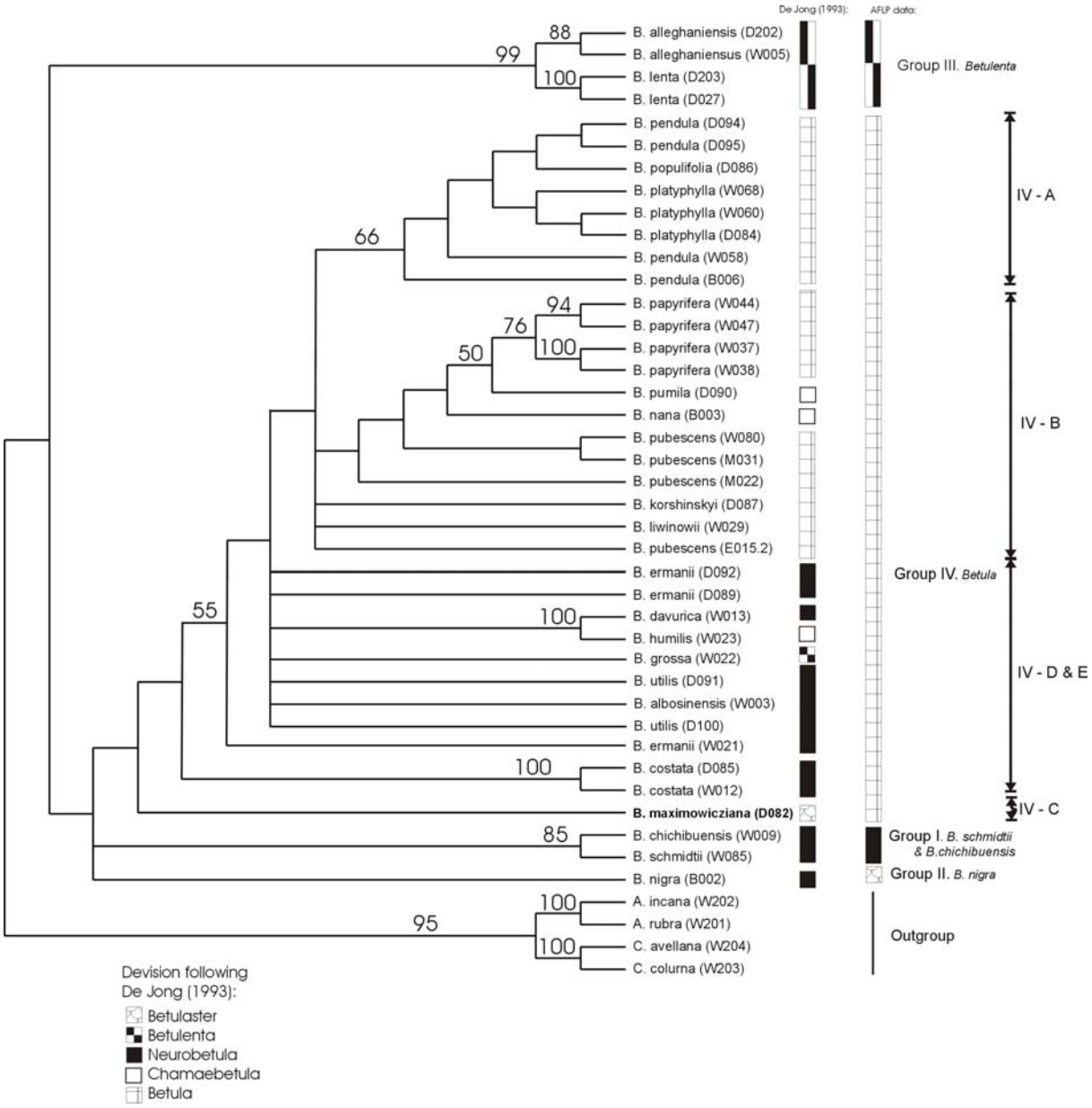
Most cultivars of hybrid origin are placed in-between the groups A to E. Their hybrid origin was confirmed by comparing the AFLP profile and/or ploidy level. Notable exceptions were the triploid hybrid 'Elegans Pendula', and *B.* x *koehnei* (a hybrid between *B. pendula* and *B. papyrifera)*, which was located among the *B. pubescens*-like accessions in group B. The cultivars 'Long Trunk' and 'Edinburgh' were positioned within group E, close to *B. utilis* and *B. albosinensis*. 'Long Trunk' was originally described as a *B. pendula* cultivar, but clusters with *B. utilis '*Doorenbos' in the NJ dendrogram and, based on the AFLP profile, appears to represent either a hybrid between *B. utilis* and *B. pendula* or a true *B. utilis* cultivar. The parental species of the hybrid cultivars could be established for some accessions. For example, the AFLP profile of *B.* x 'Elegans Pendula' shared a high number of bands with *B. pendula*. However, the close relatedness among the species in subgenus *Betula* and presence of species with a hybrid origin complicated this analysis.

To test the separation into groups and the presence of admixture in hybrid accessions within subgenus *Betula*, we used a Bayesian population clustering approach implemented in the program STRUCTURE (Falush *et al.* 2003; Pritchard *et al.* 2000). The structure analysis provided strong support for three species groups, with large and consistent improvements in the probability function [Ln *P*(D)] for runs with $K = 3$ relative to $K = 2$. Values of Ln *P*(D) improved modestly for $K = 4$ relative to $K = 3$, while $K > 4$ runs did not consistently produce further improvements. The three species groups that were distinguished were largely in line with the groups IV-A, IV-B and IV-C+D that were distinguished by the PCO (Fig. 5.2a-b). Some hybrids were actually clustered within the three distinguished groups. For example, *B.* x *koehnei*, *B.* x 'Elegans Pendula' and *B.* x *obscura* fell within group IV-B, while *B.* x *caerulea* fell within group IV-C+D. The other hybrids showed clear signs of admixture between groups. Notably, the species *B. albosinensis*, *B. utilis* and *B. grossa* that were attributed to group IV-E in the PCO plot were not distinguished as a separate group in the cluster analysis and showed admixture between group IV-B and IV-C+D, which is consistent with their position relative to these groups in the PCO plot (Fig. 5.2a).

Phylogenetic analysis of AFLP data

Cultivars with a mixed species background (interspecific hybrids, listed in Table 5.1) were excluded from the phylogenetic analysis, which was performed with 43 accessions. The dataset included 297 variable bands, 211 of which were parsimony informative. The g_1 statistic for the data set was -0.52. This value is considerably lower than the corresponding critical value of -0.09 (p=0.01; Hillis and Huelsenbeck, 1991) indicating the presence of ample phylogenetic signal. The initial parsimony analysis resulted in 12 MPTs of 721 steps on 11 different islands. The trees had a Consistency Index (Kluge *et al.* 1969) of 0.412 and a Retention Index (Farris 1989) of 0.585. Additional branch swapping did not yield any extra trees. The strict consensus of the MPTs is shown in Fig. 5.3.

Figure 5.3 - Strict consensus of 12 MPTs based on 297 AFLP markers and 22 Betula *species. Bootstrap percentages ≥ 50 are shown on the branches. The subgeneric division as proposed by De Jong (1993) is shown on the left vertical bar; the groups suggested by the AFLP data are shown on the right bar.*



The consensus tree from the data set without hybrids (Fig. 5.3), and the NJ tree of the data set including hybrids (Fig. 5.1) have a similar topology regarding the accessions that are present in both trees. Several groups can be identified when both trees are considered: (I) *B. schmidtii* and *B. chichibuensis* form a distinct and supported cluster in both the NJ (71% bootstrap support) and MP tree (85%). Both are Asian species from subgenus *Neurobetula.* These species are relatively

divergent from the other *Betula* accessions, with similarities between 0.33 and 0.48. (II) *B. nigra.* This North-American species is classified in subgenus *Neurobetula*, and is clearly separated from all other *Betula* species. It was the most divergent accession in the NJ tree with a similarity of 0.32 to 0.39 relative to the other *Betula* accessions. (III) *B. lenta* and *B. alleghaniensis* are two closely related North-American species that are classified in subgenus *Betulenta*. *B. lenta* is diploid, while *B. alleghaniensis* is hexaploid. The clade is supported in both the NJ (76%) and the MP tree (99%).

B. maximowicziana is an Asian species that is the only representative of subgenus *Betulaster* included in our study. It has a basal position in Group IV in both the NJ and MP tree and is placed in Group IV-C (Figs. 5.1 and 5.3). *B. costata* is the next species to branch of in the MPT, while being placed among the other representatives of group IV-D in the NJ tree. The remaining species of group IV-D and IV-E branch off sequentially within Group IV. The groups D and E do not form supported groups in the MPT. The species *B. ermanii*, *B. davurica*, *B. utilis* and *B. albosinensis* were previously classified in subgenus *Neurobetula. B. grossa* (subgenus *Betulenta)* is also placed within this group. In addition, *B. humilis* (subgenus *Chamaebetula*) is shown to be closely related to *B. davurica*. The clade with group IV-A and IV-B includes all examined species from subgenus *Betula*, namely *B. pendula*, *B. plathyphylla*, *B. populifolia*, *B. pubescens* and *B. papyrifera,* and two species from subgenus *Chamaebetula,* namely *B. nana* and *B. pumila*. *B. pumila* clusters with *B. papyrifera* in both the NJ tree and the MPT. Group IV-A from the PCO contains the *Betula* species *B. pendula*, *B. plathyphylla* and *B. populifolia,* and also forms a separate clade in the MPT.

Species with higher ploidy levels may represent natural interspecific hybrids. To exclude the effects of hybridization as much as possible, we repeated the analysis with only the diploid species. In the resulting NJ tree and MPTs (not shown), five groups could be identified. The first three groups (I to III) were similar to the groups discussed above, while the clustering within group IV was slightly different. In Fig. 5.3, *B. maximowicziana* is the most basal clade in group IV. In the diploid tree, *B. maximowicziana* and *B. costata* clustered together and had a basal position relative to *B. nana*, *B. pendula*, *B. plathyphylla*, and *B. populifolia*. The grouping of *B. maximowicziana* with *B. costata* was also suggested by the STRUCTURE analysis.

Discussion

AFLP markers for phylogeny

The taxonomy of the genus *Betula* is controversial. Although ample morphological variation exists in characters such as leaf shape, bark colour and shape of the catkins, attempts to reconstruct species relationships using morphological characters failed to produce a reliable classification. The occurrence of polyploidisation (Nagamitsu *et al.* 2006), hybridisation and introgression (Palme *et al.* 2004; Thórsson *et al.* 2001; Williams *et al.* 2001), and the fact that morphological characters may have evolved independently more than once in *Betula* (Li *et al.* 2005), may account for this.

Up to now, *Betula* taxonomy had been studied using morphological characters, flavonoid composition, and nuclear and chloroplast DNA sequences. The main limitations of morphological markers are selection pressure on morphological markers and hybrid morphology not always being intermediate (Thórsson *et al.* 2001). The main limitation of the sequence markers is their limited variation. We applied AFLP as an alternative, because AFLP generates polymorphic markers at a high frequency, has a high reproducibility, has genome wide sampling, and markers are not under direct selection pressure (Jones *et al.* 1997; McGregor *et al.* 2000; Russell *et al.* 1997).

More than 200 parsimony informative AFLP markers were generated in a data set of 87 unique *Betula* accessions. The AFLP data distinguished four subgenera and four groups within the largest subgenus (*Betula*). However, we could not unambiguously resolve relationships among these groups. Due to the extensive hybridisation and introgression within the genus *Betula*, good support for the relationships is not always to be expected, although bootstrap support was generally higher for the AFLP-derived groups compared to groups identified based on ITS sequences (Li *et al.* 2005; Nagamitsu *et al.* 2006), and AFLP was able to provide resolution on clades that were unresolved with ITS data. Apart from differences in support and resolution, congruence between ITS data and AFLP data was high for the genus *Betula*. This general congruence between AFLP and ITS tree topologies is found across a wide range of taxonomic groups (Koopman 2005). Besides yielding detailed information on *Betula* relationships, AFLPs provided complementary information on hybridization events. Such events were reflected in the AFLP profiles by bands shared between the hybrid and parental species, and for some hybrids in the STRUCTURE analysis. In ITS sequences the parental information may be lost or misleading (Álvarez *et al.* 2003), even when fragments are cloned before sequencing (Nagamitsu *et al.* 2006).

Phylogeny of the genus *Betula*

In summary, the AFLP results indicate the presence of four subgenera in *Betula*, as opposed to the five subgenera originally recognised by De Jong (1993). Each of the original subgenera is discussed separately below.

Subgenus *Betulenta* (De Jong 1993), which is synonymous with the *Lentae* of Regel (1865), was represented by *B. lenta*, *B. alleghaniensis*, *B. medwediewii* and *B. grossa* in our study. Winkler (1904) merged this group with subsection *Costatae*, but our results support the division of De Jong (1993), and show that the *Betulenta* sensu Winkler (1904) would be paraphyletic. *B. lenta* and *B. alleghaniensis* grouped together with good support in both the AFLP NJ and MP analysis. The close relationship of *B. lenta* and *B. alleghaniensis* (together with *B. medwediewii*) is supported by ITS and *matK* sequence data (Järvinen *et al.* 2004; Li *et al.* 2005). The decaploid *B. medwediewii* was not included in our analysis due to an excess of bands in the AFLP profile. Our AFLP results are not in line with data on ITS sequences that indicated a close relationship of *B. alleghaniensis* with the

non-*Betulenta* species *B. costata* (Li *et al.* 2005). Morphologically, *B. costata* and *B. alleghaniensis* are very distinct, and we therefore consider the AFLP results to be more reliable. The position of *B. grossa* in the *Betulenta* is supported by data on phenolic variation (Keinänen *et al.* 1999). However, both our AFLP data and data on ITS sequences (Nagamitsu *et al.* 2006) suggest that *B. grossa* is not positioned within *Betulenta*. The position of this species is thus uncertain. *B. lenta*, *B. alleghaniensis* and *B. medwediewii* are maintained within *Betulenta*.

Subgenus *Betulaster* was represented by a single accession of *B. maximowicziana*. According to the AFLP data, this accession is positioned at the periphery of subgenus *Betula* in group C. Li *et al.* (2005) pointed out that *B. maximowicziana* did not cluster with other representatives of subgenus *Betulaster*. Given that *B. alnoides* is the type species of subgenus *Betulaster*, the status of *Betulaster* would depend on the position of *B. alnoides*, which was not included in our study. *B. nigra* did cluster with *B. alnoides* according to Li *et al.* (2005). In our AFLP trees, *B. nigra* formed a separate-and well-supported group. The above therefore suggests that subgenus *Betulaster* may, in fact, be a distinct group whose position in our AFLP trees is represented by *B. nigra*.

In our data set, subgenus *Neurobetula* was represented by seven species (excluding *B. nigra*) which separated in two major groups. Group I (*B. chichibuensis* and *B. schmidtii*) was well-defined and well-supported, while the other group (Group IV-D+E in Figs. 5.1-3) was more loosely defined and contained the species *B. costata*, *B. ermanii, B. davurica*, *B. utilis* and *B. albosinensis*. Previous studies concluded that subgenus *Neurobetula* is a heterogeneous and polyphyletic group (De Jong 1993; Li *et al.* 2005). However, an alternative division was not proposed, because conflicts in morphological markers and low variation among DNA markers hampered an unambiguous conclusion. The position of *B. schmidtii* as a close relative to *B. chichibuensis* was in line with studies on ITS sequences (Li *et al.* 2005; Nagamitsu *et al.* 2006) and a recent morphological study by Skvortsov (2002). On the other hand, Keinänen *et al.* (1999) and Järvinen *et al.* (2004) found that *B. schmidtii* was closely related to subgenus *Betula*. Our results indicate that *B. costata*, *B. ermanii, B. davurica*, *B. utilis* and *B. albosinensis* should be merged with subgenus *Betula*. The common existence of hybrids between the above species and species from subgenus *Betula* also support placement in a single subgenus. All these species are of Asian origin and a more extensive sampling from their natural range will be needed to resolve interspecific relationships.

Subgenus *Chamaebetula* (De Jong, 1993) was represented in our study by *B. humilis, B. nana* and *B. pumila*. According to De Jong (1993) subgenus *Chamaebetula* is polyphyletic, and artificially grouped based on the single morphological character of having a shrubby habitus. Our results indicate that these species should be placed within subgenus *Betula*. For *B. nana*, a close relationship to subgenus *Betula* is supported by sharing of chloroplast haplotypes between *B. nana* and *B. pendula* (Maliouchenko *et al.* 2007; Palme *et al.* 2004). *B. nana*/*B. pumila* and *B.*

humilis are placed in different groups (IV-B *vs.* IV-D) within subgenus *Betula*, confirming the polyphyletic nature of the *Chamaebetula*. The above suggests that subgenus *Chamaebetula* is superfluous.

Subgenus *Betula* was represented in our study by *B. pendula*, *B. plathyphylla, B. populifolia, B. pubescens*, *B. litwinowii, B. korshinsky*i and *B. papyrifera*. Although not supported by high bootstrap values, the species originally placed within subgenus *Betula* do consistently group together in both NJ tree (Fig. 5.1) and MPTs (Fig. 5.3). The PCO analysis distinguished four species groups in the subgenus *Betula* (Fig. 5.2a), while the STRUCTURE analysis recognized three groups. The species *B. pendula*, *B. plathyphylla* and *B. populifolia* clustered together in group A, and showed hardly any genetic differentiation. In fact, the AFLP data failed to differentiate between these species, suggesting that they are conspecific. Their morphology is also very similar and Skvortsov (2002) already considered *B. platyphylla* to be synonymous with *B. pendula*. Only one *B. populifolia* accession was included, so further sampling within the natural range of this species will be necessary to confirm its status. Group B consisted of the potentially conspecific *B. pubescens*, *B. litwinowii* and *B. korshinskyi, B. papyrifera* and two species from the *Chamaebetula*, namely *B. nana* and *B. pumilla*. *B. pubescens*, *B. litwinowii* and *B. korshinskyi* do not separate in the PCO, but more extensive sampling is required to establish their status. Group D comprised *B. maximowicziana* as discussed above, while groups D and E comprised several species that were originally attributed to subgenus *Neurobetula*. The groups C and D formed a single group in the STRUCTURE analysis. Group E was shown to have an intermediate position between the groups B and D. This group contained only polyploid species, consistent with a potential hybrid origin. In summary, all previously assigned species were retained in subgenus *Betula*, while subgenus *Chamaebetula*, part of the species from *Neurobetula*, and *B. maximowicziana* were also placed in this subgenus.

Evolution

Although our results indicate that four major taxonomic groups can be recognised within the genus *Betula*, the relationship between them remains unclear. The most obvious explanation for the lack of support is the occurrence of hybridization and introgression, which would have a homogenizing effect on the relationships between species. Several types of hybrids may occur, and they can be classified as “newly formed (F1) hybrids”, “later generation hybrids” and “hybrid species” (Vriesendorp *et al.* 2005). Hybrid cultivars are likely to fall within the first two groups and we excluded the cultivated hybrids from our MP analysis. However, naturally occurring hybrid species may also exist, and may in fact make up a significant proportion of all *Betula* species. Their presence in the data set may explain the low bootstrap values within the *Betula* clade. Species such as *B. ermanii*, *B. humilis*, *B. utilis* and *B. pubescens* each have ploidy levels higher than 2n and to some extent contain AFLP bands that can be regarded as diagnostic bands for *B. pendula*.

However, we can not determine to which extent this reflects shared evolution or shared parental species in a natural hybridisation process. An alternative explanation for the lack of support relates to a situation in which the major speciation events took place within a very short time frame. This would result in so called bush-like clades that are characterised by short stems relative to the length of the branches (Rokas *et al.* 2006). Under these circumstances, homoplasy may limit the phylogenetic resolution by overwhelming the true phylogenetic signal. This could also explain why we can not determine the relationships between subgenera. If this is the case, these relationships may remain unresolved.

Screening for allergenicity

Birch pollen is a major cause of hay fever. The species *B. pendula* is known to cause allergic complaints throughout Europe. Several protein variants of the major birch allergen (Bet v 1) are found in this species (Breiteneder *et al.* 1989; Schenk *et al.* 2006), and these are known to differ in allergenicity (Ferreira *et al.* 1997). Phylogenetic relationships may be a useful predictor of the allergenicity of particular birch species, because a high similarity between protein variants increases the chance that they share particular epitopes. For example, close relatives of *B. pendula*, such as *B. populifolia* and *B. plathyphylla,* are probably allergenic as well because they are likely to contain Bet v 1 protein variants similar to those of *B. pendula*. All hybrids were also positioned in subgenus *Betula*. Several were tightly linked to *B. pendula*, which may be one of the parental species of these hybrids. These hybrids will most likely share the genetic background to be allergenic. The species from the other subgenera (I-III), such as *B. lenta*, *B. nigra* or *B. schmidtii,* represent good candidates for germplasm screening for non-allergenic protein variants of Bet v 1.

CHAPTER 6

Characterization of PR-10 genes from eight Betula species and detection of Bet v 1 isoforms in birch pollen

Martijn Schenk, Jan Cordewener, Twan America,
Wendy van 't Westende, René Smulders and Luud Gilissen

To be submitted

Abstract

Bet v 1 is an important cause of hay fever in northern Europe. Bet v 1 isoforms from the European white birch *(Betula pendula)* has been investigated extensively. The allergenic potency of other birch species is unknown. Cloning and sequencing of PR-10 genes was performed on eight birch species to establish the presence of these genes. Q-TOF LC-MS^E was applied to identify which PR-10/Bet v 1 genes are expressed in pollen and to determine the relative abundance of individual isoforms in the pollen proteome. PR-10 genes were found in all examined birch species. In total, 134 unique sequences were recovered. Sequences were attributed to different genes or pseudogenes that were subdivided into seven subfamilies. Five subfamilies were common to all birch species. Protein analysis of pollen from five birch species revealed that the genes of two subfamilies were expressed in pollen, while each species expressed a mixture of isoforms with least 4-5 different isoforms. Isoforms that were similar to isoforms with a high IgE-reactivity (Bet v 1a =PR-10.01A01) were abundant in all species except *B. lenta*, while the hypoallergenic isoform Bet v 1d (=PR-10.01B01) was restricted to *B. pendula* and close relatives of *B. pendula*. Q-TOF LC-MS^E allows fast screening of Bet v 1 isoforms in pollen by determining their presence and relative abundance. *B. pendula* contains a Bet v 1 mixture in which both isoforms with a high and low IgE-reactivity are abundant. The presence of isoforms with high IgE-reactivity is apparently of determining influence for the allergenicity of this species. The birch species which express variants that are similar to Bet v 1a are predicted to be allergenic as well.

Introduction

Birch trees grow in the temperate climate zone of the northern hemisphere and release large amounts of pollen during spring. This pollen is a major cause of Type I allergies. The main birch allergen in northern Europe is a pathogenesis-related class 10 (PR-10) protein from the European white birch *(Betula pendula)* termed Bet v 1 (Breiteneder *et al.* 1989; Jarolim *et al.* 1989). Pollen of other Fagales species contains PR-10 homologues that share epitopes with Bet v 1 (Niederberger *et al.* 1998), as do several fruits, nuts and vegetables (Bollen *et al.* in press; Fritsch *et al.* 1998; Karlsson *et al.* 2004; Scheurer *et al.* 1999). An IgE-mediated cross-reaction to these food homologues causes the so called oral allergy syndrome (Bohle *et al.* 2003; Ferreira *et al.* 2004). PR-10 proteins constitute the largest group of aeroallergens and are among the four most common food allergens (Breiteneder *et al.* 2000).

The genus *Betula* encompasses over 30 tree and shrub species that are found in diverse habitats in the boreal and temperate climate zone of the Northern Hemisphere. The taxonomy of the *Betula* genus is debated as is the number of recognized species. The genus is divided into three, four or five groups or subgenera (De Jong 1993; Järvinen *et al.* 2004; Schenk *et al.* submitted). *B. pendula* occurs in Europe and is the only species whose relation to birch pollen allergy has been extensively investigated. Sensitization to birch allergy is also reported across Asia and North America, where *B. pendula* is not present (Abe *et al.* 1997; Eriksson *et al.* 1998). Other *Betula* species occur in these areas, but their allergenic potency is unknown. *Betula* species may vary in their allergenicity as variation in allergenicity has been found for different cultivars from apple (Bolhaar *et al.* 2005a; Marzban *et al.* 2005), peach and nectarine (Ahrazem *et al.* 2007), and olive trees (Castro *et al.* 2003).

PR-10 proteins are reported as a multigene family in many higher plants ranging from Gymnosperms to Monocots and Dicots (Ekramoddoullah *et al.* 2000; Gao *et al.* 2005a; Huang *et al.* 1997). The classification as PR-proteins (Van Loon and Van Strien, 1999) is based on the induced expression in response to pathogen infections by viruses, bacteria or fungi (Pühringer *et al.* 2000; Robert *et al.* 2001; Swoboda *et al.* 1995c), to wounding (Poupard *et al.* 1998) or to abiotic stress (Moons *et al.* 1997; Utriainen *et al.* 1998). Some members of the PR-10 gene family are constitutively expressed during plant development (Walter *et al.* 1996) or expressed in specific tissues (Huang *et al.* 1997). Multiple PR-10 genes have been reported for *B. pendula* as well (Schenk *et al.* 2006). mRNAs of these genes have been detected in various tissues, including birch pollen (Breiteneder *et al.* 1989; Friedl-Hajek *et al.* 1999; Swoboda *et al.* 1995b), roots and leaves (Poupard *et al.* 1998; Utriainen *et al.* 1998), and in cells that were grown in a liquid medium in the presence of microbial pathogens (Swoboda *et al.* 1995c). PR-10 genes share a high sequence similarity and form a homogeneous group. Homogeneity is believed to be maintained by

concerted evolution (Wen *et al.* 1997). Arrangements of *PR-10* genes into clusters, such as found for Mal d 1 genes in apple, may facilitate concerted evolution (Gao *et al.* 2005a).

Several Bet v 1 isoforms are described for *B. pendula* (Breiteneder *et al.* 1989; Friedl-Hajek *et al.* 1999; Hoffmann-Sommergruber *et al.* 1997; Schenk *et al.* 2006; Swoboda *et al.* 1995b), including both allergenic and hypoallergenic isoforms (Ferreira *et al.* 1996). Individual *B. pendula* trees have the genetic background to produce a mixture of Bet v 1 isoforms with varying IgE-reactivity (Schenk *et al.* 2006). The abundance of particular isoforms at the protein level will influence the allergenicity of the pollen. Molecular masses and sequences of tryptic peptides from Bet v 1 can be determined by Q-TOF MS/MS (Helsper *et al.* 2002). The recently developed Q-TOF LC-MS^E method enables peptide identification, but has the additional advantage of determining relative abundances of peptides (Silva *et al.* 2005). The quantification of isoforms with a known IgE-reactivity (Ferreira *et al.* 1996) can be used to predict the allergenicity of particular birch species and potentially to identify species with a reduced allergenic.

Knowledge on the diversity and expression of PR-10 genes is required to characterize the allergenicity or birch species, and to facilitate selection and breeding of hypoallergenic birch trees. To investigate the presence of Bet v 1-like genes and allergenic potency in *Betula* species that are potential crossing material, we: (*I*) cloned and sequenced *PR-10* alleles from eight species from four groups/subgenera in the *Betula* genus to detect the *PR-10* genes at the genomic level, (*II*) applied Q-TOF LC-MS^E to identify the pollen-expressed Bet v 1 genes, and (*III*) determined the relative abundance of isoforms in the pollen proteome and compare their sequence with that of isoforms with a known IgE-reactivity.

Material and Methods

Plant Material

We collected young leaves from eight *Betula* species (Table 5.1). The four groups that were identified in a recent phylogenetic analysis were represented by at least one species (Schenk *et al.* submitted). Four species from the largest group (subgenus *Betula*) were included to cover the variation within this group. Plant material was collected from the botanical collections of PPO Boskoop (Boskoop, the Netherlands), the Botanical Garden of Wageningen (Wageningen, the Netherlands), and the Von Gimborn Arboretum (Doorn, the Netherlands). Fresh leaf samples were analyzed by flow cytometry (Plant Cytometry Services, Schijndel, The Netherlands) to estimate the ploidy level. Diploid *(B. pendula)* and tetraploid *(B. pubescens)* controls were included. All examined accessions were diploid, keeping the number of expected sequences per accession low. We collected pollen from the same trees that were used to extract DNA for the species *B. nigra*, *B. chichibuensis*, *B. lenta*, *B. costata* and *B. pendula* (representing all four subgenera). Pollen was collected during the flowering period of birch in April-May 2004.

PCR, cloning and sequencing

DNA was extracted using the DNeasy Plant Mini kit (Qiagen) according to the manufacturer's instructions. *PR-10* alleles were amplified from genomic DNA with two primer pairs that had been tested and used in previous research on *B. pendula* (Schenk *et al.* 2006). PCR amplification with both primer pairs was performed in 20 µl reactions containing 0.1 mM dNTP, PCR Reaction buffer (Eurogentec), 1.5 mM $MgCl_2$, 0.6 µM forward primer, 0.6 µM reverse primer, 0.5U *Taq* polymerase (Goldstar), and 20-80 ng template DNA. PCR reactions started with a heating step at 95˚C for 15 minutes, followed by 16-24 cycles of denaturation at 94˚C for 30 s, annealing at 50˚C for 45 s, and extension at 72˚C for 2 min. A final extension step of 10 min at 72˚C was added after the last cycle. To reduce the number of PCR recombination artifacts, we used as few PCR cycles as possible. The minimum number of cycles required to generate sufficient product for cloning was assessed by visual inspection of the amplified products on agarose gel.

PCR products were purified with the MinElute PCR Purification Kit (Qiagen). Purified samples were ligated into the pGEM-T easy Vector (Promega) and established in *Escherichia coli* Subcloning Efficiency DH5α cells (Invitrogen) according to the manufacturer's instructions. White colonies were picked from agar plates and grown overnight at 37˚C in freeze medium. PCR-based screening was performed with vector-specific M13 primers. PCR products were purified with Sephadex G-50 (Millipore). The DYEnamic™ ET Terminator Cycle Sequencing Kit (Amersham) was used for the sequence reaction. Sequence products were analyzed on a 96-capillary system (ABI 3730xl). Genomic *Betula* sequences have been submitted to GenBank.

Phenetic/phylogenetic analysis

Potential PCR artifacts (strand switching and base misincorporation) were excluded by retaining only those sequences that were confirmed in independent PCRs. We included one reference sequence per *B. pendula* gene to the dataset for comparison with previous results in *B. pendula* (Schenk *et al.* 2006). Nucleotide sequences were aligned using CLUSTAL W with a gap penalty of 10 and a gap extension penalty of 2. Primer traces and introns were excluded from further analysis. A Neighbor Joining (NJ) tree was constructed with Kimura two-parameter distances. Gaps were treated as missing characters. Bootstrapping was carried out with 1,000 replicates in PAUP 4.0b10 (Swofford 2003). The outgroup was composed of *PR-10* sequences from *Castanea sativa* (AJ417550) and *Fagus sylvatica* (AJ130889), which are two related Fagales species. Parsimony analysis was conducted in PAUP as a heuristic search with the following options: 100,000 random additions while holding one tree at each step, TBR branch swapping, the MulTrees option switched on, and ACCTRAN for character optimization. A strict consensus tree was calculated for all of the most parsimonious trees. Branch support was assessed by bootstrap analysis comprising 10,000 replicates consisting of 10 random addition sequences with TBR branch swapping. The results were highly similar for both analyses; therefore, only the results from the NJ analysis are shown.

Protein search database

Nucleotide sequences were aligned codon-by-codon. We analyzed general selection patterns at the molecular level using DnaSp 4.00 (Rozas *et al.* 2003). The number of non-synonymous (K_a) and synonymous substitutions (K_s) per site were calculated from pair wise comparisons with incorporation of the Jukes-Cantor correction. Nucleotide data were translated into amino acids, and the resulting protein sequences were placed in a Fasta database to be used as a search database in the Q-TOF LC-MS^E analysis. As sequence information for the primer region was not available, we used the GenBank sequences X15877 (subfamily 01), X77265 (02), X77600 (03), and X77601 (05) to fill these gaps in sequences from the respective subfamilies. The initiating Methionine is removed during PR-10 protein synthesis (Ferreira *et al.* 1996; Utriainen *et al.* 1998) and was therefore removed from the predicted proteins. Protein sequences of birch PR-10 isoforms in GenBank (overview in: Schenk et al., 2006), keratin, trypsin and Bet v 7 (GenBank AJ311666) were added to the database.

Protein extraction

Fifty mg of pollen were suspended in 1 ml of 0.05 M Tris-HCl (pH 7.5) following Cadot *et al.* (1995), who found that yield and diversity of the extracted allergens are optimal at pH 7.5 for birch pollen. After incubation for 1 hr at room temperature under constant shaking, the pollen extract was centrifuged at 10.000 rpm for 5 min and the pellet was ground with an Eppendorf-fitting pestle. The extract was shaken for another 1 hr. After centrifugation (10.000 rpm; 5 min), the supernatant was collected and freeze dried for storage.

SDS-PAGE

The freeze dried protein extract was redissolved in 0.05 M Tris-HCL (pH 7.5) and analyzed with SDS-PAGE for the presence of Bet v 1-type proteins. Proteins were separated on a 15% w/v acrylamide SDS-PAGE gel with a 5% w/v stacking gel using the Mini-Protean II gel system (Bio-Rad). After staining with Coomassie BB R-250, the gels were scanned and analyzed by Quantity One (Bio-Rad) scanner software. SDS-PAGE Standards broad range (Bio-Rad) markers were used to determine relative molecular masses.

The protein bands at a relative molecular mass of 16-18 kDa were cut out of the SDS-PAGE gel and processed essentially according to Shevchenko (1996). Bands were sliced into 1 mm^3-pieces. Bands at 14, 19 and 35 kDa were cut out and analyzed as well. Proteins were reduced with DTT and alkylated with iodoacetamide. Gel pieces were dried under vacuum, and swollen in 0.1 M $NaHCO_3$ containing sequence-grade porcine trypsin (10 ng/µl, Promega). After digestion at 37 °C overnight, peptides were extracted from the gel with 50% v/v acetonitrile, 5% v/v formic acid and dried under vacuum.

Q-TOF LC-MS/MS and Q-TOF LC-MS^E

Tryptic digests were analyzed by one-dimensional LC-MS in high-throughput configuration using the Ettan^{TM} MDLC system (GE Healthcare) directly connected to a Q-TOF-2 Mass Spectrometer (Waters Corporation, UK). Samples (5 μl) were loaded on 5 mm x 300 μm ID Zorbax^{TM} 300 SB C18 trap columns (Agilent Technologies), and the peptides were separated on 100 μm i.d. x 15cm Chromolith CapRod monolithic C18 capillary columns (Merck) at a flow rate of approximately 1 μl/min. Solvent A contained an aqueous 0.1% formic acid solution and solvent B contained 84% acetonitrile in 0.1% formic acid. The applied gradient consisted of isocratic conditions at 5% B for 10 min, a linear gradient to 30% B over 40 min, a linear gradient to 100% B over 10 min, and then a linear gradient back to 5% B over 5 min. MS analyses were performed in positive mode using ESI with a NanoLockSpray source. As lock mass, [Glu^1]fibrinopeptide B (1 pmol/μl) (Sigma) was delivered from the syringe pump (Harvard Apparatus, USA) to the reference sprayer of the NanoLockSpray source at a flow rate of 1 μl/min. The lock mass channel was sampled every 10 s.

For the initial identification of the 14, 16-18, 19 and 35 kDa bands, the Q-TOF-2 was operating in MS/MS mode for data dependent acquisition. The mass spectrometer was programmed to determine charge states of the eluting peptides, and to switch from the MS to the MS/MS mode for $z \geq 2$ at the appropriate collision energy for Argon gas-mediated CID. Each resulting MS/MS spectrum contained sequence information on a single peptide. Processing and database searching of the MS/MS data set was performed using ProteinLynx Global SERVER (PLGS) v2.3 (Waters Corporation) and the NCBI non-redundant protein database, while taking fixed (carbamidomethylation) and variable (oxidation of Methionine) modifications into account.

After identification of multiple Bet v 1 isoforms in the 16-18kDa band, we analyzed the tryptic digest from this band with Q-TOF LC-MS^E. The Q-TOF-2 was programmed to alternate between low and elevated levels of collision energy. Collision energy was 5eV in MS mode and and increased in two steps from 28 to 40eV in MS^E mode. Measuring time in both modes was 0.9 s with an interscan delay of 0.1s. Unfragmented precursors predominate in low energy mode, while fragmented ions of the precursors are observed in high energy mode. Digests were analyzed in duplicate. MS^E data were analyzed according to the procedure described by Silva *et al.* (2005) with the Expression module in PLGS. Different peptide components were detected with an ion detection algorithm, and then clustered by mass and retention time, followed by normalization of the data. The described PR-10 protein search database was used to identify peptides, while taking fixed (carbamidomethylation) and variable (oxidation of Methionine) modifications into account. After processing by PLGS, the so-called Exact Mass and Retention Time (EMRT) table was exported and reclustered using the in-house developed PACP tool (De Groot *et al.* 2008) in order to correct potential misalignment and split peak detection errors. The retention time was normalized and the reclustered EMRT table was further analyzed in Excel.

Results

PR-10 subfamilies

We examined eight *Betula* species for the presence of PR-10 genes by sequencing 1029 individual clones in both directions (Table 6.1). Sequences that contained PCR artifacts were excluded by combining information from independent PCRs. The Open Reading Frame (ORF) of the sequences was highly conserved and the alignment was straightforward. The consensus sequence of the exon had 452 positions, excluding the 31 bp that were located in the primer region. 274 consensus positions were variable, of which 228 positions were phylogenetically informative. The sequences were grouped in seven well-supported clusters in the Neighbor Joining (NJ) tree (Fig. 6.1). Five clusters coincided with the division between subfamilies as found for *B. pendula* (Schenk et al. 2006). Two new subfamilies were identified. Contrary to the previously described subfamilies 01 to 05, genes of subfamilies 06 and 07 were not identified in all species. They were only found in *B. schmidtii* and *B. chichibuensis*, and in *B. schmidtii* and *B. lenta,* respectively (Table 6.1). All sequences contained an intron that was positioned between the first and second nucleotide of codon 62. Introns were highly variable in length and composition, which was an additional characteristic for inferring the proper subfamily. The intron was excluded from the phenetic/phylogenetic analysis because it evolves at a different speed than the ORF.

PR-10 sequences and genes

Between 12 and 25 unique PR-10 sequences were recovered per species, adding up to 146 sequences in total (Table 6.1). We found 134 unique sequences, and over 100 sequences are described here for the first time. *B. pendula*, *B. plathyphylla* and *B. populifolia,* which are all members of the subgenus *Betula,* had multiple sequences in common. Another member of this subgenus, *B. costata*, had one allele in common with both *B. pendula* and *B. populifolia*. We attributed the sequences to different genes according to a predefined cut-off level of 98.5%, while allowing maximally two alleles per gene per species. These criteria go together in the majority of cases, but the *B. chichibuensis* sequences in the large cluster in subfamily 03, the *B. lenta* sequences in subfamily 02, and the genes 02A/02B and 03C/03D in *B. pendula* are more than 98.5% similar, while belonging to different genes. The total number of identified genes per species is shown in Table 6.1. The previously identified *PR-10* genes from *B. pendula* are indicated in Table 6.1 and Fig. 6.1 as *B. pendula* reference. Alleles from 11 of the 13 reference genes were recovered from the newly sequenced *B. pendula* cultivar. No new genes were identified. The majority of genes are, thus, recovered by sequencing over 100 clones per species, but some genes might still be missing from the dataset.

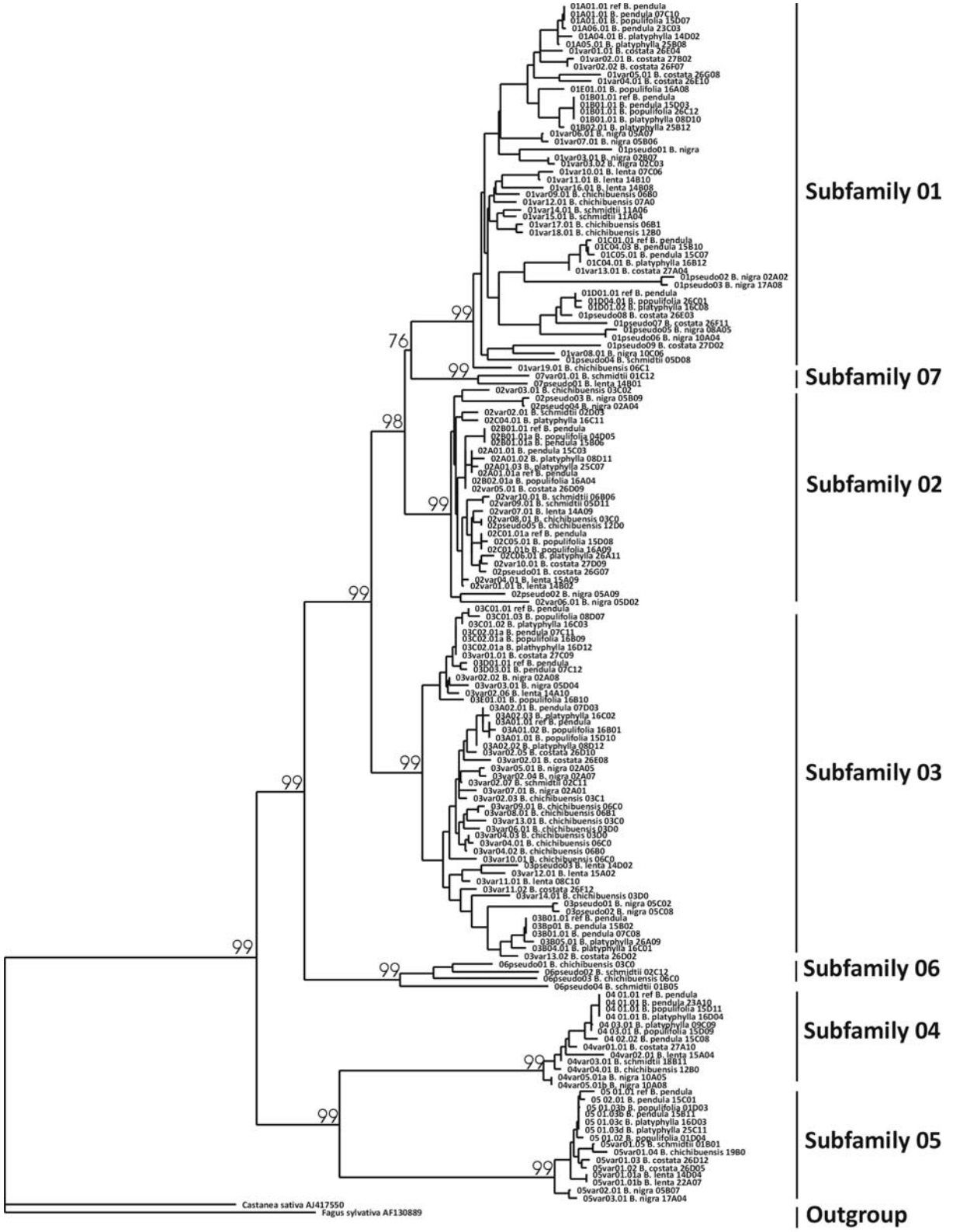
TABLE 6.1 – The examined Betula *species, the number of clones sequenced in both directions, and the number of identified sequences and genes per species.*[1] *The distinction of subfamily 01 to 05 follows Schenk* et al. *(2006), while subfamily 06 and 07 are new. Subfamily 01 and 02 were predicted to be expressed in pollen based on homology to mRNA sequences.*

Species	Number of sequenced clones	Subfamily 01		Subfamily 02		Subfamily 03		Subfamily 04		Subfamily 05		Subfamily 06		Subfamily 07		Total	
		Seqs.	Genes	Seqs	Genes	Seqs	Genes	Seqs	Genes	Seqs	Genes	Seqs	Genes	Seqs	Genes	Seqs	Genes
Subgenus *Betulaster:*																	
B. nigra	155	10	6	4	3	7	4	2	1	2	1	-	-	-	-	25	15
Subgenus *Neurobetula:*																	
B. chichibuensis	170	5	4	3	2	10	7	1	1	1	1	2	2	-	-	22	17
B. schmidtii	184	3	2	3	2	1	1	1	1	1	1	2	2	1	1	12	10
Subgenus *Betulenta:*																	
B. lenta	106	3	2	3	2	4	4	1	1	2	1	-	-	1	1	14	11
Subgenus *Betula:*																	
B. costata	103	9	8	3	2	5	5	1	1	2	1	-	-	-	-	20	17
B. pendula	102	5	3	2	2	5	4	2	1	2	1	-	-	-	-	16	11
B. plathyphylla	103	6	4	4	3	6	3	2	1	2	1	-	-	-	-	20	12
B. populifolia	106	4	4	4	2	5	3	2	1	2	1	-	-	-	-	17	11
B. pendula reference[2]	-	-	4	-	3	-	4	-	1	-	1	-	-	-	-	-	13

1 Species were diploid (2n) as measured by flow cytometry. The identification of alleles of a single gene is based on the criterion of having >98.5% sequence similarity, and by allowing maximally two alleles per gene.

2 Genes identified in *B. pendula* (Schenk *et al.* 2006).

Figure 6.1 – Clustering of the PR-10 sequences from eight Betula *species in a Neighbor Joining tree with Kimura two-parameter distances. The sequences are grouped into seven subfamilies. Support for these groups is indicated by displaying the Bootstraps percentages on the branches.*



Homologues of the PR-10 genes in *B. pendula* were identified in *B. populifolia* and *B. plathyphylla*. The labeling of sequences from these species was made according to the procedure described by Gao *et al.* (Gao *et al.* 2005a), because it had previously been used for *B. pendula* (Schenk *et al.* 2006). Labels consist of the subfamily's number, followed by a letter for each distinct gene, then a number for each unique protein variant and an additional number referring to silent mutations. When applicable, an additional letter indicates variations in the intron. It was neither possible to differentiate between homologues and paralogues of the *B. pendula* PR-10 genes in *B. lenta*, *B chichibuensis*, *B. nigra,* and *B. schmidtii*, nor to identify homologous genes among these four species. Only the PR-10 genes in *B. costata* displayed a considerable degree of homology to the genes in *B. pendula,* but differentiating homologues and paralogues was not always possible. Instead of developing a denomination scheme for each species separately, we labeled the species with the subfamily's number, followed by a number for each unique protein variant and an additional number referring to silent mutations. This facilitates the protein analysis in which isoforms are distinguished rather than separate sequences or genes.

The observed number of genes varies among birch species. Some genes are probably missing from the dataset, but the PR-10 gene copy number also varies because of evolutionary processes such as duplications, extinctions, and recombination due to unequal crossing over. The overall clustering pattern appears to reflect a combination of such events. On several locations in the NJ tree (Fig. 6.1), genes from the same species tend to group close to each other. Examples are the clusters of highly similar sequences from *B. costata* in subfamily 01 and from *B. chichibuensis* in subfamily 03, which either reflect unequal crossing-over, gene conversion or duplication events. Two clear examples of unequal crossing-over are found in *B. populifolia*. The allele *01E01.01* is a recombination between the *01A* gene and the *01B* gene. The first part matches exactly to allele *01A01.01*, while the second part differs by 1 SNP from *01B01.01* with position 267 of the ORF as the point of recombination. Both original genes were also present. Similarly, the *03E01.01* allele is a recombination between the *03B* gene and the *03D* gene. In this case, the recombination probably occurred without gene duplication, since representative sequences of the original *03B* gene, as present in *B. pendula*, were absent.

PR-10 protein predictions

The sequence information indicates that some alleles will not be expressed as a full-sized protein. Twenty-two sequences contain an early stop codon or an indel in the ORF that results in a frame shift followed by an early stop codon. These sequences are denoted as pseudogenes, leaving 112 unique sequences with an intact ORF. The pseudogenes were excluded from the PR-10 protein search database although it cannot be excluded that they produce truncated proteins. We calculated K_a/K_s ratios within each subfamily. K_a/K_s ratios in the subfamilies 01, 02 and 03 were higher among the suspected pseudogenes than among alleles with an intact ORF (Table 6.2). This

points at an alleviated selection pressure in the pseudogenes. The other PR-10 subfamilies do not contain sufficient numbers of both genes and pseudogenes to make this comparison. The majority of the sequences had 5' splicing sites of AG:GT and 3' splicing sites of AG:GC, AG:GT or AG:GA. This is in concordance with known motifs for plant introns. Notable exceptions were: an AC:GT (*B. schmidtii*, 01pseudo04) and an AG:AT (*B. nigra*, 04var05.01a) 5' splicing site, an AC:GC (*B. schmidtii*, 01pseudo04) and a TG:GC (*B. nigra*, 02pseudo04) 3' splicing site, and two deletions (*B. costata*, 01pseudo05 and 02pseudo01) at the 3' end of the intron. Except for the AG:AT splicing site, all exceptions belonged to sequences that were denoted as pseudogenes, providing additional evidence for these designations.

TABLE 6.2 – Sequence conservation within subfamilies of the PR-10 family among eight Betula species. n = number of unique sequences. Ka/Ks ratio = ratio between non-synonymous and synonymous mutations. Range substitutions = minimum and maximum number of amino acid substitutions in the ORF in pair wise comparisons between sequences. n. d. = not determined

Subfamily	01	02	03	04	05	06	07
Sequences with an intact ORF							
n =	33	19	39	6	14	0	1
K_a/K_s ratio	0.18	0.27	0.10	0.36	0.09	*n. d.*	*n. d.*
Range substitutions	0 - 16	0 - 9	0 - 8	0 - 6	0 - 4	*n. d.*	*n. d.*
Average # substitutions	7.0	3.1	2.8	3.3	0.9	*n. d.*	*n. d.*
Pseudogene sequences							
n =	9	5	3	0	0	4	1
K_a/K_s ratio	0.38	0.30	0.20	*n. d.*	*n. d.*	0.57	*n. d.*

For sequences with an intact ORF, K_a/K_s ratios were between 0.09 and 0.36, depending on the subfamily (Table 6.2). This indicates the occurrence of strong purifying selection. The PR-10 alleles in birch encode a putative protein that consists of 160 amino acids, yielding a relative molecular mass of approximately 17 kDa. The only exception is 01var17.01 in *B. chichibuensis*, which contains an indel that results in the deletion of two amino acids. Consistent with the low K_a/K_s ratios, the allelic variation is lower at the protein level than at the nucleic acid level. Hence, the 112 unique genomic sequences encode 80 unique isoforms. The *PR-10.05* gene is an extreme example. Only four isoforms are predicted for this gene despite the presence of 14 allelic variants. One of the isoforms is predicted for all species except *B. nigra*. Parts of the PR-10 protein sequences are highly conserved, as is demonstrated in the alignment of the protein sequences of five PR-10 genes (one per subfamily) predicted for *B. pendula* (Fig. 6.2). The most prominent region lies between Glu_{42} and Ile_{56} in which only a single amino acid variation is found among all 80 isoforms. This region is characterized by a phosphate-binding loop with the sequence motive GxGGxGx. Additional conserved Glycine residues are found at positions 88, 89, 92, 110 and 111.

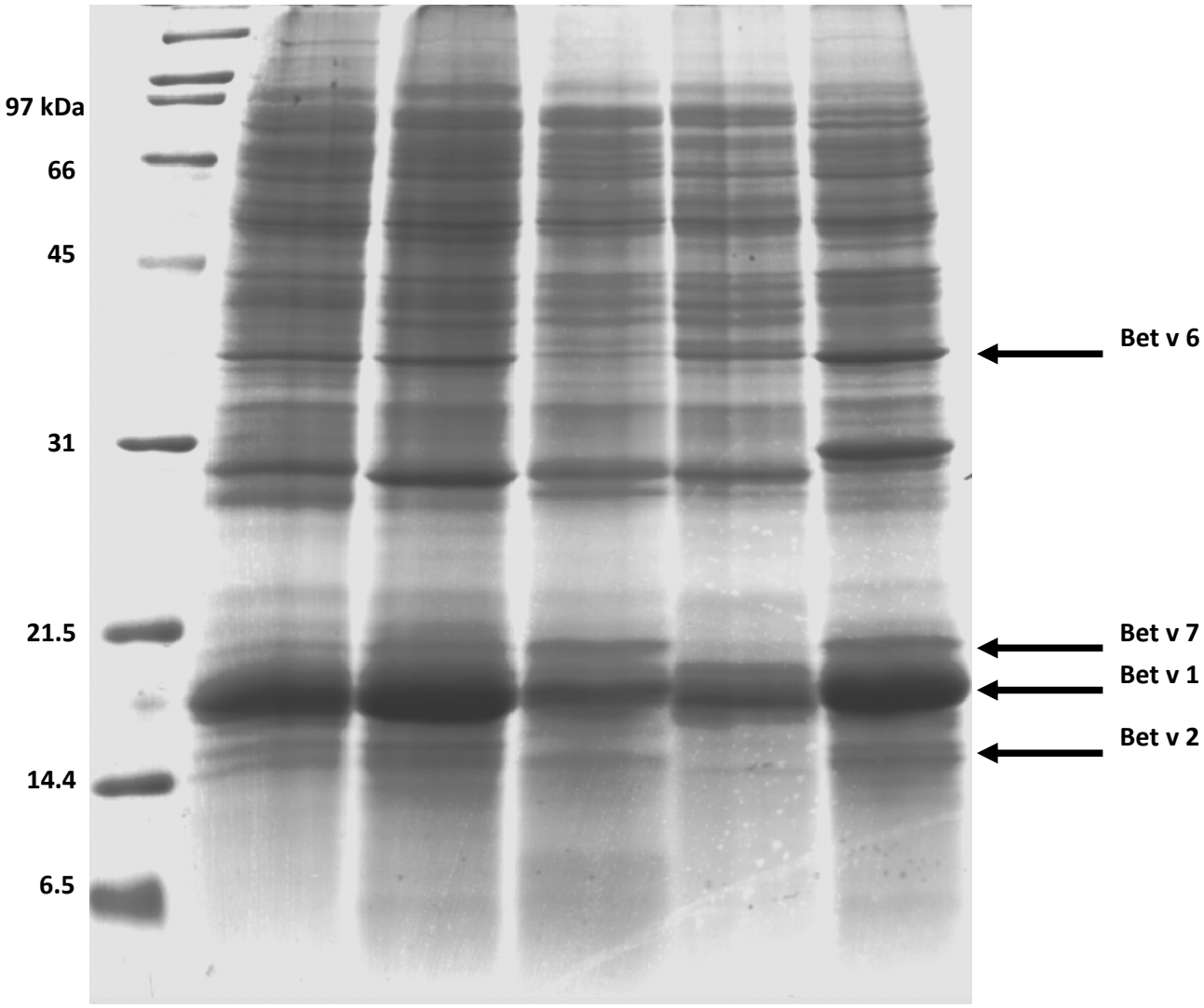
Figure 6.2 – Alignment of the theoretical tryptic peptides of the PR-10 proteins in B. pendula *'Youngii'. For simplification, one amino acid sequence is shown per subfamily. The labels of peptide fragments and their variants that are large enough too be detected are indicated in the figure. Variable amino acids are marked in black.*

```
       Fragment I                  III                 IV
       position 1-17               21-32               33-55
01A01  Ia:(M)GVFNYETETTSVIPAAR  LFK  IIIa: AFILDGDNLFPK  IVa: VAPQAISSVENIEGNGGPGTIK(K)
02A01  Ij:(M)GVFNYESETTSVIPAAR  LFK  IIIe: AFILDGDNLIPK  IVa: VAPQAISSVENIEGNGGPGTIK(K)
03A02  Iz:(M)GVFDYEGETTSVIPAAR  LFK  IIIe: AFILDGDNLIPK  IVz: VAPQAVSCVENIEGNGGPGTIK(K)
04 01  Iy:(M)GVFNDEAETTSVIPPAR  LFK  IIIz: SFILDADNILSK  IVx: IAPQAFK SAENIEGNGGPGTIK(K)
05 01  Ix:(M)GVFNYEDEATSVIAPAR  LFK  IIIy: SFVLDADNLIPK  IVv: VAPENVSSAENIEGNGGPGTIK(K)

       V                   VII                 VIII
       56-65               69-80               81-97
01A01  Va: ISFPEGFPFK  YVK  VIIa:(DR)VDEVDHTNFK  VIIIa: YNYSVIEGGPIGDTLEK  ISNEIK
02A01  Ve: ITFPEGSPFK  YVK  VIIk:(ER)VDEVDHANFK  VIIIk: YSYSMIEGGALGDTLEK  ICNEIK
03A02  Ve: ITFPEGSPFK  YVK  VIIz:(ER)IDEVDHVNFK  VIIIz: YSYSVIEGGAVGDTLEK  ICNEIK
04 01  Vz: ITFVEGSHFK  HLK  VIIy:(QR)IDEIDHTNFK  VIIIy: YSYSLIEGGPLGDTLEK  ISK EIK
05 01  Vy: ITFPEGSHFK  YMK  VIIx:(HR)VDEIDHANFK  VIIIx: YCYSIIEGGPLGDTLEK  ISYEIK

       X                                                   XVI             XVII
       104-115                                             138-145         146-159
01A01  Xa: IVATPDGGSILK  ISNK  YHTK  GDHEVK  AEQVK  ASK  XVIa: EM GETLLR  XVIIa: AVESYLLAHSDAYN
02A01  Xg: LVATPDGGSILK  ISNK  YHTK  GDHEMK  AEHMK  AIK  XVIb:(EK)GETLLR  XVIIa: AVESYLLAHSDAYN
03A02  Xz: IVAAPGGGSILK  ISNK  YHTK  GNHEMK  AEQIK  ASK  XVIz:(EK)AEALFR  XVIIa: AVESYLLAHSDAYN
04 01  Xy: IAAAPDGGSILK  FSSK  YYTK  GNISINQEQIK    AEK  XVIy:(EK)GAGLFK  XVIIz: AIEGYLL???????
05 01  Xx: IVAAPGGGSILK  ITSK  YHTK  GDISLNEEEIK    AGK  XVIx:(EK)GAGLFK  XVIIx: AVENYLVAHPNAYN
```

Figure 6.3 – SDS-PAGE with birch pollen extracts from (1) B. chichibuensis*, (2)* B. costata*, (3)* B. nigra*, (4)* B. lenta *and (5)* B. pendula*. Bands that were analyzed with Q-TOF LC-MS/MS and their identifications in the* B. pendula *sample are indicated by arrows. (M) LMW size marker proteins.*



Bet v 1 expression in pollen

The presence of Bet v 1 proteins was examined in pollen of *B. nigra*, *B. chichibuensis*, *B. lenta*, *B. costata* and *B. pendula* 'Youngii'. Pollen proteins were solubilized in an aqueous buffer and analyzed by SDS-PAGE. An intense protein band was observed after CBB-staining in each of the samples at the expected molecular mass of Bet v 1, between 16-18 kDa (Fig. 6.3), while other intense bands were visible at 28 kDa and 35 kDa. The efficiency of the protein extraction for Bet v 1 was investigated by SDS-PAGE analysis of the pellet that remained after extraction. No 16-18 kDa band was visible when the pellet content was separated by SDS-PAGE (not shown).

To establish the identity of the protein(s) in the 16-18 kDa band, this band was cut out from the lane of *B. pendula* (Fig. 6.3) and in-gel digested with trypsin. Q-TOF LC-MS/MS analysis of the tryptic peptides yielded multiple Bet v 1 isoforms (details given below). The bands just above and below the 16-18 kDa band were also cut out for sequencing and checked for the presence of Bet v 1. The lower band at 14 kDa was identified as birch profilin (Bet v 2; GenBank AAA16522; 2 peptides, coverage 24%) and did not contain any Bet v 1 fragments. The higher band at 19 kDa was identified as birch cyclophilin (Bet v 7; CAC841116; 3 peptides, coverage 28%) and contained minor traces of Bet v 1 (Bet v 1a; CAA33887; 1 peptide, coverage 14%). Bollen *et al.* (in press) detected a band of ~35 kDa when purified Bet v 1 was analyzed by SDS-PAGE, consisting of (dimeric) Bet v 1. We identified the intense band at ~35 kDa in our *B. pendula* extract as isoflavone reductase (Bet v 6; GenBank AAG22740; 19 peptides, coverage 49%) and did not detect any Bet v 1 fragments in this band.

Analysis of Bet v 1 isoforms by Q-TOF LC-MS^E

To elucidate the expression of separate Bet v 1 isoforms in pollen extracts, the respective 16-18 kDa bands were in-gel digested with trypsin. Trypsin cleaves proteins exclusively at the C-terminus of Arginine and Lysine. Fig. 6.2 shows the fragments I-XVII that are theoretically formed after tryptic digestion of isoforms from the five different subfamilies in *B. pendula.* Several fragments can be used to discriminate between isoforms of the five subfamilies on the basis of peptide mass and sequence. For Bet v 1 isoforms within a subfamily, the number of discriminating fragments becomes lower. A new mass spectrometric technique called Q-TOF LC-MS^E was used for simultaneous identification and quantification of peptides. With Q-TOF LC-MS^E, collision levels are alternating between low and elevated energy levels. The quantification of peptides occurs at low collision energy, while fragmentation spectra of the peptides are acquired at elevated collision energy to facilitate identification. A distinguished feature of LC-MS^E procedure is that information is obtained for all peptides. This contrasts MS/MS, in which a subset of peptides is selected for fragmentation. The software analysis used a search database for interpretation of the fragmentation spectra. This database contained the sequence information on all PR-10 isoforms

described in this paper and on previously described PR-10 isoforms from *B. pendula* (Schenk *et al.* 2006).

The results for the identification of separate isoforms are summarized in Table 6.3. Only a subset of the PR-10 proteins that were predicted from the gDNA sequence data was actually expressed in the pollen of the five investigated birch species. No evidence was found for expression of isoforms from subfamilies 03 to 07. We identified 22 Bet v 1 peptide fragments in *B. pendula* (Table 6.3). All identified peptides were predicted from the gDNA sequences. Eight detected peptides could be used to distinguish between subfamily 01 and 02. Among the seven *B. pendula* genes in subfamily 01 and 02 as found by gDNA sequencing, the expression of four (*01A*, *01B*, *01C*, *02C*) was confirmed (Table 6.3). Four peptides were specific for isoform 01B01, while one peptide was specific for isoform 02C01. Two peptides were specific for both isoforms of gene *01A*, while two others were specific for both isoforms of gene *01C*. Expression of the isoforms 02A01 and 02B01 cannot be distinguished, but one of them or both are expressed. Table 6.3 also shows identified peptide fragments that were not specific for particular isoforms, genes or subfamilies, and peptide fragments that were long enough to be detected, but were not observed in the tryptic digest. Information on absent fragments can be used to exclude the expression of particular isoforms. For example, multiple peptides that were specific for isoform 01D01 were not detected, excluding its expression. Finally, Table 6.3 shows that the sequence coverage of the expressed isoforms amounted to 71 to 79%.

In total, the expression of 14 isoforms was confirmed by unique peptides in the five birch species (Table 6.3). An additional 15 isoforms were potentially expressed, but lacked unique peptides to distinguish them from other isoforms. Several of these must be expressed, because part of the recovered peptides were traced back to two potentially expressed isoforms, as in the above example of the isoforms 02A01 and 02 B01 in *B. pendula*. The expression of five isoforms was ruled out, because multiple unique peptides from these variants were missing from the peptide mixture. Altogether, at least 4 to 6 isoforms were expressed in each of the five birch species. Two identified peptides in *B. costata* and one peptide from *B. nigra* did not match to sequences recovered from these species and are shown in Table 6.3 as belonging to an “unknown isoform”. This indicated that the alleles encoding these isoforms were missing from the sequence dataset of the species in question. Conflicting evidence was found for the expression of the isoforms 01var10 and 01var11 in *B. lenta*. Two peptides that were unique for these isoforms were detected, while three peptides that were expected if the isoforms would be expressed, were lacking. The most likely explanation is the expression of an allele that is missing from our database rather than the expression of 01var10 or 01var11.

TABLE 6.3 – Peptides fragments of PR-10 isoforms that were identified in pollen extracts from five Betula *species by Q-TOF LC-MS^E. Each isoform is displayed on a separate line. Some isoforms are encoded by the same gene, which is indicated in the third column. The gene label consists of the number of the subfamily to which the gene belongs, followed by a letter for each distinct gene. It was not possible to identify gene homologues across species, so these labels do not correspond to the same gene in different species. Theoretical tryptic peptide fragments are shown on top of the table and are labeled with Roman numbers as indicated in Fig. 6.2, which also displays representative amino acid sequences of the isoforms 01A01 and 02A01. Each variant of the tryptic peptide fragments is indicated with a letter in the table. Bold capital letters indicate that a fragment is unique for the isoforms from a particular gene. Bold italic letters indicate that a fragment is unique for the isoforms from a particular subfamily. Letters displayed between brackets indicate that a particular fragment was* ***absent*** *in the PR-10 mixture. Finally, the sequence coverage is shown in the last column and displays coverage of the total protein sequence, including also the fragments that are too small to be detected (II, VI, IX, XI, XII, XIII, XIV, XV).*

Species	Isoform	Fragment Gene	I	III	IV	V	VII	VIII	X [*2]	XVI	XVII	Sequence coverage
B. pendula	01A01	1A	**A**	***a***	a	***a***	***a***	**A**	a?	***a***	a	79%
	01A06	1A	**A**	***a***	a	***b***	***a***	**A**	a?	***a***	a	79%
	01B01	1B	**B**	**B**	a	***b***	***a***	**C**	**B**	***a***	a	79%
	01C04	1C	**D**	***a***	a	***a***	***a***	**D**	***c?***	***a***	a	71%
	01C05	1C	**D**	***a***	a	***a***	***a***	**D**	***c?***	***a***	a	71%
	01D01	1D	**(E)**	***a***	a	**(C)**	**(C)**	**(E)**	a?	***a***	a	-
	02A01	2A	***j***	e	a	e	***k***	***k***	***g?***	**(B [*3])**	a	74%
	02B01	2B	***j***	e	a	e	***k***	***k***	***g?***	***(c [*3])***	a	74%
	02C01	2C	***j***	***F***	a	e	***k***	***k***	a?	***(c [*3])***	a	74%
	03 *[1]		**(C), (*z*)**	e	**(*z*), (Y)**	e	**(Z), (*y*)**	**(*z*), (Y)**	**(z), (Y)**	**(*z*)**	a	-
	04 *[1]		**(Y)**	**(Z)**	**(X), (W)**	**(Z)**	**(X)**	**(X)**	**(X)**	**(Y)**	**(Z), (Y)**	-
	05 *[1]		**(X)**	**(Y)**	**(V)**	**(Y)**	**(W)**	**(W)**	**(W)**	**(X)**	**(X)**	-
B. chichibuensis	01var09	1A	***a***	***a***	a	***a***	***a***	***d***	a?	***a***	a	79%
	01var12	1B	**C**	***a***	a	***a***	***a***	***d***	a?	***a***	a	79%
	01var17	1C	***a***	***a***	a	***a***	**H**	**(H)**	***c?***	***a***	a	-
	01var18	1C	***a***	***a***	a	***a***	**H**	***d***	***c?***	***a***	a	71%
	01var19	1D	**(E)**	***a***	a	**(C)**	**(J)**	**(E)**	a?	***a***	a	-
	02var03	2A	***j***	e	**B**	e	***k***	***k***	a?	(c [*3])	a	74%
	02var08	2B	***j***	**F**	a	e	***k***	***k***	a?	(c [*3])	a	74%

Species	Isoform	Fragment Gene	I	III	IV	V	VII	VIII	X [*2]	XVI	XVII	Sequence coverage
B. costata	01var01	1A	**F**	***a***	a	***a***	***a***	***c***	a	***a***	a	79%
	01var02	1B	***a***	***a***	a	***a***	***a***	***c***	a	***a***	a	79%
	01var04	1C	***a***	**(C)**	a	***b***	**(E)**	***c***	**(D)**	***a***	a	-
	01var05	1D	***a***	***a***	a	***b***	***a***	***c***	**(E)**	***a***	a	71%
	01var13	1E	**D**	***a***	a	***a***	***a***	**D**	a	***a***	a	79 %
	Unknown		**E**									?
	02var05	2A	***j***	**(J)**	a	e	***k***	***k***	a	***(c[*3])***	a	-
	02var10	2B	***j***	**G**	a	e	***k***	***k***	a	***(c[*3])***	a	74%
	Unknown		**F**									?
B. lenta	01var10	1A	**G**	**A**	a	**(A)**	**(F)**	***d***	a	***(a[*3])***	a	60%
	01var11	1A	**G**	**A**	a	**(A)**	**(G)**	***d***	a	***(a[*3])***	a	60%
	01var16	1B	**B**	**D**	a	**D**	**A**	***d***	a	***(a[*3])***	a	74 %
	02var01	2A	***j***	e	a	e	***k***	***k***	a	***(c[*3])***	a	74%
	02var04	2B	***j***	**H**	a	e	***k***	***k***	a	***(c[*3])***	a	74%
	02var07	2C	***j***	**F**	a	e	***k***	***k***	a	***(c[*3])***	a	74%
B. nigra	01var03	1A	**B**	***a***	a	**B**	***a***	**F**	a	***a***	a	79 %
	01var06	1B	***a***	***a***	a	***a***	-	**C**	a	***a***	a	74 %
	01var07	1C	***a***	***a***	a	***a***	***a***	**C**	a	***a***	a	79 %
	01var08	1D	***a***	***a***	a	***a***	***a***	**D**	**F**	***a***	a	79 %
	02var06	2A	**J**	**K**	a	e	**K**	**K**	a	***(c[*3])***	a	74 %
	unknown			**F**								?

1 No specific peptide fragments were detected for any of the isoforms in subfamily 03 to 05 and these are, therefore, summarized into a single row and not displayed for the other species

2 Fragments X_a and X_g have exactly the same mass and cannot be distinguished. The peak of peptide X_c overlaps with the first isotope peak of peptide $X_{a=g}$ because they differ exactly 1Da in size and have the same charge, and X_c cannot be identified separately as a result

3 The XVI-peptides are not always detected because of their small size

Quantification by Q-TOF LC-MS[E]

We determined the peak intensities of Bet v 1 peptides in the tryptic digest for *B. pendula* 'Youngii' (Table 6.4). The relative amounts of individual isoforms can be deduced from these peak intensities. Not all identified fragments could be used for quantification, because the peak detection algorithm groups some peaks from different fragments together that are highly similar in mass and retention time. For example, fragment I_a (1854,91 Da) and VII_a (1854,89 Da) have a retention time that is only marginally different, causing a strong overlap in peak area. The relative amounts of two isoforms can be estimated directly: peptide III_f is unique for isoform 02C01 and comprises 17% of all fragment III-variants, while peptides III_b and X_b are unique for 01B01 and comprise 18-19% of all fragment III and X-variants. The isoforms 02A01 and 02B01 could not be

distinguished, but together are estimated to comprise 13% of the mixture based on fragment III_e. The relative amounts of the other isoforms were estimated indirectly. Isoform 01A06 and 01B01 share fragment V_b, which comprises 23% of all fragment V-variants. 01A06 is thus estimated to comprise 4-5% of the mixture. The ratio between 01B01 and 01C04 plus 01C05 can be deduced from fragment I_b. 01C04 plus 01C05 are thus estimated to comprise 6% of the mixture. This leaves 40-41% of the total amount of Bet v 1 for isoform 01A01.

Isoform 01A01 is identical to the isoform Bet v 1a, which was shown to have the highest IgE-reactivity by Ferreira *et al.* (1996). Pollen of *B. costata*, *B. nigra* and *B. chichibuensis* contained isoforms that are highly similar to this isoform. The isoforms 01var09 in *B. chichibuensis*, 01var06 and 01var07 in *B. nigra,* and 01var01 and 01var02 in *B. costata* differ by 1-3 amino acids from Bet v 1a. We determined the expression of individual isoforms in the tryptic digest from the Bet v 1 band from these species in a similar fashion as reported for *B. pendula*. The Bet v 1a-related isoform were estimated to comprise 38% (*B. chichibuensis*), 36-44% (*B. nigra*) and 36-41% (*B. costata*) of the total amount of Bet v 1. *B. lenta* differed from the other species, because the isoform most similar to Bet v1a differed by seven amino-acids. Expression of this isoform was estimated at 12-19%. Another major difference between *B. lenta* and the other species was the expression of subfamily 01 isoforms relative to subfamily 02 isoforms. In *B. lenta*, subfamily 02 accounted for 74-83% of the total amount of Bet v 1, while this was 25-40% in *B. pendula*, *B. nigra* and *B. chichibuensis* and 49-56% in *B. costata*.

Discussion

PR-10 gene family organization and evolution

We examined eight birch species for the presence and diversity of PR-10 genes by cloning and sequencing more than 100 clones per species. The eight species represent four different subgenera/groups in the genus *Betula* (Schenk *et al.* submitted) and cover a large part of the existing variation within the genus. PR-10 genes were found in each birch species, as was expected given the broad range of plant species in which PR-10 genes are found (Ekramoddoullah *et al.* 2000; Gao *et al.* 2005a; Huang *et al.* 1997). The PR-10 genes were organized into subfamilies, as previously reported for *B. pendula* (Schenk *et al.* 2006). Five subfamilies were recovered from all species. Two new subfamilies were identified, but these were restricted to two species and were mostly composed of pseudogenes.

TABLE 6.4 – Quantification of identified peptides in the pollen from B. pendula *'Youngii' by Q-TOF LC-MSE. Numbers indicate the relative amount of a peptide variant compared to the total amount of homologues fragments; amounts are averaged over the two duplicates. Quantification was not possible for part of the peptide variants[1,2] The displayed abundances indicate the relative amounts among those variants that could be quantified. n.q. = not possible to quantify.*

Species	Isoform	Fragment Gene	I *1	III	IV	V	VII	VIII *1	X *2	XVII	Direct estimate	Indirect estimate	Subfamily	Direct estimate
B. pendula	01A01	1A	Ia: n.q.	IIIa: 51	IVa: 100	Va: 46	VIIa: 75	VIIIa: n.q.	Xa+g+c: 82	XVIIa: 100	-	4-41%	01	68-75%
	01A06	1A	Ia: n.q.	IIIa: 51	IVa: 100	Vb: 23	VIIa: 75	VIIIa: n.q.	Xa+g+c: 82	XVIIa: 100	-	4-5%		
	01B01	1B	Ib: 69	**IIIb: 19**	IVa: 100	Vb: 23	VIIa: 75	VIIIc: n.q.	**Xb: 18**	XVIIa: 100	18-19%	-		
	01C04/ 01C05	1C	Id: 31	IIIa: 51	IVa: 100	Va: 46	VIIa: 75	VIIId: 100	Xa+g+c: 82	XVIIa: 100	-	6%		
	01D01	1D	**Ie: 0**	IIIa: 51	IVa: 100	**Vc: 0**	**VIIc: 0**	**VIIIe: 0**	Xa+g+c: 82:	XVIIa: 100	0%	-		
	02A01/ 02B01	2A	Ia: n.q.	**IIIe: 13**	IVa: 100	Ve: 32	VIIk: 25	VIIIk: n.q.	Xa+g+c: 82	XVIIa: 100	13%	-	02	25-32%
	02C01	2C	Ia: n.q.	**IIIf: 17**	IVa: 100	Ve: 32	VIIk: 25	VIIIk: n.q.	Xa+g+c: 82	XVIIa: 100	17%	-		

1 Quantification was only possible for part of the peptide variants, because of a similar mass of I_a (1854,91 Da) and $VIII_a$ (1854,89 Da), and I_j (1840,89 Da) and $VIII_C$ (1840,88 Da). Fragment $VIII_k$ overlaps with a keratin peptide.

2 Fragments X_a and X_g have exactly the same mass and cannot be distinguished. The peak of peptide X_C overlaps with the first isotope peak of peptide $X_{a=g}$ because they differ exactly 1Da in size and have the same charge, and X_C cannot be identified as a result

The genomic organization within the PR-10 subfamilies was complex. Differentiating between paralogues and homologues was not possible beyond closely related species. One likely explanation is concerted evolution, for which cladistic evidence was found (Fig. 6.1). Concerted evolution causes genes to evolve as a single unit whose members (occasionally) exchange genetic information through gene conversion or unequal crossing-over (Liao, 1999). Tandemly arranged genes have increased conversion rates, while this arrangement is a prerequisite for the occurrence of unequal crossing-over (Ohta 2000). Most *PR-10* genes in apple are arranged in a duplicated cluster (Gao *et al*, 2005), thus facilitating the main mechanisms for concerted evolution. We obtained two alleles that appear the direct result of unequal crossing-over between Bet v 1 genes. On a higher taxonomic level, cladistic evidence for concerted evolution is present in the overall gene tree of the PR-10 family (Wen *et al*, 1997), as sequence divergence is generally smaller between different genes from the same species than between genes from different species.

Nei and Rooney (2005) suggested that a mixture of recent gene duplications and purifying selection could also explain why tandem gene duplicates appear similar. In their model of birth-and-death evolution of genes, new genes arise due to gene duplications, evolve independently while undergoing purifying selection, and go extinct after becoming non-functional. Pseudogenes are characteristic for this process. The low K_a/K_s ratios clearly point at the occurrence of purifying selection. Pseudogenes are a common feature among the PR-10 genes in birch, since we recovered them from six out of eight species. As much as one-third of the recovered alleles in *B. nigra* had an interrupted ORF. We have not determined the potential expression of all these variants, since the truncated isoforms would have migrated outside the 16-18 kDa band in the SDS-PAGE. None were, however, detected in the 14 kDa band. Basically, all ingredients for the "birth-and-death" model are present, except that independent evolution is questionable. When duplicates arise from unequal crossing-over they combine genetic information from two genes, for which we found evidence in two instances. Also, the clustering of for example the *B. chichibuensis* alleles (Fig. 6.1) would suggest an extremely high number of recent duplications. Both processes of "birth-and-death" and concerted evolution could, therefore, be active in the PR-10 gene family. Regardless of the evolutionary processes, its outcome is clear: PR-10 proteins are homogenous as a group and even stronger so within subfamilies.

Bet v 1 expression

Which PR-10 genes and isoforms are actually expressed in pollen and are thereby the true Bet v1 allergens? We used Q-TOF analysis of peptides in the tryptic digest from Bet v 1 isoforms to investigate expression in pollen of five *Betula* species. Only peptides originating from isoforms in subfamily 01 and 02 were recovered from birch pollen, confirming predictions based on mRNA expression (Breiteneder *et al.* 1989; Friedl-Hajek *et al.* 1999; Swoboda *et al.* 1995b). Koistinen *et al.* (2002) showed that the genes from subfamily 03 and 05 were nevertheless expressed in roots

and leaves of *B. pendula*. The single gene in subfamily 05 that was sequenced in all eight birch species is homologous to *ypr10c,* which has a high basal transcription level in roots and a relatively lower basal transcription level in leaves (Poupard *et al.* 2001; Poupard *et al.* 1998; Swoboda *et al.* 1995c). Its expression is induced by copper stress (Utriainen *et al.* 1998) and during senescence in leaves (Valjakka *et al.* 1999). Several genes were obtained for subfamily 03 by screening gDNA, ranging from one gene in *B. schmidtii* to seven genes in B*. chichibuensis*. The genes PR-10.03C and 03D (=*ypr10a* and *ypr10b)* in *B. pendula* become transcriptionally upregulated upon infection of the leaves with fungal pathogens (Swoboda *et al.* 1995c). Their transcription is induced by wounding or auxin treatment in roots (Poupard *et al.* 2001; Poupard *et al.* 1998). No data have been reported about the expression of the sequenced PR-10 genes in subfamilies 04, 06 and 07.

The pollen expressed Bet v 1 genes are transcribed during the late stages of anther development (Swoboda *et al.* 1995a), but which factors induce transcription is unknown. Bet v 1 is an abundant pollen protein that is estimated to encompass 10% of the total protein content of the pollen from *B. pendula* (Larsen 1995). The Bet v 1 band was the most intense band in the SDS-PAGE of birch pollen extracts. Its exact abundance is difficult to estimate due to differences in extraction efficiency between different proteins. However, given the intensity of the Bet v 1 band and the low amount of proteins that were observed in the pellet, our results suggested that the abundance of Bet v 1 is higher than 10% of the total protein content and is likely to exceed 20%. The occurrence of Bet v 1 isoforms in *B. pendula* has been studied in a mixture of pollen from different trees by Swoboda *et al.* (1995b). They analyzed tryptic digests of purified Bet v 1 isoforms by Plasma Desorption Mass Spectrometry (PDMS), a technique which only reveals peptide masses. We examined pollen from individual trees and analyzed the tryptic digest of the Bet v 1 protein band by Q-TOF LC-MS^E, which reveals both peptide masses and underlying sequences. The ability to determine the peptide sequences, yields more accurate information on the expression of individual isoforms. We demonstrated that at least 4 to 6 isoforms were expressed in the pollen of the birch species *B. pendula*, *B. nigra*, *B. lenta* and *B. costata*. The actual number is likely to be higher since it was not possible to discriminate each individual isoform due to the high similarity between some isoforms.

Based on the PDMS peak area of peptides, Swoboda *et al.* (1995b) estimated that the relative amount of Bet v 1a (=PR-10.01A01) in the pollen mixture from *B. pendula* was at least 50% of the total amount of Bet v 1. Ferreira *et al.* (1996) estimated the relative amounts of different Bet v 1 isoforms by NH_2-terminal sequencing of purified natural Bet v 1. They report a ~2:2:1 ratio for isoforms that respectively contain Ser, Thr and Ile at the 7^{th} amino acid position. In *B. pendula* 'Youngii' pollen, this would corresponds to expression of the isoforms 02A01+02B01+02C01 : 01A01+01A06 : 01B01+01C04+01C05, respectively. We used Q-TOF LC-MS^E to quantify the relative expression of Bet v 1 isoforms. This requires that peptide fragments from these isoforms are of

equal length and do not have major difference in amino acid composition. Differences in detection efficiency of the mass spectrometer would otherwise distort quantification. This prerequisite is met by the PR-10 proteins due to their high homogeneity. We estimated that relative expression of 01A01 plus 01A06 in pollen of *B. pendula* 'Youngii' was 44-46%, while the abovementioned ratio was ~30%:45%:25% for *B. pendula* 'Youngii'. The similarity between our results and previously obtained estimates indicates that the quantities obtained *B. pendula* 'Youngii' are also representative for other *B. pendula* trees.

Except for the protein separation on SDS-PAGE, no purification step was included in our extraction procedure. Although this minimizes the chance that certain isoforms are lost during purification, the tryptic digest of the Bet v 1 protein band might be contaminated with peptides of other pollen proteins of a similar mass. Three peptides of the pollen allergen Bet v 7 were detected in the 16-18 kDa band, but the amount of Bet v 7 was estimated to be less than 2% of the amount of Bet v 1, based on the peak intensities of these peptides. All peptides with a high peak intensity could be attributed to Bet v 1 isoforms. Full sequence coverage of Bet v 1 isoforms cannot be achieved by only using trypsin as a protease, as smaller peptides will be lost during the extraction of peptides from the SDS-PAGE gel. Proteases that cleave at other sites would yield peptides that cover part of the missing protein sequence. Coverage with Q-TOF LC-MS^E was 71-79% for the *B. pendula* isoforms, which is slightly higher than the 57-60% coverage reported for Q-TOF MS/MS (Helsper *et al.* 2002).

Allergenicity

Ferreira *et al.* (1996) distinguished isoforms with a low, intermediate, and high IgE-reactivity. For these isoforms, expression of 01B01 (=Bet v 1d, low IgE-reactivity), 02C01 (=Bet v 1c, intermediate IgE-reactivity), 01C04 (=Bet v 1f, intermediate IgE-reactivity) and 01A01 (=Bet v 1a, high IgE-reactivity) was confirmed in the pollen of *B. pendula* 'Youngii' by identification of unique peptides (Table 6.3). Isoforms of all three types of IgE-reactivity were shown to be abundant and encompassed 35-38% (high), 22-24% (intermediate) and 18-19% (low) of the total amount of Bet v 1. This leaves 17-22% of the total Bet v 1 for isoforms with an unknown IgE-reactivity. We observed similar quantities in two other *B. pendula* cultivars as well (results not shown). Since *B. pendula* is known to be highly allergenic, the presence of isoforms with a high IgE-reactivity is apparently of determining influence on the allergenicity of the pollen. Interestingly, people that are not yet sensitized to *B. pendula* pollen come into contact with several abundant isoforms. However, the ability of these isoforms to provoke an IgE-mediated response varies in patients that have become sensitized (Ferreira *et al.* 1997; Wagner *et al.* in press). The factors that cause one isoform to develop into having a high IgE-reactivity and another isoform into having a low IgE-reactivity are currently unknown. The abundance of the isoforms may play a role, but as isoforms

with an intermediate or low IgE-reactivity still have considerable expression levels, this is unlikely to be the only factor.

The outlook for identifying birch trees that express only hypoallergenic isoforms is limited. The isoforms Bet v 1l and Bet v1d (=01B01) are currently known as hypoallergenic (Ferreira *et al.* 1996). The crystal-structure of Bet v 1l has been determined (Markovic-Housley *et al.* 2003) and its allergenicity has recently been tested on a large group of patients (Wagner *et al.* in press). Bet v 1l was, however, not identified in any of the examined species and may represent a sequencing artifact or unexpressed allele. 01B01 was only found in *B. pendula*, *B. populifolia* and *B. plathyphylla*. The most similar isoforms in the other species differed by at least five amino acids. In contrast, the highly allergenic isoform 01A01 was picked up in *B. pendula* 'Youngii', both as a gDNA sequence and as an expressed protein. *B. populifolia* also contains the 01A01 sequence. The other *Betula* species do not, but *B. chichibuensis*, *B. costata* and *B. nigra* contain highly similar isoforms that differ only by 1-3 amino acids from 01A01. A high similarity between isoforms increases the chance that they share epitopes. On the other hand, a few amino acid substitutions may influence the allergenicity drastically (Ferreira *et al.* 1998; Neudecker *et al.* 2003; Son *et al.* 1999). In all species, these isoforms are abundant as well. *B. lenta* forms an exception in which the closest relative to 01A01 has 95.5% similarity and differs by seven amino acids.

Conclusion

We identified 12 to 25 unique PR-10 sequences in eight different birch species. The genomic organization of the PR-10 gene family was similar between all examined birch species as five major subfamilies were found in all species. Application of Q-TOF LC-MS^E showed that genes from two large subfamilies (01 and 02) were expressed in birch pollen. We showed that Q-TOF LC-MS^E allowed fast screening of Bet v 1 isoforms in birch pollen by determining presence and relative abundance of individual Bet v 1 isoforms. The pollen of species from three out of four subgenera in the *Betula* genus contained a mixture of Bet v 1 isoforms, with abundant levels of isoforms that were similar to isoforms with a high IgE-reactivity. We predict that the allergenic potency of these species will be high. *B. lenta* (subgenus *Betulenta*) formed an exception and lacked isoforms that were similar to known isoforms with a high IgE-reactivity. This species and related species represent the most promising candidates for further screening of allergenicity by for example skin prick tests or a nasal challenge.

CHAPTER 7

Assessment of allergenicity of birch pollen from several species of the genus Betula

Martijn Schenk, Jan Cordewener, Twan America,
Jeroen Peters, René Smulders and Luud Gilissen

To be submitted

Abstract

Pollen of the European white birch (*Betula pendula*) is a major cause of hay fever, particularly in northern and central Europe. The allergenic potency of other birch species is unknown. Twenty-four birch trees, representing 15 species and covering a large part of the genetic variation in the genus *Betula*, were tested for their allergenicity. SDS-PAGE showed that a 17 kDa band, representing the major birch allergen Bet v 1, was the most abundant protein in all examined pollen extracts, while immunoblotting with pooled sera showed that this band was highly immune-reactive in all pollen extracts. The degree of immune-reactivity correlated well with the total amount of Bet v 1 in the extracts, which amounted to 44% to 61% of the total amount of extracted protein. Pollen extracts from different trees varied in the total protein content, presumably the result of variation in pollen quality due to pollen rupture and varying hydration during the extraction as was confirmed by flow cytometry and microscopic observations confirmed. Tryptic peptide fragments of Bet v 1 isoforms from pollen of 11 birch species were analyzed and quantified by Q-TOF LC-MS^E. Amino acid residues that are associated with high IgE-reactivity were detected in all species and were abundant among the isoforms in the Bet v 1 mixture. Overall, differences in immune-reactivity between birch species were too small to have clinical relevance, implying that all examined birch species will be allergenic. Development of hypoallergenic birch trees must thus rely on other approaches, such as selection or introduction of male sterility.

Introduction

The genus *Betula* encompasses over 30 tree and shrub species that are found in diverse habitats in the boreal and temperate climate zone of the Northern Hemisphere. Birch pollen is released in large amounts during spring and is a major cause of Type I allergies. Birch pollen contains various allergens that are termed Bet v 1 to Bet v 8. Bet v 1 is the major birch allergen in Northern Europe, and is a pathogenesis-related class 10 (PR-10) protein from the European white birch *(Betula pendula)* (Breiteneder *et al.* 1989; Jarolim *et al.* 1989). Other Fagales species, such as alder and hazel, produce pollen with PR-10 homologues that share epitopes with Bet v 1 (Niederberger *et al.* 1998). PR-10 proteins consequently constitute the largest group of aeroallergens (Breiteneder *et al.* 2000). Towards the south of Europe, profilin becomes more prevalent as an aeroallergen relative to PR-10 allergens (Moverare *et al.* 2002), although its clinical relevance remains under debate (Wensing *et al.* 2002). Birch also contains profilin (Bet v 2). The clinical relevance of the other birch allergens is unknown.

B. pendula occurs in Europe and is extensively investigated in relation to birch pollen allergy. Sensitization to birch is also reported across Asia and North America, where *B. pendula* is not present (Abe *et al.* 1997; Eriksson *et al.* 1998), but where other *Betula* species do occur. Various birch cultivars and species are also used as ornamental trees and planted outside their natural distribution range. Different *Betula* species and cultivars may vary in their allergenicity as variation in allergenicity has been found for different varieties from apple (Bolhaar *et al.* 2005a; Carnés *et al.* 2006; Marzban *et al.* 2005), peach and nectarine (Ahrazem *et al.* 2007), date palms (Kwaasi *et al.* 2000), and olive trees (Castro *et al.* 2003). However, variation in allergenicity is not always present as was shown for peanut (Koppelman *et al.* 2001). Allergenicity can be reduced as a result of a reduced concentration of allergens (Ahrazem *et al.* 2007; Carnés *et al.* 2006; Castro *et al.* 2003; Marzban *et al.* 2005), but also due to the absence of isoforms with a high allergenicity (Ferreira *et al.* 1996; Wangorsch *et al.* 2007).

There is limited information on potential differences in allergen composition among birch species and cultivars. Bet v 1 was estimated to account for 10% of the total protein content of *B. pendula* pollen (Larsen 1995). Various Bet v 1 isoforms are transcribed and expressed in *B. pendula* pollen (Breiteneder *et al.* 1989; Schenk *et al.* 2006; Swoboda *et al.* 1995b), including both isoforms with a low and with a high IgE-reactivity (Ferreira *et al.* 1996). A single *B. pendula* tree expresses a mixture of Bet v 1 isoforms with varying IgE-reactivity (Schenk *et al.* in preparation). The relative abundance of individual isoforms with varying IgE-reactivity within the mixture may affect the total allergenicity of the pollen. Q-TOF LC-MS^E can be used to quantify peptides (Silva *et al.* 2005), and this technique was previously shown to be able to distinguish and quantify tryptic peptides of Bet v 1 isoforms (Schenk *et al.* in preparation). Site-directed mutagenesis has shown that several amino acid residues of Bet v 1 were crucial for IgE-reactivity (Ferreira *et al.* 1998; Son *et al.* 1999).

Quantification of the abundance of these amino acid residues among Bet v 1 isoforms my aid in the prediction of allergenicity.

The aim of this study was to assess the allergenicity of 22 *Betula* species and cultivars by: (i) determining the antigenic and allergenic profile by SDS-PAGE and Western-blotting, (ii) determining the amount of Bet v 1 in a fluorescent array using polyclonal antibodies, and (iii) applying Q-TOF LC-MS^E to quantify the abundance of amino acid residues that affect IgE-reactivity of Bet v 1 isoforms among the isoforms in the pollen mixture.

*TABLE 7.1 – Birch (*Betula*) species and cultivars included in the study and their taxonomic classification*

Species/Subspecies/Cultivar	Subsection (Winkler 1904)	Subgenus (De Jong 1993)	Subgenus (Schenk *et al.* submitted)	Examined by LC-MS^E
B. chichibuensis	-	*Neurobetula*	Group I	Yes
B. nigra	*Costatae*	*Neurobetula*	Group II	Yes
B. lenta subsp. *lenta*	*Costatae*	*Betulenta*	Group III	Yes
B. alleghaniensis	*Costatae*	*Betulenta*	Group III	Yes
B. medwediewii	*Costatae*	*Betulenta*	Group III	Yes
B. pendula 'Youngii' (1)	*Albae*	*Betula*	Group IV-A	Yes
B. pendula 'Youngii' (2)	*Albae*	*Betula*	Group IV-A	Yes
B. pendula 'Fastigiata'	*Albae*	*Betula*	Group IV-A	Yes
B. pendula 'Tristis'	*Albae*	*Betula*	Group IV-A	Yes
B. platyphylla subsp. *szechuanica*	*Albae*	*Betula*	Group IV-A	Yes
B. platyphylla subsp. *mandshurica*	*Albae*	*Betula*	Group IV-A	No
B. pubescens	*Albae*	*Betula*	Group IV-B	Yes
B. pubescens subsp. *tortuosa*	*Albae*	*Betula*	Group IV-B	No
B. papyrifera	*Albae*	*Betula*	Group IV-B	Yes
B. papyrifera subsp. *cordifolia*	*Albae*	*Betula*	Group IV-B	No
B. costata	*Costatae*	*Neurobetula*	Group IV-C	Yes
B. davurica	*Albae*	*Neurobetula*	Group IV-C	Yes
B. humilis	*Nanae*	*Chamaebetula*	Group IV-C	No
B. albosinensis	*Costatae*	*Neurobetula*	Group IV-C	No
B. utilis	*Costatae*	*Neurobetula*	Group IV-C	No
B. utilis subsp. *jacquemontii*	*Costatae*	*Neurobetula*	Group IV-C	No
B. grossa	*Costatae*	*Betulenta*	Group IV-C	No
B. x Edinburgh	Hybrid	Hybrid	Hybrid	No
B. x purpusii	Hybrid	Hybrid	Hybrid	No

Methods

Pollen samples

Pollen samples were collected in April-May 2006 and 2007. Pollen samples from 24 birch trees were examined, representing fifteen different species and two hybrids (Table 7.1). The *Betula* genus is divided into three to five groups or subgenera (De Jong 1993; Järvinen *et al.* 2004; Schenk *et al.* submitted). In this study, all groups are represented by at least one species except subgenus *Betulaster* (De Jong 1993). Trees were located in the Botanical Garden of Wageningen (Wageningen, The Netherlands) or planted as ornamental trees along streets (Ede, The Netherlands). Pollen samples were collected per individual tree from multiple branches by shaking the flowering catkins inside 50 ml blue caps. Harvested pollen was sieved through a 90 µm mesh filter to eliminate extraneous plant material. Other plant parts were estimated to account for < 5% of the pollen mixture by light microscopy observations. Pollen samples were freeze dried before storage at 4 °C.

Flow cytometry

To establish the quality of the collected pollen, we analyzed pollen samples by flow cytometry. Twenty-five mg of pollen was suspended in 500 µl 0.1 M PBS (pH 7.4). Pollen samples were analyzed on a DakoCytomation CyAn high-performance flow cytometer, using Summit 4.0 software (DakoCytomation). Scatter characteristics and autofluorescence were recorded.

Protein extraction

Protein extracts were obtained by suspending 50 mg of pollen sample into 1 ml of 0.05 M Tris-HCl buffer (pH 7.5) following the protocol described by Schenk *et al.* (in preparation). After the final extraction step, the supernatant was collected and stored at -20 °C until use. The Bradford assay was used to determine the total amount of protein in the extract.

SDS-PAGE and Immunoblotting

To determine the antigenic profile of each pollen protein extract, crude extracts were run in duplicate on 15% w/v acrylamide SDS-PAGE gels with a 5% w/v stacking gel using the Mini-Protean II gel system (Bio-Rad). One set of gels was stained with Coomassie BB R-250, while the duplicate set was used for immunoblotting**.** SDS-PAGE Standards broad range (Bio-Rad) markers were used to determine relative molecular masses.

Serum samples (kindly provided by Y. Vissers, Wageningen University) were stored at -80 °C until use. Specific IgE levels to birch were quantified using the ImmunoCAP system (Phadia, Uppsala, Sweden). Sera were characterized for reactivity against Bet v 1, Bet v 2, and Bet v 4 plus Bet v 6. A serum pool was prepared with sera from four individuals that had total specific IgE levels to birch exceeding 50 kU/l. Reactivity against individual allergens by individual sera ranged from 41 to >100

kU/l for Bet v 1, ranged from 2.52 to <0.10 kU/l for Bet v 2, and was <0.10 kU/l for Bet v 4 plus Bet v 6.

For Western blotting, proteins were transferred to a nitrocellulose membrane (Millipore, USA). Transfer was performed with using a Mini Trans-Blot electrophoresis transfer cell (Bio-Rad) at a constant voltage of 100 V for 1 h. Membranes were blocked with blocking buffer (2% w/v milk powder + 0.1 M PBS; pH 7.4) and incubated for 1 h at room temperature. Pooled serum was diluted 1:3 in 0.1% w/v milk powder and 0.1 M PBS (pH 7.4) and added as primary antibody. Membranes were incubated at 15 °C overnight. Membranes were subsequently incubated for 1 h with mouse anti-human IgE (1:1000, Sigma-Aldrich) in dilution buffer (0.1% w/v milk powder + 0.1 M TBS; pH 8.2), followed by incubation with goat anti-mouse IgG antibodies (1:1000) conjugated with alkaline phosphatase (Sigma-Aldrich) in dilution buffer for 1 h. Membranes were washed after each incubation step. The phosphatase was detected using freshly prepared substrate solution, consisting of 5% w/v 5-bromo-4-chloro-3-indolylphosphate and 5% w/v nitroblue tetrazolium in DMF added to 0.1 M Tris-HCl, 5 mM $MgCl_2$, 100 mM NaCl at pH 9.6. Gel images were scanned and intensity of the bands was quantified by densitometry with ImageJ 1.38 software (http://rsb.info.nih.gov/ij/). One control sample was blotted on each Western blot to correct for possible differences in immunostaining between blots.

Microsphere Immuno Assay (MIA)

Polyclonal antibody (pAb) raised against Bet v 1 protein (kindly provided by L. Zuidmeer, Sanquin) was coupled to fluorescent microspheres. 5.0 x 10^6 carboxylated microspheres (Luminex MAGPlex^{TM}-C; Luminex Corp) were placed in a micro centrifuge tube, washed with dH_2O, and resuspended in 80 µl of 0.1 M monobasic sodium phosphate (pH 6.2). The bead suspension was activated by addition of 10 µl of 50mg/mL sulfo-NHS and 50 µl of 50 mg/mL 1-Ethyl-3-(3-Dimethylaminopropyl)-Carbodiimide hydrochloride (EDC) and incubation for 20 minutes. Excess of activating agents was removed by washing the beads twice with 50 mM -(*N*-morpholino)ethane sulfonic acid buffer (MES; pH 5.0). Beads were resuspended in 50 mM MES, and pAb antibodies were added to the bead suspension and incubated while mixing (by rotation) at room temperature for 2 hours. The beads were washed twice in storage buffer (PBS-TBN; pH 7.4) by incubation for 30 minutes while mixing by rotation, and subsequently stored in the dark at 4 °C.

Pollen protein extracts were diluted in PBS-TBN (Phosphate-Buffered Saline (pH 7.4) containing 0.02% Tween 20, 0.1% Bovine serum albumin, 0.02% azide). Extracts were assayed at dilutions of 1/250 and 1/400. 10^3 pAb-coupled microspheres were added to each sample. The solution was vortexed to assure that microspheres were homogenously dispersed throughout the solution and incubated for 60 min in the dark while gently mixing. After incubation, the beads were captured on a Dynal MPC 96S magnet and washed with PBS-TBN. Beads were resuspended in 100 µl PBS-

TBN containing biotinylated anti-Bet v 1 as a secondary antibody and incubated for 30 minutes while gently mixing. The beads were again captured and washed, and then resuspended in 100 µl PBS-TBN containing 2 µg/ml Streptavidin rPhycoerythrin. After 15 minutes of incubation, the samples were directly analyzed on the Luminex 100 machine using Applied Cytometry Systems software. A total number of 100 beads were measured per sample.

rBet v 1a (Biomay, Austria) was diluted 1/100 in assay buffer and then serially diluted 1/10 to give a final dilution of 1/100.000 in order to generate the control curve. Negative controls contained assay buffer without rBet v 1a. The fluorescent signal of the beads was quantified, and the control curve (Fig. 7.1) was determined by using a Logistic-5PL curve fit (Brendan Technologies). For each pollen extract, median fluorescent intensity was calculated and the control curve was used to establish the concentration of Bet v 1.

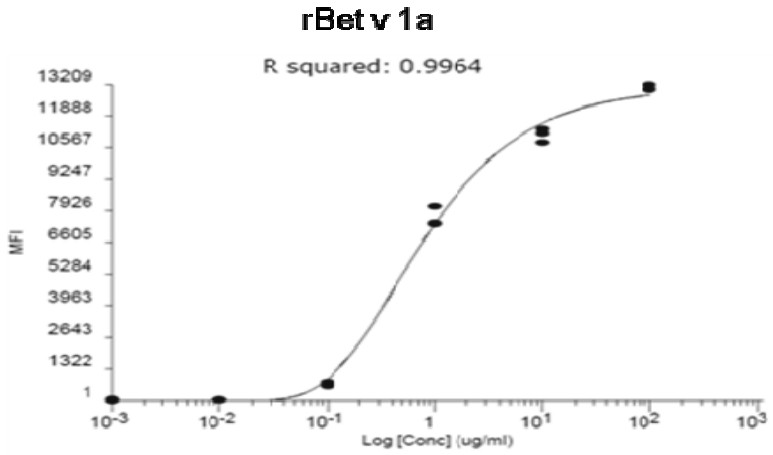


Figure 7.1 – Control curve of the microsphere immuno assay using polyclonal AntiBodies (pAB) raised against Bet v 1. The curve was generated with rBet v 1a and measured in three replicates

Q-TOF LC-MS^E of Bet v 1

For 14 birch trees, Q-TOF LC-MS^E was performed (indicated in Table 7.1). The protein bands at a relative molecular mass of ~17 kDa were cut out of the SDS-PAGE gel and processed essentially according to Shevchenko (1996). The band was sliced into 1 mm^3 -pieces. Proteins were reduced with DTT and alkylated with iodoacetamide. Gel pieces were dried under vacuum, and swollen in 0.1 M $NaHCO_3$ containing 0.05 M calcium chloride and sequence-grade porcine trypsin (10 ng/µl, Promega). After in-gel digestion at 37 °C overnight, peptides were extracted from the gel with 50% v/v acetonitrile, 5% v/v formic acid and dried under vacuum.

Tryptic digests were analyzed in high-throughput configuration using the Ettan^{TM} MDLC system (GE Healthcare) directly connected to a Q-TOF-2 Mass Spectrometer (Waters Corporation, UK). Samples (5 µl) were loaded on 5 mm x 300 µm ID Zorbax^{TM} 300 SB C18 trap columns (Agilent Technologies), and the peptides were separated on 100 µm i.d. x 15cm Chromolith CapRod monolithic C18 capillary columns (Merck) at a flow rate of approximately 1 µl/min. Solvent A

contained an aqueous 0.1% formic acid solution and solvent B contained 84% acetonitrile in 0.1% formic acid. The applied gradient consisted of isocratic conditions at 5% B for 10 min, a linear gradient to 30% B over 40 min, a linear gradient to 100% B over 10 min, and then a linear gradient back to 5% B over 5 min. MS analysis was performed in positive mode using ESI with a NanoLockSpray source. [Glu^1]fibrinopeptide B (1 pmol/µl) (Sigma) was delivered from the syringe pump (Harvard Apparatus, USA) to the reference sprayer of the NanoLockSpray source at a flow rate of 1 µl/min as lock mass. The lock mass channel was sampled every 10 seconds.

The Q-TOF-2 Mass Spectrometer was programmed to alternate between low and elevated levels of collision energy. Collision energy was 5eV in MS mode and increased in two steps from 28 to 40eV in MS^E mode. Measuring time in both modes was 0.9 s with an interscan delay of 0.1s. Unfragmented precursors predominate in low energy mode and fragmented ions of precursors are recovered in high energy mode. Each sample was analyzed in duplicate. MS^E data were analyzed according to the procedure described by Silva *et al.* (2005) using ProteinLynx Global SERVER (PLGS) v2.3 (Waters Corporation). Different peptide components were detected with an ion detection algorithm, and then clustered by mass and retention time, followed by normalization of the data. The search database described in Schenk *et al.* (in preparation) was used to identify peptides. Fixed (carbamidomethylation) and variable modifications (oxidation of Methionine) were taken into account. After processing by PLGS, the Exact Mass and Retention Time (EMRT) table was exported and reclustered using the PACP tool (De Groot *et al.* 2008) in order to correct potential misalignment and split peak detection errors, and to normalize retention time.

Results

Pollen harvesting and quality parameters

Birch pollen was harvested from 24 trees. The aim of this research was to compare allergenicity of birch pollen, and we therefore tried to minimize environmental factors that could influence pollen quality and/or allergenicity. Pollen was collected in the morning, at a similar height of the tree (1.5 to 4 meters), and covering branches around the tree. Most samples were collected at the same location (a botanical garden). However, onset of flowering varied between different birch species and pollen was consequently collected over a time span of one month. As varying weather conditions during this period may have affected pollen quality, we examined the pollen by light microscopy and noted several differences in appearance. Pollen appeared swollen and hydrated in some samples, while being dehydrated in other samples. In addition, some samples contained high amounts of pollen that appeared ruptured. Pollen was also analyzed by flow cytometry and two types of pollen were observed: the first type had a higher Forward SCatter (FSC) and a higher Side SCatter (SSC) than the second type. A higher FSC indicates a larger size, while a higher SSc indicates a higher cell complexity, so the second type is likely to represent the ruptured pollen particles. The abundance of both types of pollen particles varied among samples (data not shown).

SDS-PAGE electrophoresis of pollen proteins

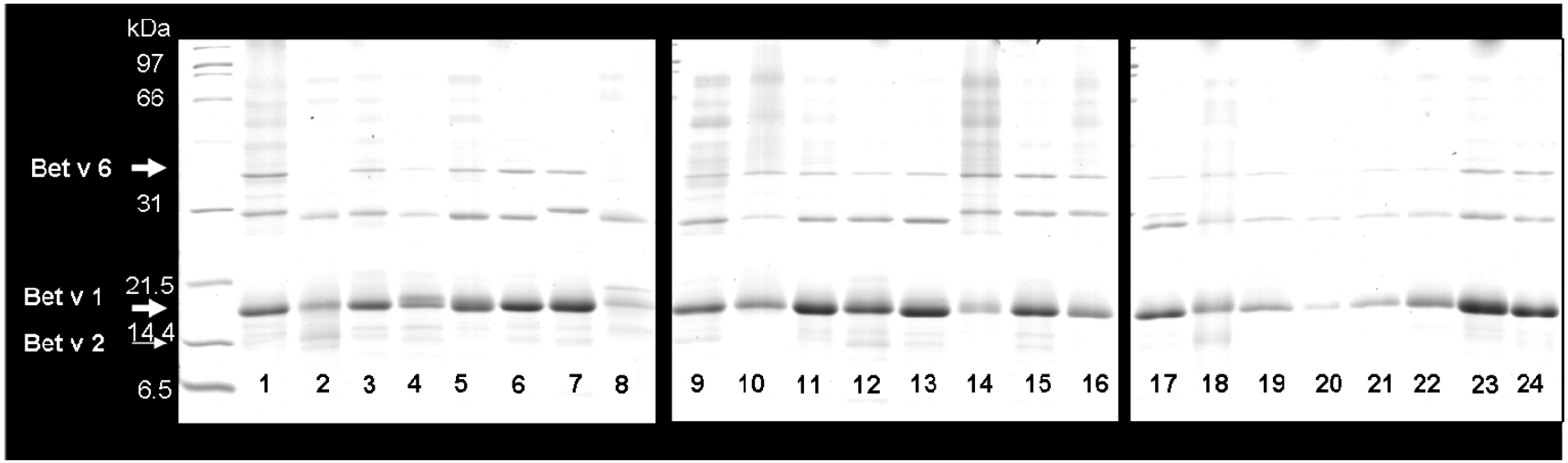
Similar profiles were observed when crude protein extracts from pollen of 24 birch trees were analyzed by SDS-PAGE (Fig. 7.2). The three most intense bands were found at approximately 17, 29 and 35 kDa. The 17 kDa band was detected in pollen extracts of all trees and its size matches to the expected size of Bet v 1. The 35 kDa band was observed in 22 out of 24 extracts and matches exactly to the expected size of Bet v 6. The 29 kDa band was observed in all extracts, but its size does not match to any known birch pollen allergen. Apart from the intense bands, a double band was observed in 22 of the 24 trees at 13 and 14 kDa, which matches to the expected size of Bet v 2 (14 kDa). The identity of the proteins at 13, 14, 17 and 35 kDa was previously confirmed by in gel tryptic digestion of the corresponding bands from extracts of *B. lenta* and *B. pendula* 'Youngii', and peptide sequencing using Q-TOF LC-MS/MS (Schenk *et al.* in preparation). The bands at 13 and 14 kDa both contained peptides of Bet v 2, while the band at 35 kDa was identified as Bet v 6. Due to the previously mentioned differences between pollen batches in extraction yield, the total amount of protein that was loaded on the gel varied between pollen extracts. In for example *B. costata*, a lower amount of total protein may have resulted in bands with intensities below the detection limit of CBB-staining, such as absence of the Bet v 2 and Bet v 6 bands in that sample (Fig. 7.2).

The ~17 kDa band varied in intensity between pollen extracts from different trees (Fig. 7.2). We examined whether these differences were related to differences in the Bet v 1 content or to differences in the total amount of extracted protein. Therefore, the amount of Bet v 1 relative to the total amount of pollen protein was determined for a subset of the samples using an independent method. The Bet v 1 concentration was determined in a Microsphere Immuno Assay (MIA) using polyclonal AntiBody (pAB) raised against rBet v 1. The relative amount of Bet v 1 amounted to 46 to 64% for the examined pollen samples (Table 7.2).The amount of Bet v 1 that was loaded onto the gel as determined in the MIA correlated well with the intensity of the 17 kDa band in the SDS-PAGE gels as determined by densitometry (results not shown).

TABLE 7.2 – Bet v 1 content of birch pollen extracts as determined in a Microsphere Immuno Assay using pAB. Total protein was determined by the Bradford Assay. Values are mean ± SE of replicates measured in duplicate at two different dilutions.

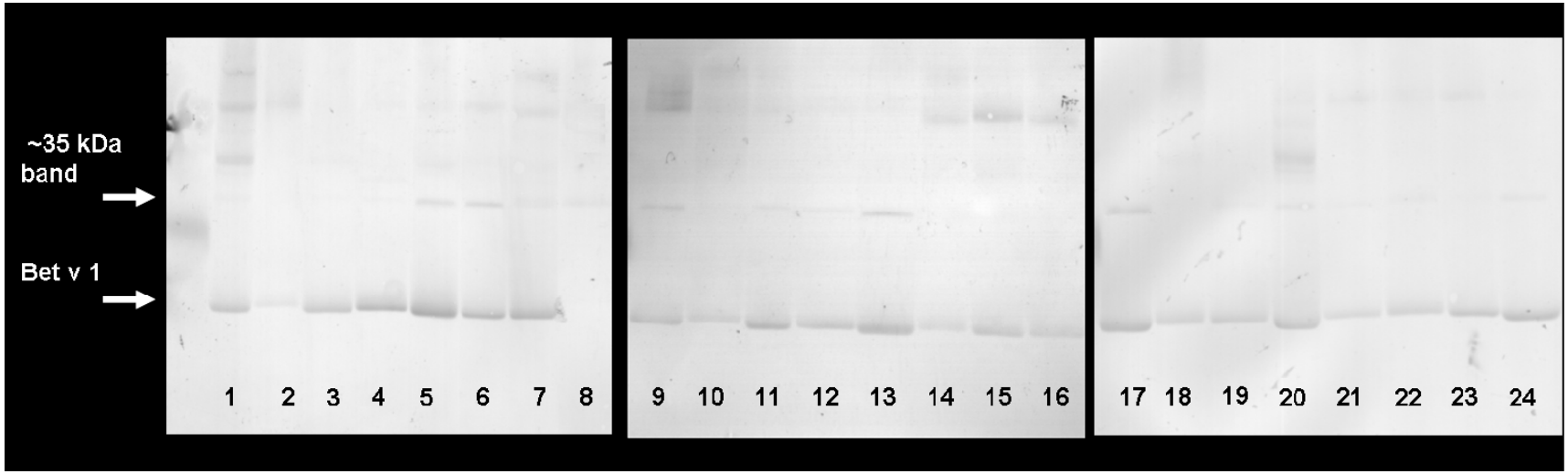
Species	Total protein (ng/µl)	Bet v 1 (ng/µl)	% Bet v 1
B. chichibuensis	4114	2459 ± 415	60
B. nigra	4058	2260 ± 181	56
B. lenta subsp. *lenta*	3762	1717 ± 87	46
B. pendula Youngii (1)	4027	2567 ± 228	64
B. costata	4448	2589 ± 162	58

Figure 7.2 - SDS-PAGE with pollen extracts of 24 birch trees. Bands matching to the expected size of the allergens Bet v 1, Bet v 2 and Bet v 6 are indicated by arrows. LMW size marker proteins are indicated in the first lane on the left.



(1) *B. pendula* 'Youngii'; (2) *B. nigra*; (3) *B. chichibuensis*; (4) *B. lenta*; (5) *B. alleghaniensis*; (6) *B. pendula* 'Fastigiata'; (7) *B. pubescens*; (8) *B. costata*; (9) *B. medwediewii*; (10) *B. davurica*; (11) *B. papyrifera*; (12) *B. pubescens tortuosa*; (13) *B. utilis*; (14) *B. platyphylla mandshurica*; (15) *B. pendula* 'Youngii'; (16) *B. pendula* 'Tristis'; (17) *B. albosinensis*; (18) *B. grossa*; (19) *B. humilis*; (20) *B. utilis*; (21) *B. plathyphylla szechuanica*; (22) B. x 'Edinburgh'; (23) B. x *purpusii*; (24) *B. papyrifera cordifolia*

Figure 7.3 – Western blots with pollen extracts of 24 birch trees. Bands that are recognized by a pool of sera coming from birch allergic patients indicated by arrows.



(1) *B. pendula* 'Youngii'; (2) *B. nigra*; (3) *B. chichibuensis*; (4) *B. lenta*; (5) *B. alleghaniensis*; (6) *B. pendula* 'Fastigiata'; (7) *B. pubescens*; (8) *B. costata*; (9) *B. medwediewii*; (10) *B. davurica*; (11) *B. papyrifera*; (12) *B. pubescens tortuosa*; (13) *B. utilis*; (14) *B. platyphylla mandshurica*; (15) *B. pendula* 'Youngii'; (16) *B. pendula* 'Tristis'; (17) *B. albosinensis*; (18) *B. grossa*; (19) *B. humilis*; (20) *B. utilis*; (21) *B. plathyphylla szechuanica*; (22) B. x 'Edinburgh'; (23) B. x *purpusii*; (24) *B. papyrifera cordifolia*

Western blot analysis of allergenicity across birch species

Western blots were prepared after SDS-PAGE using the same amount of crude protein extract as used for the CBB-stained gels shown in Fig. 7.2. After protein transfer to nitrocellulose, blots were incubated with a pool of sera of birch allergic patients. A ~17 kDa was recognized by IgE in the serum pool in all but one of the 24 birch pollen extracts (Fig. 7.3). No IgE-binding was observed in the extract of *B. costata*, but the total amount of protein loaded on the gel and, consequently, the amount of Bet v 1 loaded on the gel were very low and may have been under the detection limit for immunoblotting. An additional band at ~35 kDa was recognized in a subset of the extracts, as were several faint bands at higher molecular masses. The intensity of the 17 kDa band in the Western blots was determined with densitometry. The amount of blotted Bet v 1 was determined in a MIA using pAB. The blotted amount of Bet v 1 was correlated (R^2=0.55) with the intensity of the 17 kDa band in the Western blots (Fig. 7.4).

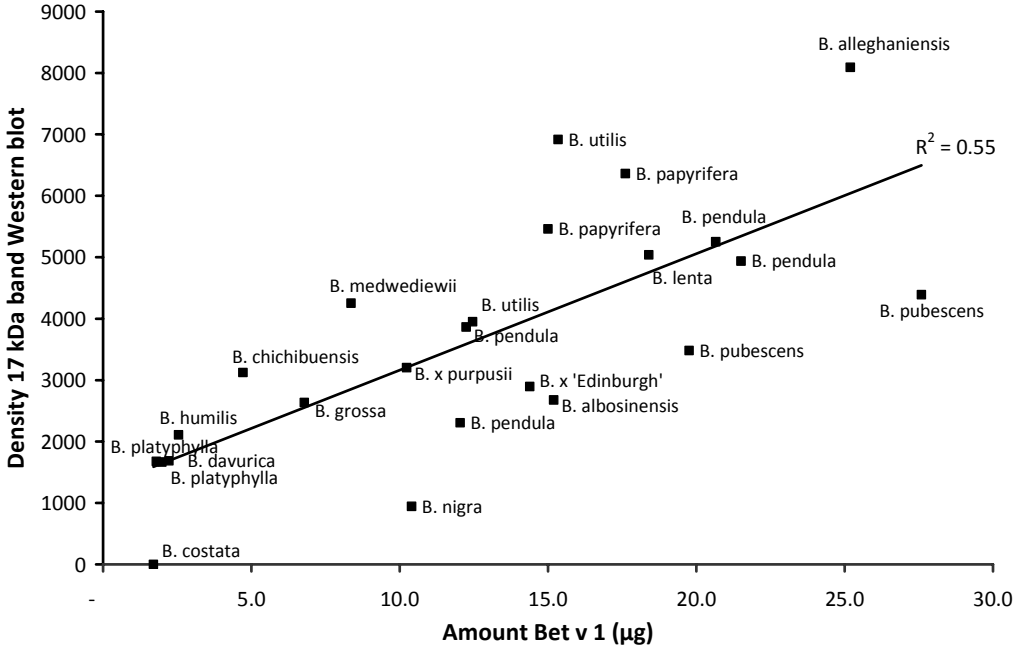


Figure 7.4 – Allergenicity of pollen extracts from 24 birch trees. The intensity of the 17 kDa band that is recognized in the Western blot by a pool of sera of birch allergic patients was determined with densitometry, while the blotted amount of Bet v 1 (µg) was determined in a Microsphere Immuno Assay with pAB against rBet v 1.

Presence and abundance of Bet v 1 isoforms

The 17 kDa bands from pollen extracts of 14 trees were in-gel digested with trypsin and analyzed with Q-TOF LC-MS^E. These trees represented 11 species and were selected across the genus *Betula*, and included one biological replicate of the *B. pendula* cultivar 'Youngii'. Tryptic Bet v 1

peptides were identified and quantified. Attention was focused on peptides that contained amino acid residues that are known to be associated with either a low or high IgE-reactivity. Using site-directed mutagenesis it has been demonstrated that at least 11 amino acid residues in the sequence of Bet v 1 or one of its PR-10 homologues have an effect on IgE-reactivity (Ferreira *et al.* 1998; Neudecker *et al.* 2003; Son *et al.* 1999; Spangfort *et al.* 2003). The presence and abundance of ten of these amino acid residues can potentially be examined by detection of tryptic peptides with Q-TOF LC-MS^E (the eleventh amino acid residue resides in a tryptic peptide that is too small for detection with the used LC-MS system). However, only five of the ten amino acid residues that affect IgE-reactivity were found to vary among the various isoforms that were detected in the investigated birch species. These five amino acids were located at position 28, 30, 57, 112 and 113 on the tryptic peptide fragments III, V and X (Fig. 7.5). The relative amounts of these fragments in the tryptic digests of the 14 pollen extracts were deduced from peak intensities in the Q-TOF LC-MS^E analysis.

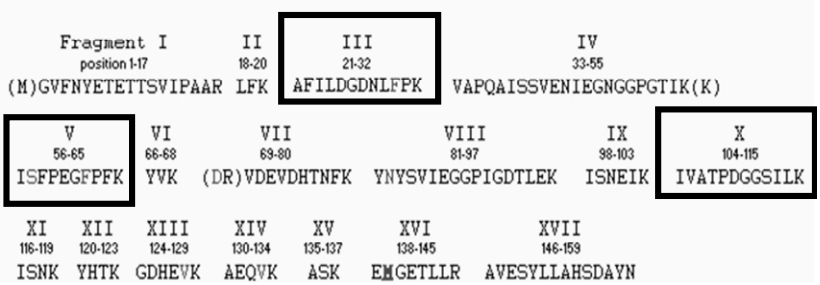


Figure 7.5 – Amino acid sequence and theoretical tryptic peptides of Bet v 1a (Acc. No. X15877). Tryptic fragments are indicated with Roman numbers. Other Bet v 1 isoforms contain Lysine and Arginine residues at the same positions, and will consequently yield tryptic peptides with the same number of amino acids as in the depicted sequence of Bet v 1a. The fragments that are marked with a box contain amino acid residues which affect IgE-reactivity and which are variable among the investigated birch species.

TABLE 7.3 (next page)- Quantification by Q-TOF LC-MS^E of Bet v 1 tryptic peptides containing amino acids that are associated with a high or low IgE-reactivity. Bet v 1 isoforms were extracted from pollen of 14 Betula trees (representing 11 species) and were in-gel digested by trypsin before analysis by Q-TOF LC-MS^E. Identified tryptic peptides (III, V and X) are labeled with Roman numbers. Variants of these peptides are indicated with an additional letter. Percentages in the table indicate the relative amount of each fragment variant compared to the total amount of variants in that species (i.e., per column); amounts are averaged over two duplicates. Relative amounts over 50% are highlighted in grey. The block of peptide variants is followed by a summary of relative amounts of amino acid residues associated with high IgE-reactivity, with low IgE-reactivity, or with an unknown effect.

Peptide Sequence	Fragment variant	Association with high or low IgE-reactivity	*B. pendula* 'Youngii' (1)	*B. pendula* 'Youngii' (2)	*B. pendula* 'Tristis'	*B. pendula* 'Fastigiata'	*B. platyphylla*	*B. pubescens*	*B. papyrifera*	*B. davurica*	*B. nigra*	*B. chichibuensis*	*B. lenta*	*B. alleghaniensis*	*B. medwediewii*	*B. costata*
AFILDGDNLFPK	IIIa		51	45	44	48	18	47	56	59	39	60	6	32	40	23
AFILGGDNLFPK	IIInew						18									
AFILDGDILFPK	IIId						3						19			
AFILDGDNLIPK	IIIe		13	27	20	41	25	38	39	29	20	23	43	14	28	60
AFILEGDNLIPK	IIIf		17	17	25		22	2	3			17	10		15	3
AFILDGDNVIPK	IIIh												22	48	5	
AFILDGNNLIPK	IIIk										42					
AFILDGDNLVPK	IIIb		19	12	11	11	14	14	3	4						
AFILEGDTLIPK	IIIg									8				6	12	14
(K)ISFPEGFPFK	Va		46	39	45	43	33	31	37	33	46	61		52	13	40
(K)ISFPEGFPFR	Vd												20		24	
(K)INFPEGFPFK	Vb		23	15	11	14	12	24	12	4	16			20	10	5
(K)INFPEGIPFK	Vc						5			11					10	
(K)ITFPEGSPFK	Ve		32	46	44	42	49	45	51	52	38	39	80	28	43	56
IVATPDGGSILK LVATPDGGSILK IVATPNGGSILK	Xa+Xg+Xc [1]		82	95	80	87	86	91	95	81	86	100	100	72	89	100
IVETPNGGSILK	Xf										14					
IVATPDGGCILK	Xnew								5	19				7	9	
IVATPDGGCVLK	Xb		18	5	20	13	14	9						21	2	
Residue 28	N=Asn	High	100	100	100	100	97	100	100	92	100	100	81	94	88	86
	T=Thr	Low								8				6	12	14
	I=Ile	Unknown					3						19			
Residue 30	F=Phe	High	51	45	44	48	39	47	56	59	39	60	25	32	40	23
	V=Val	Low	19	12	11	11	14	14	3	4						
	I=Ile	Unknown	30	44	45	41	47	40	42	37	62	40	75	68	60	77
Residue 57	S=Ser	High	46	39	45	43	33	31	37	33	46	61	20	52	37	40
	N=Asn	Low	23	15	11	14	18	24	12	15	16			20	20	5
	T=Thr	Unknown	32	46	44	42	49	45	51	52	38	39	80	28	43	56
Residue 112	S=Ser	High	82	95	80	87	86	91	95	81	100	100	100	72	89	100
	C=Cys	Low	18	5	20	13	14	9	5	19				28	11	
Residue 113	I=Ile	High	82	95	80	87	86	91	100	100	100	100	100	79	98	100
	V=Val	Low	18	5	20	13	14	9						21	2	
Average residue at position 28,30,57,112,113		High	72	75	70	73	68	72	78	73	77	84	65	66	70	70
		Low	16	7	12	10	12	11	4	9	3			15	9	4
		Unknown	12	18	18	17	20	17	18	18	20	16	35	19	21	27

1 Fragments X_a and X_g have exactly the same mass and cannot be distinguished. The peak of peptide X_c overlaps with the first isotope peak of peptide $X_{a=g}$ because they differ exactly 1Da in size and have the same charge, and X_c cannot be identified separately as a result

Table 7.3 displays the identified variants of the tryptic peptides III, V and X, and their relative quantities. The first amino acid position which is known to influence IgE-reactivity is located at position 28 in Fragment III. For this fragment, the variants III_a and III_e were present in all 14 examined birch trees, while the isoforms III_b and III_f to III_h were only detected in a subset of trees. Variant III_d was only observed in *B. platyphylla* and *B. lenta,* and variant III_k only in *B. nigra.* One new variant of fragment III was detected by LC-MS^E, displaying an amino acid substitution at position 25 (Gly^{25} for Asp^{25} or Glu^{25}). Most detected variants of fragment III contained the amino acid Asn^{28}. Having this amino acid at position 28 is associated with a high IgE-reactivity. The variants III_d and III_e form an exception and contain Ile^{28} and Thr^{28}, respectively. Having the amino acid Thr^{28} is associated with a low IgE-reactivity, while the effect of Ile^{28} on IgE-reactivity is unknown. The variants III_d and III_e were observed in pollen of six trees, but their relative abundance was low in each of these trees, and varied between 3 and 19% of all detected variants. Similar comparisons can be made for the other amino acid positions that influence IgE-reactivity. These results are summarized in the lower part of Table 7.3, which displays the relative abundance of the amino acids that are associated with either a high, low or unknown IgE-reactivity. All amino acids residues that are associated with high IgE-reactivity were identified in each of the examined species and their abundance was also high. The residues Ser^{112} and Ile^{113} were even present in 72%-100% and 79%-100% of the Bet v 1 isoforms, respectively. Amino acids that were associated with low IgE-reactivity were rare (Table 7.3). For example, *B. chichibuensis* did not contain any of these amino acid residues. Amino acids associated with a high IgE-reactivity were present in 65% to 84% of the isoforms when averaged over all five amino acid positions.

Discussion

SDS-PAGE analysis of pollen extracts revealed a protein band representing the major birch allergen Bet v 1 in each of the 24 birch trees that were examined. These trees represented 15 different birch species and four subgenera in the genus *Betula* (Schenk *et al.* submitted) and consequently cover a large part of the variation in the genus *Betula*. This strongly suggests that all *Betula* species will contain Bet v 1 in their pollen as was already predicted from the sequencing data of eight species (Schenk *et al.* in preparation).

Total amount of Bet v 1 protein across species

Bet v 1 is abundantly expressed in pollen of the examined species, and is visible as an intense band at 17 kDa after SDS-PAGE analysis and CBB-staining. The exact quantification of the amount of Bet v 1 was not straightforward, as the total amount of extractable protein from birch pollen was highly variable between different trees. Variation in protein content between pollen that was harvested from different birch trees is most probably due to variation in pollen quality. Several factors may account for this variation in pollen quality. Although care was taken to minimize variation caused by the collection procedure, not all variation could be excluded. Variation in

flowering time between different birch species could not be circumvented as individual trees produced pollen for only a period of ± 5 days, while differences in the onset of flowering were much larger than 5 days. For example, *B. lenta* started flowering two weeks after *B. plathyphylla* had started, while *B. medwediewii* started flowering even four weeks later than *B. plathyphylla*. As a result, pollen has been exposed to different weather conditions before sampling. Especially contact with water may have lead to differences in pollen quality, as birch pollen are known to display abortive pollen germination upon contact with rainwater (Grote *et al.* 2003). Two types of pollen grains were observed by light microscopy and flow cytometry, one being large in size and being one small in size. Both types varied in abundance between samples. The small type of grains appeared to represent ruptured pollen based on light microscopy observations, and may have released its protein content prior to extraction. A second source of variation is that some pollen grains were hydrated prior to extraction in some samples, resulting in a lower protein concentration per pollen grain.

We applied a mild extraction procedure for the isolation of Bet v 1 from pollen, which was shown to yield good results in *Betula* species (Cadot *et al.* 1995; Schenk *et al.* in preparation). The efficiency of the protein extraction was investigated by SDS-PAGE analysis of the pellet that remained after extraction to exclude the possibility that the extraction procedure was not rigorous enough to extract all Bet v 1. Several faint bands of high molecular weight proteins were observed in the pellet, but no 17 kDa band was visible (results not shown). Overall, we consider it most likely that differences in total protein content were the result from differences in pollen quality at the time of collection.

Differences in the Bet v 1 content

When differences in Bet v 1 content are observed between pollen of different *Betula* species, this could have a genetic basis, but may also be due to environmental influences. The pollen-expressed Bet v 1 genes are transcribed during the late stages of anther development (Swoboda *et al.* 1995a), but which factors induce transcription is unknown. Factors that may influence the allergen composition of pollen grains are the orientation on the trees, the location of the flowers on the tree, and hours of sunlight. We tried to collect all pollen in the morning, at approximately the same height (1.5 to 4 meters), while sampling around the tree, but this did not necessarily remove all environmental influence. For Ole e 1, an olive pollen allergen, the allergen content was tested for five consecutive years (Fernández-Caldas *et al.* 2007). Allergen levels were variable between different olive varieties, but stable within each variety for consequetive years and it was concluded that in this case, variation among trees was due by genetic differences.

Based on the intensity of the 17 kDa band after SDS-PAGE and CBB-staining, we concluded that Bet v 1 was by far the most abundant protein in all examined pollen extracts. For different birch

species, variation in the relative amount of Bet v 1 was limited. Bet v 1 was estimated to account for as much as 44 to 61% of total protein in our extracts. Larsen (1995) estimated the Bet v 1 content of *B. pendula* pollen to be 10%. The large difference between our and their results can in part be explained by overestimation of the concentration in our extracts due to non-extracted proteins (that were observed as faint bands after SDS-PAGE of the pellet), but the faintness of the bands suggests that their amount was only minor and the real Bet v 1 concentration is likely to be between both estimates.

IgE-reactivity of different Bet v 1 isoforms

Bet v 1 isoforms differ in their IgE-reactivity (Ferreira *et al.* 1996; Wagner *et al.* in press). We identified and quantified tryptic peptides of Bet v 1 in the pollen of 14 *Betula* trees (covering 11 species) by Q-TOF LC-MS^E. Several *Betula* species were previously shown to contain a mixture of Bet v 1 isoforms (Schenk *et al.* in preparation). Notably, *B. lenta* showed relatively large differences in isoform composition compared to the other species. Rather than focusing on isoforms as a whole, the current analysis focused on particular amino acids that affect IgE-reactivity and their abundance among the isoforms.

Eleven amino acids positions are known to affect IgE-reactivity of Bet v 1 isoforms (Ferreira *et al.* 1998; Neudecker *et al.* 2003; Son *et al.* 1999; Spangfort *et al.* 2003). Five of these positions were found to vary among the examined species, while one position could not be evaluated. The other five positions were invariable and each contained the residue that was associated with high IgE-reactivity. Regarding the variable positions, the residues associated with high IgE-reactivity were recovered in pollen of all examined birch species, and were on average, the most abundant residues (65% to 84%) among the Bet v 1 isoforms in the pollen mixture. This suggests that the largest part of the Bet v 1 mixture is composed of Bet v 1 isoforms with a high IgE-reactivity. However, all mutations were studied in the context of Bet v 1a, and many of the other species contain isoforms with additional amino acid substitutions, which effect on allergenicity is unknown. Potentially, more amino acids positions can be identified that affect allergenicity and are variable under natural circumstances. In either case, the similarities between the Bet v 1 mixtures of different species are very large, and are unlikely to result in significant differences in allergenicity.

Allergenicity of birch pollen extracts

A Microsphere Immuno Assay was used to measure Bet v 1 amounts. A similar method was previously used to quantify six indoor allergens (Earle et al., 2007), who indicate a high correlation (R^2=0.90-0.99) between the results obtained by the fluorescent array and an ELISA. Densitometry was used to quantify the intensity of the Bet v 1 band in the immunoblots. A correlation coefficient of 0.55 (Fig. 7.4) between Bet v 1 amounts and immune-reactivity can be considered as

a good correlation given that the densitometry method is not highly accurate. Small variations are observed in immune-reactivity between pollen of different species. For example, *B. nigra* scored slightly lower, while *B. alleghaniensis* scored slightly higher than would be expected based on the amount of Bet v 1 in the pollen extract. Koppelman *et al.* (2001) indicated that clinically relevant differences for the major peanut allergen should be in the order of a 100x to 1000x reduction in IgE-reactivity. The largest difference that we observed in the immunoblot was less than a 10x difference in immune-reactivity, suggesting that the differences between pollen of different *Betula* species were too small to have clinical relevance.

A faint signal was obtained with IgE in the serum pool for a band at approximately 35 kDa in some extracts. This band either represented dimeric Bet v 1 (Bollen *et al.* in press) or the birch allergen Bet v 6=isoflavone reductase (Karamloo et al., 1999). In pollen extracts of *B. lenta* and B*. pendula, t*he 35 kDa band was shown to be composed of Bet v 6 (Schenk *et al.* in preparation). However, the sera that contributed to the serum pool were characterized for IgE-reactivity against Bet v 6 and no specific IgE against this allergen was present in any of the individual sera, excluding this option. Bollen *et al.* (in press) detected a 35 kDa band with a high immune-reactivity in purified natural Bet v 1, even under reducing conditions in the SDS-PAGE. This was shown to be a dimer of Bet v 1. This dimer may not have been detected by Q-TOF LC-MS/MS because the amount of dimer was too small.

Conclusion

All examined birch species contain large amounts of a mixture of Bet v 1 isoforms in their pollen. Isoforms that are predicted to have a high IgE-reactivity predominate in the mixture of each species, despite variations in the composition of the mixture between species. As a result, immune-reactivity is expected to be high for pollen of all investigated birch species, as is confirmed by immunoblotting experiments. Concerning the allergenicity of pollen of different birch species, we could not determine any differences that are large enough to have clinical relevance, implying that all examined birch species are allergenic. Development of hypoallergenic birch trees must, thus, rely on other approaches, such a selection or introduction of male sterility.

CHAPTER 8

General discussion

General context: birch pollen allergy

Birch pollen allergy and the related Oral Allergy Syndrome (OAS) affect a substantial part of the population, particularly in northern and central Europe. Selection or development of birch trees with reduced allergenicity may contribute to alleviation of birch pollen-associated complaints. Birch pollen contains multiple allergens that are termed Bet v 1 to Bet v 8. The majority of patients who are allergic to birch react to the major birch allergen Bet v 1. In the research presented here, the emphasis has been on this major allergen, which is a member of the PR-10 gene family (Breiteneder *et al.* 1989; Jarolim *et al.* 1989). Allergic reactions to Bet v 1 have a strong link to OAS, due to an IgE-mediated cross-reaction between Bet v 1 and PR-10 proteins in plant foods (Ferreira *et al.* 2004; Wensing *et al.* 2002). The Mal d 1 allergen in apple is a good example of this cross-reactivity, as up to 70% of the people who suffer from birch pollen allergy experience OAS, particularly after eating apples (Ebner *et al.* 1991; Ortolani *et al.* 1988). The Mal d 1 allergens have been well-characterized at the genomic level (Gao *et al.* 2005a). Apple cultivars differ in allergenicity and the cultivar Santana has been shown to have a low allergenicity (=hypoallergenic) (Bolhaar *et al.* 2005a). Next to selection of hypoallergenic apples from existing cultivars, Genetic Modification (GM) can be applied to develop apples with reduced allergenicity. Silencing of Mal d 1 genes has been carried out by RNA interference (Gilissen *et al.* 2005). The introduction of hypoallergenic food products should take the preferences and responses of allergic consumers into account, particularly when their development involves GM.

This chapter will discuss the molecular characterization of the PR-10 genes and the finding that birch pollen contains a mixture of Bet v 1 isoforms with varying IgE-reactivity, which has relevance to sensitization and cross-reactivity. The perspectives for selection and breeding of hypoallergenic birch trees will be subsequently discussed. This is followed by a discussion on the development of hypoallergenic products, with apple as a major example. Consideration will be made of how allergic consumers react to novel hypoallergenic foods, particularly when GM is applied to develop these products.

Validation of allergen sequences

Investigating diversity within a group of closely related sequences, such as the PR-10 gene family, requires accurate sequences to distinguish between individual family members. In Chapter 4, *Pfu* polymerase was used for PCR amplification. *Pfu* polymerase has proofreading capacity, which reduces the number of mis-incorporated bases. However, several recovered sequences appeared to be recombinations between other sequences (Fig. 8.1). These sequences occupied intermediate positions between well-defined clusters in a phylogenetic tree. We also identified clusters of

multiple sequences that differed by only one or two SNPs, suggesting the mis-incorporation of bases during the PCR. A strategy of validating sequences in independent PCRs was therefore adopted to determine which sequences could be confirmed and which ones were potential PCR artifacts. In addition, reducing the number of cycles, as suggested by Zylstra *et al.* (Zylstra *et al.* 1998) proved an efficient solution to lower the occurrence of recombination artifacts (Chapter 4). Recombination artifacts may originate from incomplete primer extension and re-annealing to a different template (Judo *et al.* 1998) or from strand switching between different templates (Odelberg *et al.* 1995). As *Pfu* polymerase is known to generate more, and more complex recombination artifacts than *Taq* polymerase (Whinnett *et al.* 2003; Zylstra *et al.* 1998), we switched to amplification by *Taq* polymerase in the follow-up study described in Chapter 6. Indeed, this reduced the occurrence of recombination artifacts from 23% (=22+1%) to 7% (=5+2%) (Table 8.1). Meanwhile, the number of PCR artifacts with mis-incorporated bases increased from 14% to 28%. The total frequency of artifacts was thus similar. However, sequences with recombination artifacts combine phylogenetic signals from two sequences and are more difficult to identify than sequences with a mis-incorporated base (if the data set is sufficiently redundant). As most allergens are part of gene families (Breiteneder *et al.* 2004) and multiple sequences are likely to be amplified in the same PCR, the application of independent PCRs to facilitate sequence validation is recommended when sequencing allergens.

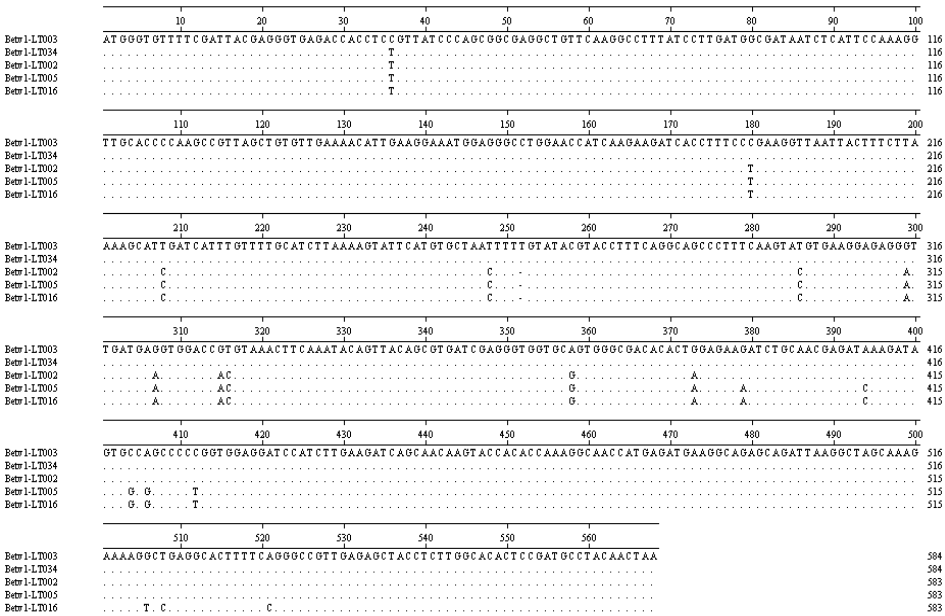


Figure 8.1 – Sequences that illustrate the potential occurrence of PCR recombination artifacts. The upper (BpLt003) and lower most (BpLt016) sequence are confirmed in independent PCRs, while the other sequences are recovered from one PCR only and appear to represent recombination artifacts. The three artifact sequences take an intermediate position between the confirmed sequences with position 180, 379 and 506 as the respective points were recombination can be detected.

TABLE 8.1 – Occurrence of PCR artifacts during amplification of a group of closely related sequences from the PR-10 gene family. The number of sequences with recombination artifacts, mis-incorporated bases and both types of artifacts is indicated.

Species	Number of clones examined	Artifacts			
		Recombination	Base Mis-incorporation	Both	**Total**
Study 1 (Chapter 4)					
B. pendula Schneverdinger Goldbirke	120	35 (29%)	24 (20%)	0	**59 (49%)**
B. pendula Tristis	108	24 (22%)	13 (12%)	2 (2%)	**39 (36%)**
B. pendula Long Trunk	175	26 (15%)	19 (11%)	1 (1%)	**46 (26%)**
Average		**22%**	**14%**	**1%**	**37%**
Study 2 (Chapter 6)					
B. chichibuensis	170	18 (11%)	42 (25%)	5 (3%)	**65 (38%)**
B. costata	108	2 (2%)	31 (29%)	1 (1%)	**34 (31%)**
B lenta	106	9 (8%)	35 (33%)	4 (4%)	**48 (45%)**
B. nigra	155	5 (3%)	39 (25%)	0	**44 (28%)**
B. pendula Youngii	102	6 (6%)	24 (24%)	3 (3%)	**33 (32%)**
B. platyphylla	103	5 (5%)	35 (34%)	1 (1%)	**41 (40%)**
B. populifolia	106	2 (2%)	34 (32%)	0	**36 (34%)**
B schmidtii	184	10 (5%)	46 (25%)	6 (3%)	**62 (34%)**
Average		**5%**	**28%**	**2%**	**35%**

Correct sequences are especially important when recombinant allergens are developed for application in allergy diagnostics or Specific ImmunoTherapy (SIT). The Bet v 2 allergen is an example in which a sequencing error might have been incorporated into a recombinant protein. The double band at approximately 14 kDa from the SDS-PAGE gels depicted in Fig. 6.3 and 7.1 was cut out and analyzed by Q-TOF LC-MS/MS. Four different peptide fragments were identified that matched to Bet v 2 (Fig. 8.2c). One of these fragments represented a new variant, while the other tree fragments matched exactly to peptides that were predicted from the two GenBank sequences of Bet v 2 (DQ65633 and M65179). The predicted full length protein sequences of DQ65633 and M65179 differ by a block of five amino acids (Fig. 8.2a). Interestingly, the nucleotide sequences differ by only two nucleotides. The first one causes a frame shift in M65179, while the second one shifts the frame back to its original position (Fig. 8.2b). While M65179 is used to produce recombinant Bet v 2 (rBet v 2), the most likely original configuration is found in DQ650633, because this protein sequence matched to the sequences in *Corylus avellana*, which is a closely related species. Whether the shift in M65179 occurred *in vivo* or during the PCR is not deducible. The protein fragment that contains the region under discussion is not detected with Q-TOF, because it is too large when trypsin is used as a protease. Other proteases would be required to

determine which of the isoforms is expressed. It has been noted that natural Bet v 2 (nBet v 2) has a significantly higher IgE binding capacity than rBet v 2 (Van Ree *et al.* 1999). This could either relate to the existence of multiple Bet v 2 isoforms in nBet v 2, to differences in conformation between nBet v 2 and rBet v 2, but it could, thus, also be the result of selecting an incorrect sequence for rBet v 2.

(a.)

```
DQ650633 Betula pendula   MSWQTYVDEHLMCDIDGQGQQLAASAIVGHDGSVWAQSSSFPQFKPQEIT [ 50]
M65179   Betula pendula   ..................ASNSL.......................... [ 50]
DQ663552 Corylus avellana ..........................................L..E... [ 50]
DQ663543 Corylus avellana ....A.....................................L..E... [ 50]

DQ650633 Betula pendula   GIMKDFEEPGHLAPTGLHLGGIKYMVIQGEAGAVIRGKKGSGGITIKKTG [100]
M65179   Betula pendula   ................................................. [100]
DQ663552 Corylus avellana ......D.............T........................... [100]
DQ663543 Corylus avellana ......D.............T........................... [100]

DQ650633 Betula pendula   QALVFGIYEEPVTPGQCNMVVERLGDYLIDQGL                [134]
M65179   Betula pendula   ..............................                  [134]
DQ663552 Corylus avellana ..............................                  [134]
DQ663543 Corylus avellana ............................LE....               [134]
```

(b.)

```
DQ650633 Betula pendula   [36] GTGCGATATCGACGGGCAAGGCCAGCAACTCGCTGCATCTGCGATCGTCG [85]
M65179   Betula pendula   [36] ..................CCAGCAACTCGCTG.............. [85]
DQ663552 Corylus avellana [36] ..............C..G.........................T.... [85]
DQ663543 Corylus avellana [36] ..............C..G.........................T.... [85]
```

(c.)

```
DQ650633 Betula pendula    MSWQTYVDEHLMCDIDGQGQQLAASAIVGHDGSVWAQSSSFPQFK     [ 43]

DQ650633 Betula pendula    PQEITGIMK DFEEPGHLA-PTGLHLGGIK YMVIQGEAGAVIR GK K [ 99]
Betula pendula Youngii (1)           .........-.......... ............
Betula pendula Youngii (2)           ....GP...G..........
Betula lenta                         .........-.......... ............

DQ650633 Betula pendula    GSGGITIK K TGQALVFGIYEEPVTPGQCNMVVER LGDYLIDQGL   [134]
Betula pendula Youngii (1)                                    ..........
```

Figure 8.2 – Alignment of sequences from profilin. (a.) protein sequences from B. pendula *and* C. avellana *and their respective GenBank accession numbers. (b.) DNA sequence of* B. pendula *and* C. avellana *at the location of the potential sequence artifact. (c.) Theoretical tryptic peptides of birch profilin (Bet v 2). Tthe tryptic peptides that were sequenced from* B. pendula *'Youngii' and* B. lenta *using Q-TOF LC-MS/MS are underlined.*

Chapter 6 describes the application of Q-TOF LC-MS^E to analyze tryptic digests of PR-10 proteins in birch pollen. The tryptic peptide sequences can be used to evaluate two issues as far as the pollen-expressed genes are concerned, namely (1) whether the applied criteria for exclusion of suspected artifacts are correct and (2) whether any *B. pendula* sequences are missing from our data set. In

both cases, the tryptic digest would include peptides that originate from Bet v 1 isoforms that were not present in our search database. Expression of excluded artifact sequences could be evaluated for a limited set of sequences only, because not every mis-incorporated base leads to an amino acid substituation, and because protein sequence coverage by Q-TOF LC-MS^E is 71 to 79% for the PR-10 isoforms in *Betula* species (Chapter 6). All Bet v 1 accessions that have been deposited to GenBank/EMBL/DDJB by previous studies are from *B. pendula* or *B. platyphylla*. These GenBank sequences were included in the search database. No additional peptides were matched to any of the GenBank isoforms, although several GenBank sequences encode isoforms that contain unique tryptic peptides (Chapter 6). This indicates that the applied exclusion criteria were valid, and that there are no wide-ranging gaps in our data set.

PR-10 Gene family and Bet v 1 expression

Previous studies on Bet v 1 were primarily based on mRNA from pollen of *B. pendula* (Breiteneder *et al.* 1989; Friedl-Hajek *et al.* 1999; Swoboda *et al.* 1995b). Other PR-10 alleles were identified in roots and leaves (Koistinen *et al.* 2002; Poupard *et al.* 2001; Poupard *et al.* 1998). PR-10 isoforms that are expressed in vegetative tissues are not involved in hay fever, but information on these isoforms does contribute to understanding the variation within the PR-10 gene family. The research described in Chapter 4 used gDNA to amplify PR-10 sequences to recover as many PR-10 genes as possible, rather than focusing specifically on the pollen-expressed genes. An advantage is that the gDNA sequences include an intron, which is much more variable than the exon and which was previously shown to be useful to distinguish between different Mal d 1 genes (Gao *et al.* 2005a). Multiple PR-10 sequences were recovered from each *B. pendula* cultivar and these were attributed to at least thirteen different genes based on two criteria (Chapter 4 and 6): (1) each gene was allowed to have 1-2 allelic variants, since *B. pendula* is a diploid species, and (2) a cut-off value of 98.5% similarity was used to discriminate between different genes. Intron size and sequence were additional characters to distinguish between different genes in our study as well (Chapter 4). The two predefined criteria to discriminate different genes go together in the majority of cases, but Chapter 6 describes some exceptions in which three or more sequences from a single tree had more than 98.5% similarity. Therefore, the estimate of 13 genes for *B. pendula* should be considered as a conservative estimate and the real number of genes is potentially higher.

Amplification from gDNA provides no direct evidence for expression of PR-10 genes in pollen. Homology to known mRNA sequences was initially used as a criterion to predict which genes were pollen-expressed. Six out of the seven genes in subfamily 01 and 02 had homologues in the GenBank that were derived from pollen mRNA sequences. The genes in subfamily 01 and 02 were therefore considered to be pollen-expressed (Chapter 4). Three out of the five genes from subfamily 03 or 05 had homologues in the GenBank that were derived from mRNA sequences from roots and leaves. The single gene in subfamily 04 had no mRNA derived sequence in the

GenBank database and its transcription remains unknown. Deducing transcription entirely from homology to known mRNAs is rather speculative, especially when investigating other *Betula* species, for which no mRNA sequences are known. Therefore, Q-TOF LC-MS^E was used to detect PR-10 proteins in the pollen proteome of several *Betula* species in Chapter 6 and 7. We confirmed that only genes of subfamily 01 and 02 were expressed in pollen. However, not all genes from these subfamilies were expressed. For example, the *PR-10.01D* gene was sequenced from *B. pendula* 'Tristis', but was neither expressed in the two 'Tristis' trees, nor in the two other *B. pendula* cultivars that were examined (Chapter 7).

The PR-10 gene family is present as a multigene family in many species. Members of the PR-10 family in *B. pendula* can be subdivided into seven different subfamilies, five of which were common to all birch species. The remaining two subfamilies were each restricted to two out of the eight examined species, and mainly composed of sequences that have an interrupted open reading frame and are considered pseudogenes. An organization of *PR-10* genes into subfamilies was also reported for *Malus domestica* (Gao *et al.* 2005a) and for *Pinus monticola* (Liu *et al.* 2004). Differences in expression appear to coincide with the division between subfamilies. Subfamilies 01 and 02 were shown to be pollen-expressed (Chapter 6 and 7), while the genes in subfamily 03 and 05 are expressed in roots and leaves (Koistinen *et al.* 2002). Transcription of the genes in subfamily 03 and 05 is induced by different factors. The single gene in subfamily 05 has high basal transcription levels, while genes in subfamily 03 become transcriptionally upregulated upon infection of the leaves or by wounding (Poupard *et al.* 2001; Poupard *et al.* 1998; Swoboda *et al.* 1995c). The fact that the PR-10 gene family in *B. pendula* has multiple members raises the question as to why duplicates are maintained. The classical model on gene duplication predicts that null allele formation should occur in case of redundancies. Preservation of duplicates is explained by assuming that the duplicates have acquired novel functions (Ohno 1970). In contrast, the duplication-degeneration-complementation model predicts that subfunctionalization is responsible for preservation of duplicates (Force *et al.* 1999). The subfunctionalization mechanism, in which duplicates are expressed under different conditions or in different tissues, allows the occurrence of degenerative mutations and may therefore be more common than the acquirement of novel functions, which requires rare beneficial mutations (Force *et al.* 1999). The fact that *PR-10* genes in *B. pendula* differ in the location of expression and the conditions that induce transcription, is indicative for subfunctionalization.

If subfunctionalization has occurred in the *Betula* species, one would expect that it has also occurred in related species that contain multiple PR-10 genes. Sequences and some EST sequences from several Fagales species have been deposited to the GenBank. A Neighbor Joining (NJ) tree was constructed which included these sequences and a representative sequence for each of the *B. pendula* genes (Fig. 8.3). The sequences from *Coryllus avellana* and *Carpinus betulus* did show a

division into respectively three and four subfamilies. Similarities to the clustering of the *B. pendula* genes were observed. All known pollen expressed genes were located within the large cluster at the top of Fig. 8.3. However, this cluster also included *Betula* genes of subfamily 03, which are transcribed in roots and leaves (Koistinen *et al.* 2002). The seed-expressed mRNAs from hazel, which are relevant with regard to food allergy to hazelnut (Lüttkopf *et al.* 2002), clustered with subfamily 05 from birch. Only one PR-10 sequence was present for the species *Castanea sativa*, *Fagus sylvatica*, *Casuarina glauca* and *Alnus glutinosa*, while the EST sequences from the *Juglans* hybrid are also placed in a single cluster. Overall, this analysis does indicate that subfunctionalization is present among other Fagales species as well, but it will require a more extensive sampling of Fagales species to determine similarities in the pattern of subfunctionalization among different species.

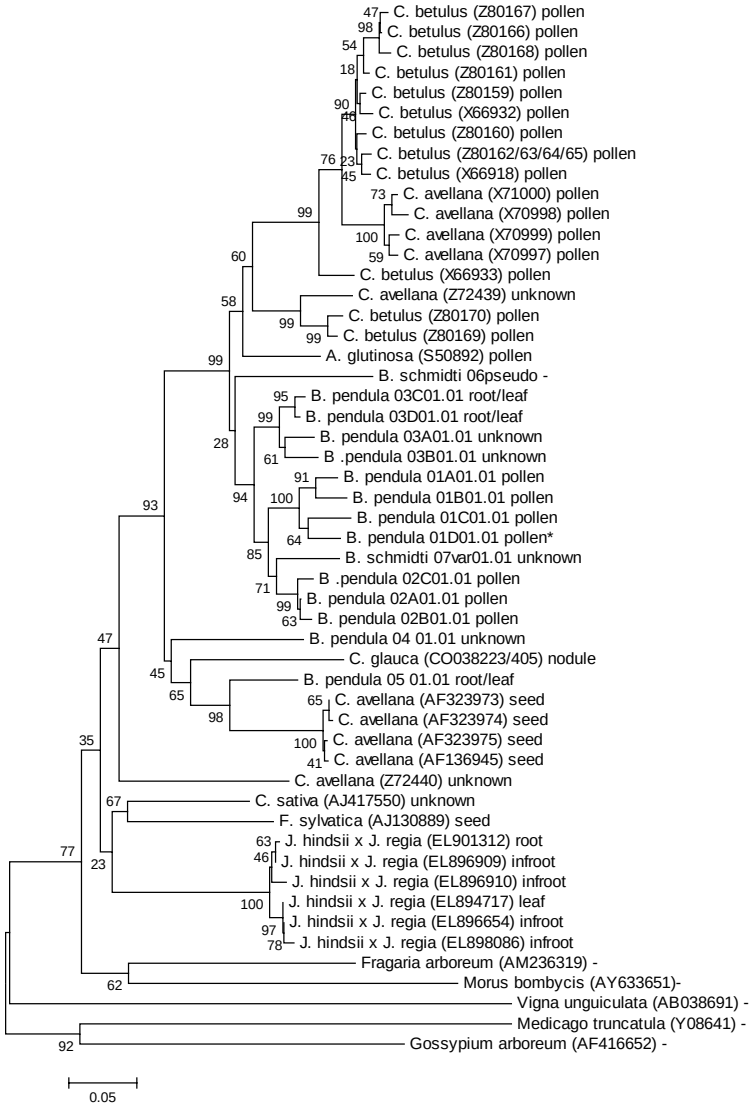


Figure 8.3 – Clustering of PR-10 sequences from Fagales species in a Neighbor Joining tree with Kimura two-parameter distances. Bootstrap percentages are shown on the branches (1,000 replicates). The location of transcription is indicated for sequences derived from mRNAs or ESTs. The included B. pendula *sequences are described in Chapter 6. Additional sequences were recovered from GenBank (Acc. No) for: European hornbeam (*Carpinus betulus*), European hazel (*Coryllus avellana*), alder (*Alnus glutinosa*), swamp oak (*Casuarina glauca*), European chestnut (*Castanea sativa*), beech (*Fagus sylvatica*) and a walnut hybrid (*Juglans hindsii x Juglans regia*). Sequences of* Fragaria arboreum*,* Vigna unguiculata*,* Morus bombycis*,* Medicago trunculata *and* Gossypium arboretum *were used as outgroup.*

Bet v 1 proteins are very abundant in birch pollen (Larsen 1995; Chapter 7). Given this abundance, it is striking that the exact function of Bet v 1 has not been established. Bet v 1 has initially been reported to act as ribonuclease (Swoboda *et al.* 1996). A more recent study suggested that the cavity inside the Bet v 1 protein is related to its function as plant steroid carrier (Markovic-Housley *et al.* 2003). Bet v 1 is reported to interact with various physiological ligands (Mogensen *et al.* 2002), but which ligands are bound inside the cavity *in vivo* is unclear. Additional suggestions for the function of PR-10 proteins in general are that they act as cytokine-binding proteins (Fujimoto *et al.* 1998), as storage proteins (Flores *et al.* 2002; Richard-Molard *et al.* 2004) or as cryoprotective proteins (Ukaji *et al.* 2004). To what extent all the suggested functions of the PR-10 proteins are compatible is also unclear, and elucidating their function is an interesting target for further studies.

A mixture of allergens in each tree: consequences for sensitization and cross-reactivity

Ferreira *et al*. (1996) divided the Bet v 1 isoforms from *B. pendula* into three groups according to their allergenicity, and confirmed the division between high, moderate, or low IgE-reactivity in a Skin Prick Test (SPT). Swoboda *et al*. (1995b) showed that pollen samples that were combined from multiple trees contain multiple Bet v 1 isoforms. We found that the seven genes that are putatively expressed in the pollen of a single *B. pendula* tree encode isoforms with different IgE-reactivity (Chapter 4). The expression of these genes was confirmed by the study in Chapter 6, which subsequently showed that other *Betula* species contain a similar mixture of isoforms. Q-TOF LC-MS^E proved to be a fast and accurate screening method to determine the presence of particular isoforms or tryptic peptides, provided that a good search database is available. In the studies described in Chapter 6 and 7, the database was formed by genomic sequences obtained from the same trees from which pollen were collected. The species investigated in Chapter 6 were selected from different subgenera in the *Betula* genus (Chapter 5) and are considered representative for the variation in this genus. All examined species contained a mixture of isoforms. Additional birch species were screened in chapter 7, all of which also contained multiple isoforms. This leads to the hypothesis that the occurrence of Bet v 1 isoform mixtures will be common to all *Betula* species, and possibly to other Fagales species with PR-10 pollen proteins.

Many birch species and birch hybrids have higher ploidy levels (tetraploid, hexaploid, octaploid and even decaploid) than the diploid *B. pendula* (Chapter 5) and will consequently contain an increased number of allergen isoforms. The *B. pendula* cultivar 'Long Trunk' as examined in Chapter 3 was, in fact, a tetraploid hybrid between *B. pendula* and another birch species. Indeed, 28 unique sequences were recovered in 'Long Trunk' opposed to 14-16 unique sequences in the diploid *B. pendula* cultivars. We also examined the tetraploid species *B. ermanii* from which 25 unique sequences were recovered (data not shown). *B. ermanii* occupies an intermediate phylogenetic position between *B. pendula* and *B. costata* (20 unique sequences; see Chapter 6).

The increased number of unique sequences in trees with higher ploidy levels may lead to an increased number of isoforms, if all of them are expressed. The extent to which this leads to an increased variation among isoforms in one tree is somewhat lower, because the parental species may often share isoforms that are similar to each other. For example, *B. medwediewii* is a decaploid (10n) species and therefore theoretically harbors approximately 5 times as many PR-10 genes as *B. pendula*. However, the difference in the number of peptide variants between both species was much smaller. Whereas *B. pendula* 'Youngii contained 21 different peptide variants in total (Table 6.3), *B. medwediewii* contained 28 different peptide variants, which was only a slight increase.

The abundance of particular isoforms will affect the overall allergenicity of the pollen mixture from a single tree. *B. pendula* is known to be highly allergenic, despite the presence of a mixture of isoforms with varying IgE-reactivity. The abundant presence of Bet v 1a (=PR-10.01A01), which is known to have a high IgE-reactivity, may already explain the high allergenicity of its pollen. Indeed, 01A01 was estimated to encompass ~40% of the total amount of Bet v 1 (Chapter 6) and thereby was the most abundant isoform in the mixture. Given that the Bet v 1 band was by far the most intense band in the SDS-PAGE; 01A01 was also the most abundant protein in the pollen grain. *B. pendula* isoforms that have an intermediate IgE-reactivity (02C01 plus 01C04), and an isoform that has a low IgE-reactivity (01B01) each comprise ~20% of the mixture, which is still a substantial level of expression. When sensitization takes place, individuals that are exposed to *B. pendula* pollen automatically come into contact with a mixture of isoforms. It appears that during the sensitization process, some isoforms develop into isoforms with a high IgE-reactivity, while other isoforms develop into isoforms with an intermediate or low IgE-reactivity (Ferreira *et al.* 1996; Wagner *et al.* in press). The factors that cause this difference in sensitizing power are of interest for future research. Since all three types of isoforms are expressed at relatively high levels, it is unlikely that differences in sensitization can be fully attributed to the amount of each isoform type. Mogensen *et al.* (2007) show that a specific isoform of the *PR-10.01A* gene is able to bind and permeabilize human membranes. Perhaps this ability varies between different Bet v 1 isoforms, causing a difference in the ability to sensitize people.

The involvement of multiple Bet v 1 isoforms has to be considered with regard to application in allergy diagnostics. Natural mixtures of allergens are difficult to standardize, in contrast to recombinant allergens. This is especially so when allergens are unstable, which is the case for several food allergens (Bohle *et al.* 2004). The use of recombinant allergens also enables to determine the sensitization profile of individual patients on the level of individual allergens rather than on food products as a whole. For diagnostic purposes, recombinant isoforms should be recognized by as many patients as possible. rBet v 1a seems an excellent candidate for diagnosis of birch pollen allergy (Pauli *et al.* 1996; Valenta *et al.* 1998; Van Ree *et al.* 1999). This isoform has a

high IgE-reactivity. Bet v 1a was found in high quantities in all examined *B. pendula* cultivars (Chapter 6 and 7). As a consequence, every individual sensitized against *B. pendula* is likely to react to Bet v 1a. *B. populifolia* (a North-American species) also contains the Bet v 1a allele (Chapter 6), while *B. plathyphylla* and *B. costata* (Asian species) express isoforms that are highly similar and will contain similar epitopes (Chapter 6 and 7), suggesting that when sensitization takes place against the latter species, patients will react to Bet v 1a as well. Other abundant species on which no sequencing data were acquired, such as *B. pubescens* (Europe), *B. papyrifera* (North-America) and *B. ermanii* (Asia) take an intermediate phylogenetic position between *B. pendula* and *B. costata* (Chapter 5) and can be expected to contain similar isoforms as well.

The involvement of multiple allergen isoforms has potential consequences with regard to the application of SIT. The use of natural mixtures of allergens may induce IgE-mediated side effects. Therefore, current investigations are focusing on the use of recombinant allergens with a low IgE-reactivity (Wagner *et al.* in press), or recombinant allergen fragments (Gafvelin *et al.* 2005). The use of recombinant allergens allows modification of allergens to reduce the risk of IgE-mediated side-effects (Valenta *et al.* 2007). Small peptides harboring the major T-cell epitopes are used to inhibit IgE responses in allergen peptide-based SIT. Such peptides can even be incorporated into a rice-based edible vaccine (Takagi *et al.* 2005). However, the use of recombinant allergen solutions or small peptides in SIT introduces the issue of choice: if one selects one specific isoform, there is a risk for selecting isoforms that lack particular T-cell epitopes. However, all *B. pendula* cultivars and *Betula* species as examined in this thesis (Chapter 7) contained a highly similar Bet v 1 mixture. Individual patients that have been exposed to pollen from different trees or even species may still display sensitization to the same Bet v 1 allergens. This increases the chance for identifying SIT applications that are universally applicable, provided SIT includes all relevant epitopes.

Cross-reactivity and OAS

Sensitization to birch pollen is accompanied by cross-reactions or co-sensitization to pollen of other Fagales species (Corsico *et al.* 2000; Gioulekas *et al.* 2004; Mari *et al.* 2003; Von Mutius *et al.* 1998). The term cross-reactivity is applied if the source of the sensitization can be identified. If not, the process is referred to as co-sensitization (Ferreira *et al.* 2004). Bet v 1 homologues are transcribed in the pollen of several Fagales species (Fig. 8.3). The high similarity at the amino acid level between Bet v 1 and these homologues provides a molecular basis for cross-reactivity (Niederberger *et al.* 1998). Alder species have a distribution that is very similar to that of species of the genus *Betula*, and *A. glutinosa* is a very common species throughout large parts of Europe. Pollen counts for *Alnus* sp. are on average about 5-fold lower than counts for birch in for example Poland and Belgium (Detandt *et al.* 2000; Weryszko-Chmielewska *et al.* 2001), which are still considerable amounts. Pollen of *Corylus* is much less abundant, but relevant with regard to hazel nut consumption. Both *Alnus* and *Corylus* flower during the winter and may prime allergic

reactions to birch (Crimi *et al.* 1990; Rak *et al.* 1992). Sensitization to Fagales species such as oak, beech and chestnut occurs (Mari *et al.* 2003), but appears more rare than sensitization to the above mentioned species as the overall sensitization to Fagales species is much lower in southern European countries, even though there is a large number of oak, beech and chestnut trees in this region.

Exposure to birch pollen may prime the reactions that result from the cross-reactivity between Bet v 1 and its homologues in plant foods. The occurrence of clinical reactions to apple increased during the birch pollen season in birch allergic individuals (Skamstrup Hansen *et al.* 2001). Reactivity against Bet v 1 homologues follows a similar geographical pattern as reactivity against birch. For example, reactivity to Mal d 1 is prevalent in north and central Europe where birch trees are common, while being less prevalent in South Europe (Fernández-Rivas *et al.* 2006). It is not clear whether all Bet v 1 isoforms are relevant with regard to cross-reactivity. One limiting factor for the identification of hypoallergenic plants is that the IgE-reactivity of many isoforms is not known. For example, over 50 unique isoforms from various birch species have been identified in this thesis (Chapter 6), but knowledge on the IgE-reactivity is limited to nine isoforms (Ferreira *et al.* 1996; Wagner *et al.* in press) and T-cell activation has only been tested thoroughly for a single isoform (Jahn-Schmid *et al.* 2005). Differences in reactivity between isoforms have led to the identification of a few amino acid residues that affect IgE-reactivity (Ferreira *et al.* 1998; Neudecker *et al.* 2003; Spangfort *et al.* 2003). However, this information is based entirely on the few isoforms that have been tested as recombinant proteins. More residues are likely to be identified when other isoforms would be tested by expressing them as recombinant proteins and using these in a SPT or T-cell activation test.

It has been suggested that cross-reactivity between allergens from different plant species reflects taxonomic distances between species (Weber 2003). A closer taxonomic relation would result in higher similarities in amino acid composition between proteins from plant species. Indeed, the PR-10 proteins in the highly cross-reactive apple are relatively more similar to Bet v 1 than those in soy or carrot for which cross-reactivity is rarer (Table 8.2). However, this does not explain observations by Van Ree *et al.* (2000), showing that reactivity of Bet v 1 sensitized patients decreases significantly from apple to cherry to peach to pear, while the overall similarity of Bet v 1a (X15877) to isoforms from these fruits is in the same order of magnitude (57-60%). PR-10 genes are transcribed in many plant foods besides those for which cross-reactivity by Bet v 1 homologues has been demonstrated. Table 8.2 lists all plant foods in which ESTs or mRNAs of PR-10 genes have been recovered from edible parts (GenBank search: September 2007). However, not all of the foods listed in Table 8.2 are known to cause OAS. One explanation would be that these allergies have not been described at the present time, but this seems unlikely for fruits such as grape en *Citrus* sp., which are consumed on a large scale. Possible explanations are that the PR-

10 proteins are expressed at very low levels in these fruits, or that cross-reactive epitopes are absent on the expressed isoforms due to absence of particular amino acids. In the latter case, differences in allergenicity are caused by the absence of similarity in the epitope region is rather than the overall similarity. Alternatively, PR-10 proteins in these fruits may have a lack of surface conservation, leading to changes in conformational or electrostatic properties (Ghosh *et al.* 2008), or the food matrix may play a role in the interaction with these PR-10 allergens.

TABLE 8.2 – PR-10 sequences transcribed in fruits, vegetables, grains, nuts, legumes and seeds sorted by the overall amino-acid similarity to Bet v 1a (X15877)

Source	EST or mRNA from edible parts	Described as allergen (GenBank Acc. No.)	Similarity to Bet v1a
Beech nut (*Fagus sylvatica*)	AJ130889	-	65 %
Apricot (*Prunus armeniaca*)	U93165	Pru ar 1 (U93165)	62 %
Cherry (*Prunus avium*)	U66076	Pru av 1 (U66076)	60 %
Peach (*Prunus persica*)	DQ251187	Pru p 1 (DQ251187)	60 %
Apple (*Malus domestica*)	X83672	Mal d 1 (X83672)	59 %
Grape (*Vitis vinifera*)	EC958497	-	58 %
Cluster bean (*Cyamopsis tetragonoloba*)	EG987166	-	58 %
Pear (*Pyrus communis*)	AF057030	Pyr c 1 (AF057030)	57 %
Hazelnut (*Corylus avellana*)	AF136945	Cor a 1.04 (AF136945)	57 %
Scarlet runner bean (*Phaseolus coccineus*)	CV541024	-	56 %
Sweet potato (*Ipomoea batatas*)	CO499980	-	55 %
Tomato (*Solanum lycopersicum*)	AK224718	-	53 %
Mandarin (*Citrus reticulata*)	DY260838	-	53 %
Valencia orange (*Citrus sinensis*)	EH406480	-	53 %
Sesame (*Sesamms indicum*)	BU668133	-	51 %
Pepper (*Capsicum annuum*)	CO912259	-	49 %
Cassava (*Manihot esculenta*)	DV457043	-	49 %
Soybean (*Glycine max*)	-	Gly m 4 (X60043)	48 %
Parsley (*Petroselinum crispum*)	X98688	-	45 %
Mungbean (*Vigna adiate*)	AY792956	Vig r 1 (AY792956)	45 %
Cowpea (*Vigna unguiculata*)	CK151444	-	45 %
Bread Wheat (*Triticum aestivum*)	CJ856226	-	43 %
Barley (*Hordeum vulgare*)	BI778636	-	43 %
Peanut (*Arachis hypogaea*)	AY328088	Ara h 8 (AY328088)	42 %
Oca (*Oxalis tuberosa*)	AF333436	-	41 %
Celery (*Apium graveolens*)	Z48967	Api g 1 (Z48967)	40 %
Asparagus (*Asparagus officinalis*)	AJ132612	-	39 %
Sorghum (*Sorghum bicolor*)	U60764	-	39 %
Carrot (*Daucus carota*)	Z81361	Dau c 1 (Z81361)	37 %
Onion (*Allium cepa*)	CF452005	-	37 %
Yam-bean (*Pachyrhizus erosus*)	AY433943	-	36 %

Breeding perspectives towards hypoallergenic birch trees

Birch trees planted in cities are generally selected and subsequently clonally propagated to maintain desirable characteristics. Although birch trees are gradually becoming less popular as lane trees, they are still planted as ornamental trees in parks and private gardens. Even outside their natural distribution range, birch has, for example, been introduced into urban areas in Australia and New-Zealand as an ornamental tree (Spellerberg *et al.* 2006). Characteristics that make birch popular are the colored bark and graceful shape of the tree. Disadvantages of birch are the abundant seed production and the superficially anchoring root system. In recent years, nature in cities has received more attention, and seed and pollen production are considered a source of food for insects or birds that live in cities. However, planting of birch trees is currently opposed by people who are suffering from birch pollen allergy, and civic authorities often receive requests to cut down birch trees. Whether this really helps patients suffering from birch pollen allergy is an area of interest for future research, and likely to be dependent on the extent to which local pollen (Clot 2001; Spieksma *et al.* 2003) and pollen distributed by long-distance transport (Koivikko *et al.* 1986; Skjøth *et al.* 2007; Sofiev *et al.* 2006) contribute to the allergic reactions. A second issue regarding exposure is the increased planting of exotic birch species as ornamental trees. Originally, only *B. pendula* and *B. pubescens* occurred in the Netherlands. Especially several *B. pendula* cultivars and *B. utilis* are often used as ornamental trees, but species such as *B. ermanii*, *B. pubescens*, *B. albosinensis*, *B. maximowicziana*, *B. nana* and *B. nigra* are also planted. These species differ in their flowering time and period (Fig. 8.4). When different species with different flowering times are planted in the same area, the local duration of the hay fever season may increase as a consequence.

Birch trees that have a reduced allergenicity may be suitable for planting within the urban environment. A reduced allergenicity of birch trees can be based on a reduced quantity of allergens or on the absence of allergen isoforms with a high IgE-reactivity. Variation in the amount of allergens has been detected in fruits and pollen of allergenic plants such as apple, peach and nectarine, peanut and olive (Ahrazem *et al.* 2007; Castro *et al.* 2003; Koppelman *et al.* 2001; Marzban *et al.* 2005). However, birch trees that produce pollen with substantial reduction levels of Bet v 1 are unlikely to be identified, given the high amount of Bet v 1-type proteins in pollen of each of the examined birch species (Chapter 6 and 7). Variation in the allergenicity of isoforms is known to exist for Bet v 1 and Api g 1 (Ferreira *et al.* 1996; Wangorsch *et al.* 2007). The isoform Bet v 1a (=PR-10.01A01) is known to have a high IgE-reactivity. Bet v 1a or isoforms that differed by only 1-3 amino acids were detected in all examined *Betula* species except *B. lenta* (Chapter 6). A high similarity between isoforms increases the chance that they share epitopes. These isoforms were also abundant in all these species, while isoforms that were similar to isoforms with a low IgE-reactivity were rare or absent in all species (Chapter 6). Breeding of birch trees that only contain isoforms with a low IgE-reactivity is complicated when the Bet v 1 genes are present in a

tandem repeat and linked. This has been found for the PR-10 genes in apple (Gao *et al.* 2005a). Mapping of the PR-10 genes in birch is currently not possible, given the low density of genetic markers on the linkage map of *B. pendula* (Pekkinen *et al.* 2005).

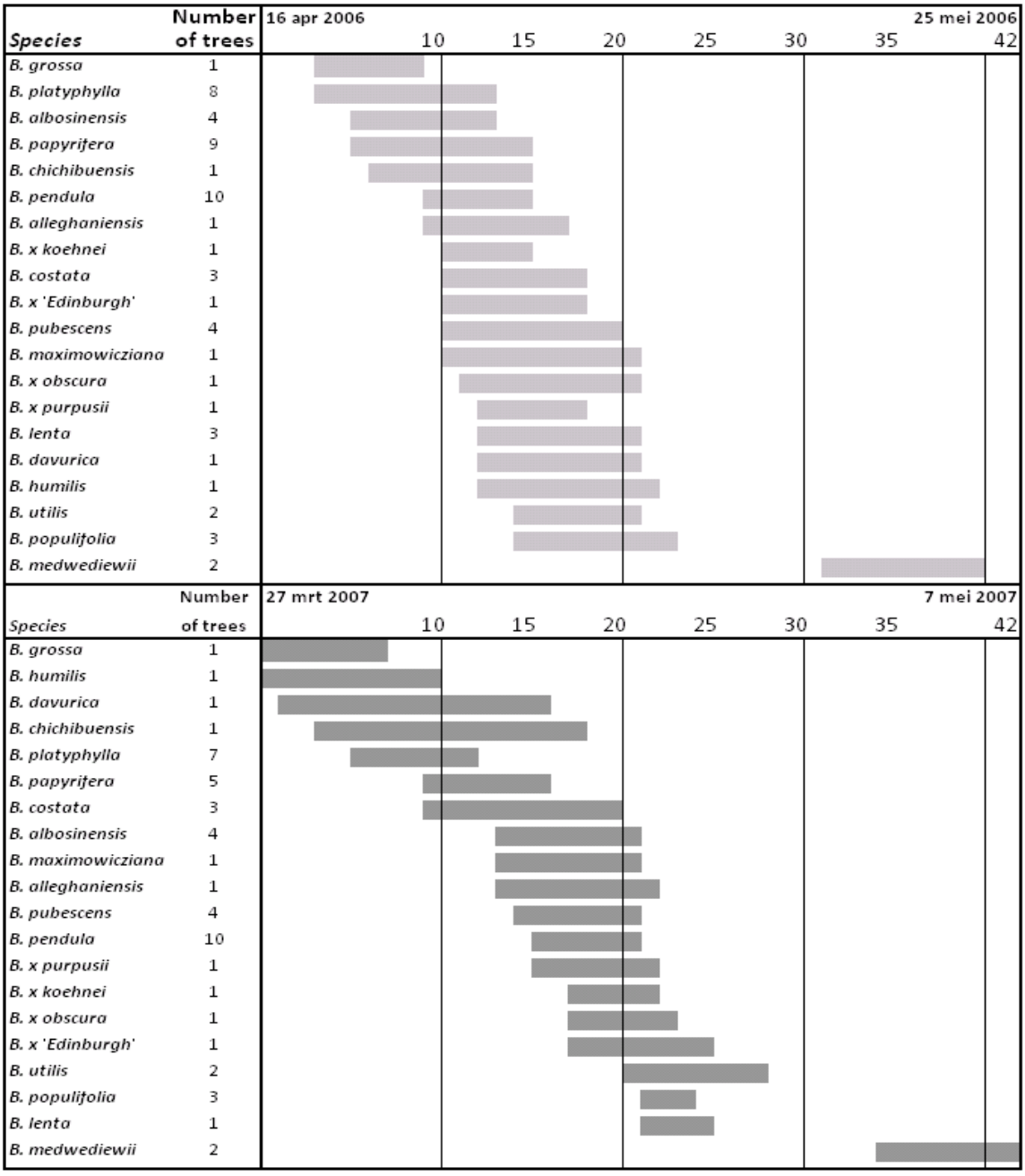


Figure 8.4 – Flowering times of different birch species at the Botanical Garden of Wageningen (Wageningen, the Netherlands) observed in 2006 and 2007. The number of trees examined for each species, average start day and end day of the flowering are indicated, as is the average length of the flowering period for each species.

When tested in a Western blot with serum from patients that were allergic to birch, all birch species displayed considerable IgE-reactivity, including *B. lenta* (Chapter 7). The experiments showed a clear correlation between the IgE-reactivity in the immunoblot and the total amount of Bet v 1, regardless of which isoforms were present in the pollen mixture. Thus, development of hypoallergenic birch trees will require other strategies than selection of trees that produce pollen with a reduced allergenicity. Feasible alternatives include the search for completely sterile or male sterile trees, which is a naturally occurring phenomenon in many plant species. In particular, hybrids between less related species often display a reduced fertility, which may result in a reduced or aborted pollen production. However, hybridization and introgression are common among *Betula* species (Nagamitsu *et al.* 2006; Palme *et al.* 2004) as can also be judged by the number of hybrids identified and included in Chapter 5. The high frequency of polyploidization may point at the occurrence of unreduced (2*n*) gametes, a potentially major mechanism of polyploid production (Jauhar 2003; Ramanna *et al.* 2003). Sterility can be introduced in birch by genetic modification (Lemmetyinen *et al.* 2004), but is perhaps already present in the natural variation. One birch tree that was monitored did not produce any male catkins for three consecutive years, while another tree produced male catkins that contained no pollen (personal observation). Another option would be the application of RNA interference (RNAi) technology, which proved to be successful to silence Mal d 1 allergens in apple (Gilissen *et al.* 2005). Previously, transgene-induced silencing had been used to silence the allergens Gly m Bd 30 in soybean (Herman *et al.* 2003) and Lol p 5 in the pollen of ryegrass (Bhalla *et al.* 1999; Petrovska *et al.* 2005). More recent approaches have used RNAi to silence the allergens Lyc e 1 in tomato (Le *et al.* 2006) and Ara h 2 in peanut (Dodo *et al.* 2007).

Food allergy management

The cultivar (cv) Santana was identified as a hypoallergenic apple (Bolhaar *et al.* 2005a), which was confirmed in a food challenge study (Kootstra *et al.* 2008). The sales pilot study in Chapter 3 shows that almost half of the apple allergic consumers self-reported that they were able to eat the cv Santana without experiencing an allergic reaction. Among the consumers that still experienced an allergic reaction, several stated that their reaction was less severe than the reaction experienced following consumption of other apple cultivars. A correlation was observed between the occurrence of an allergic reaction to the cv Santana and suffering from a more severe apple allergy (Chapter 3). However, some consumers who self-reported suffering from a mild apple allergy also self-reported experiencing a considerable allergic reaction to the cv Santana, while some consumers with a severe apple allergy self-reported no allergic reaction to the cv Santana. These results rely on self-diagnosis, and the observed correlation may be stronger when food challenges are used to classify patients. These observations raise the following issues: why do hypoallergenic products still cause complaints in some consumers, what can be done to improve these products, and are hypoallergenic products suitable to be marketed as such.

Three factors may explain why hypoallergenic products still cause complaints in some consumers: (1.) Many foods contain multiple allergens. This is for example the case in apple, but can basically be observed in any allergenic food [http://www.allergen.org/]. These allergens may even induce different clinical response profiles (Pastorello *et al.* 1999; Wensing *et al.* 2002). Hypoallergenic foods may have a reduced allergenicity for one allergen, but not for the others, causing differentiation in the response among individuals that react to different allergens in the same food. (2.) Individual consumers have different thresholds for an allergic reaction. Allergen levels that are tolerated by one consumer may still trigger an allergic reaction in others (Taylor *et al.* 2002). When the reduced allergenicity of a hypoallergenic product is based on lower allergen content, the reduction may not suffice for all consumers. (3.) Consumers may react to different isoforms of the allergen. Isoforms of Bet v 1 and the celery allergen Api g 1 vary in allergenicity (Ferreira *et al.* 1997; Wangorsch *et al.* 2007). Protein variants which trigger an allergic reaction in one patient may not do so in others (Ferreira *et al.* 1997). When the reduced allergenicity of a hypoallergenic product is based on the absence of a particular isoform, patients that recognize other isoforms will still react.

Cv Santana was identified as hypoallergenic food by screening of existing cultivars. No attempts have been made to improve the hypoallergenic properties of this apple. Other apple cultivars with a reduced allergenicity were also identified in SPTs (Bolhaar *et al.* 2005a; Carnés *et al.* 2006), perhaps providing additional material for further breeding. A potential alternative would be the development of hypoallergenic foods like apple by application of RNAi (Dodo *et al.* 2007; Gilissen *et al.* 2005). One major advantage would be that hypoallergenic properties can be introduced into existing varieties of fruit or vegetable crops that are known to have a good taste, texture, shelf-life and economic viability. When multiple (major) allergens are found in the same food, this implies that simultaneous silencing of these allergens will be required. In addition, reduction levels of allergens should be efficient (Le *et al.* 2006) and stable under variable growing conditions. Proteins that fulfill essential functions will be difficult to silence. Potential solutions would then be tissue-specific expression of the RNAi transgene, by targeting a subgroup of proteins with a tissue-specific promoter, or the introduction of hypoallergenic protein variants to restore the physiological function.

Societal concern about GM has been a major factor in delaying the development of GM foods (Gaskell *et al.* 2006). Societal attitude towards GM is influenced by case-specific characteristics of the application, such as which organism is modified and to what purpose (Frewer *et al.* 1997; Zechendorf 1994). The majority of the European consumers are positive about medical applications of GM, whilst at the same time rejecting many agricultural applications. Acceptance of hypoallergenic GM apples is quite high among apple allergic consumers, because these

consumers experience a “personal benefit” associated with hypoallergenic GM products (Chapter 3). This suggests that consumers prioritize their personal needs and beliefs and accept applications of GM when the benefits are sufficiently large. Similar results were observed for acceptance of hypoallergenic non-food products, such as hypoallergenic GM birch or grass (Chapter 2). The research in Chapter 2 shows that acceptance of GM to reduce the allergenicity of plants that cause hay fever was relatively high in comparison to applications of GM in hypoallergenic foods. In either case, the results of the studies in Chapter 2 and 3 indicate a clear consumer preference for traditional breeding over breeding by GM for the development of hypoallergenic birch trees or food products. Product development approaches which focus on traditional breeding should be explored before exploring GM as has been done for birch in the Chapters 4 to 7.

Given the individual differences among apple allergic consumers in the response to cv Santana, a cautious approach regarding interventions based on hypoallergenic foods seems wise. Not every food that causes allergy is suitable for introduction of hypoallergenic counterparts. Interventions with hypoallergenic foods are not desirable when foods may cause a very severe reaction such as anaphylactic shock, because of the potential fatal consequences in case the hypoallergenic product is not effective. Allergenic foods that cause OAS can be considered suitable targets for the development of hypoallergenic food products. The Bet v 1 homologues that cause the OAS are too labile to induce severe systemic responses (Ortolani *et al.* 1988). If an allergic reaction occurs to these foods, the consequences are rather minor and tolerable. The consequences can further be minimized by using a step-to-step protocol as has been development and utilized for the cv Santana (Chapter 3). The hypoallergenic cv Santana was, generally, well received by apple allergic consumers and rated positively by all consumers except the few consumers that experienced a strong allergic reaction following consumption (Chapter 3). For allergic consumers who did not experience an allergic reaction following consumption of the hypoallergenic food, the introduction of hypoallergenic foods contributes to a normal and healthy diet as they can now relieve part of their dietary restrictions. This is not the case for consumers with minor allergic complaints to the hypoallergenic products. These consumers were, however, positive about hypoallergenic products in the case of the cv Santana (Chapter 3). Some consumers stated that they were positive about the cv Santana because attention was being paid to their allergy problem. Whether these consumers will continue to be positive about hypoallergenic products if future hypoallergenic products still do not alleviate their allergic complaints, is questionable. There was evidence of consumer demand for the development and introduction other hypoallergenic fruits, such as cherry, strawberry, peach and pear, although ambivalence existed about the method applied to their development.

Conclusion

The research described in this thesis focuses on two issues regarding birch pollen allergy. The first issue relates to the question whether hypoallergenic birch trees can be selected or developed from the existing natural variation. Sequence validation was necessary for accurate molecular characterization of the multigene Bet v 1 allergens in *B. pendula* and in other birch species, to exclude sequencing artifacts. Q-TOF LC-MS^E was used to confirm expression of the Bet v 1 allergens by detection of tryptic Bet v 1 peptides. Pollen from individual trees of all examined birch species contained a mixture of Bet v 1 isoforms. Isoforms that were similar to isoforms with a high IgE-reactivity were abundant in this mixture as was the total amount of Bet v 1. This reduced the possibilities for identifying hypoallergenic birch trees. Indeed, immunoblotting showed that all examined *Betula* species had a high allergenicity. Development of hypoallergenic birch trees thus has to rely on other approaches, such as selection or introduction of male sterility or the application of RNAi to silence the Bet v 1 genes.

The second issue deals with the development of hypoallergenic products in general. Societal issues need to be considered, particularly when GM is applied to develop such products. The survey-based research demonstrates a clear consumer preference for traditional plant breeding compared to plant breeding utilizing GM technology. None-the-less, acceptance of hypoallergenic GM products is quite high, particularly for allergic sufferers who are allergic to the product being modified. This supports the hypothesis that consumer attitude towards GM is partly driven by recognition of specific and personally relevant benefits. Finally, the introduction of the hypoallergenic apple cultivar Santana, which was not developed by GM, was well received by allergic consumers, also by those who still experienced mild allergic reactions to the product. This suggests that there is a need and a market to develop hypoallergenic products in the future.

REFERENCES

Abe Y, Kimura S, Kokubo T, Mizumoto K, Uehara M, Katagiri M (1997) Epitope analysis of birch pollen allergen in Japanese subjects. Journal of Clinical Immunology 17: 485-493

Aberg N, Hesselmar B, Aberg B, Eriksson B (1995) Increase of asthma, allergic rhinitis and eczema in Swedish schoolchildren between 1979 and 1991. Clinical and Experimental Allergy 25: 815-819

Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF (2005) PCR-induced sequence artifacts and bias: Insights from comparison of two 16s rRNA clone libraries constructed from the same sample. Applied and Environmental Microbiology 71: 8966-8969

Ahrazem O, Jimeno L, López-Torrejón G, Herrero M, Espada JL, Sánchez-Monge R, Duffort O, Barber D, Salcedo G (2007) Assessing allergen levels in peach and nectarine cultivars. Annals of Allergy, Asthma and Immunology 99: 42-47

Alm JS, Lilja G, Pershagen G, Scheynius A (1997) Early BCG vaccination and development of atopy. Lancet 350: 400-403

Alm JS, Swartz J, Lilja G, Scheynius A, Pershagen G (1999) Atopy in children of families with an anthroposophic lifestyle. Lancet 353: 1485-1488

Althoff DM, Gitzendanner MA, Segraves KA (2007) The utility of amplified fragment length polymorphisms in phylogenetics: a comparison of homology within and between genomes. Systematic Biology 56: 477-484

Álvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution 29: 417-434

Arens P, Coops H, Jansen J, Vosman B (1998) Molecular genetic analysis of black poplar (L.) along Dutch rivers. Molecular Ecology 7: 11-18

Asero R (1998) Effects of birch pollen-specific immunotherapy on apple allergy in birch pollen-hypersensitive patients. Clinical and Experimental Allergy 28: 1368-1373

Asero R, Mistrello G, Roncarolo D, Amato S (2007) Detection of some safe plant-derived foods for LTP-allergic patients. International Archives of Allergy and Immunology 144: 57-63

Atkinson MD (1992) *Betula pendula* Roth (*B. verrucosa* Ehrh.) and *B. pubescens* Ehrh. Journal of Ecology 80: 837-870

Ballmer-Weber BK (2002) Lipid transfer protein as a potential panallergen? Allergy 57: 873-875

Bantignies B, Séguin J, Muzac I, Dédaldéchamp F, Gulick P, Ibrahim R (2000) Direct evidence for ribonucleolytic activity of a PR-10-like protein from white lupin roots. Plant Molecular Biology 42: 871-881

Barbee RA, Brown WG, Kaltenborn W, Halonen M (1981) Allergen skin-test reactivity in a community population sample: Correlation with age, histamine skin reactions, and total serum immunoglobulin E. Journal of Allergy and Clinical Immunology 68: 15-19

Bhalla PL, Singh MB (2004) Knocking out expression of plant allergen genes. Methods 32: 340-345

Bhalla PL, Swoboda I, Singh MB (1999) Antisense-mediated silencing of a gene encoding a major ryegrass pollen allergen. Proceedings of the National Academy of Science of the USA 96: 11676-11680

Björksten F, Suoniemi I, Koski V (1980) Neonatal birch-pollen contact and subsequent allergy to birch pollen. Clinical Allergy 10: 585-591

Bohle B, Radakovics A, Jahn-Schmid B, Hoffmann-Sommergruber K, Fischer GF, Ebner C (2003) Bet v 1, the major birch pollen allergen, initiates sensitization to Api g 1, the major allergen in celery: evidence at the T cell level. European Journal of Immunology 33: 3303-3310

Bohle B, Vieths S (2004) Improving diagnostic tests for food allergy with recombinant allergens. Methods 32: 292-299

Bohle B, Zwölfer B, Heratizadeh A, Jahn-Schmid B, Antonia YD, Alter M, Keller W, Zuidmeer L, van Ree R, Werfel T, Ebner C (2006) Cooking birch pollen-related food: Divergent consequences for IgE- and T cell-mediated reactivity in vitro and in vivo. Journal of Allergy and Clinical Immunology 118: 242-249

Bolhaar STHP, Tiemessen MM, Zuidmeer L, van Leeuwen A, Hoffmann-Sommergruber K, Bruijnzeel-Koomen CAFM, Taams LS, Knol EF, van Hoffen E, van Ree R, Knulst AC (2004a) Efficacy of birch-pollen immunotherapy on cross-reactive food allergy confirmed by skin tests and double-blind food challenges. Clinical and Experimental Allergy 34: 761-769

Bolhaar STHP, Van de Weg WE, Van Ree R, Gonzalez-Macebo E, Zuidmeer L, Bruijnzeel-Koomen CAFM, Fernandez-Rivas M, Jansen J, Hoffmann-Sommergruber K, Knulst AC, Gilissen LJWJ (2005a) In vivo assessment with prick-to-prick testing and double-blind, placebo-controlled food challenge of allergenicity of apple cultivars. Journal of Allergy And Clinical Immunology 116: 1080-1086

Bolhaar STHP, Van Ree R, Bruijnzeel-Koomen CAFM, Knulst AC, Zuidmeer L (2004b) Allergy to jackfruit: A novel example of Bet v 1-related food allergy. Allergy: European Journal of Allergy and Clinical Immunology 59: 1187-1192

Bolhaar STHP, van Ree R, Ma Y, Bruijnzeel-Koomen CAFM, Vieths S, Hoffmann-Sommergruber K, Knulst AC, Zuidmeer L (2005b) Severe allergy to sharon fruit caused by birch pollen. International Archives Of Allergy And Immunology 136: 45-52

Bollen MA, Aranzazu G, Cordewener JHG, Wichers HJ, G. HJPF, Savelkoul HFJ, Van Boekel MAJS (in press) Purification and characterization of natural bet v 1 from birch pollen and related allergens from carrot and celery.

Bonin A, Bellemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. Molecular Ecology 13: 3261-3273

Breiteneder H, Ebner C (2000) Molecular and biochemical classification of plant-derived food allergens. Journal of Allergy and Clinical Immunology 106: 27-36

Breiteneder H, Ferreira F, Hoffmann-Sommergruber K, Ebner C, Breitenbach M, Rumpold H, Kraft D, Scheiner O (1993) Four recombinant isoforms of Cor a I, the major allergen of hazel pollen, show different IgE-binding properties. European Journal of Biochemistry 212: 355-362

Breiteneder H, Ferreira F, Reikerstorfer A, Duchene M, Valenta R, Hoffmann-Sommergruber K, Ebner C, Breitenbach M, Kraft D, Scheiner O (1992) Complementary DNA cloning and expression in *Escherichia coli* of Aln g I, the major allergen in pollen of alder (*Alnus glutinosa*). Journal of Allergy and Clinical Immunology 90: 909-917

Breiteneder H, Hoffmann-Sommergruber K, O'Riordain G, Susani M, Ahorn H, Ebner C, Kraft D, Scheiner O (1995) Molecular characterization of Api g 1, the major allergen of celery (*Apium graveolens*), and its immunological and structural relationships to a group of 17-kDa tree pollen allergens. European Journal of Biochemistry 15: 484-489

Breiteneder H, Pettenburger K, Bito A, Valenta R, Kraft D, Rumpold H, Scheiner O, Breitenbach M (1989) The gene coding for the major birch pollen allergen BetvI, is highly homologous to a pea disease resistance response gene. EMBO Journal 8: 1935-1938

Breiteneder H, Radauer C (2004) A classification of plant food allergens. Journal of Allergy and Clinical Immunology 113: 821-830

Brenna O, Pompei C, Ortolani C, Pravettoni V, Farioli L, Pastorello EA (2000) Technological processes to decrease the allergenicity of peach juice and nectar. Journal of Agricultural and Food Chemistry 48: 493-497

Burney P, Malmberg E, Chinn S, Jarvis D, Luczynska C, Lai E (1997) The distribution of total and specific serum IgE in the European Community Respiratory Health Survey. Journal of Allergy and Clinical Immunology 99: 314-322

Cadot P, Díaz JF, Proost P, Van Damme J, Engelborghs Y, Stevens EAM, Ceuppens JL (2000) Purification and characterization of an 18-kd allergen of birch (*Betula verrucosa*) pollen: Identification as a cyclophilin. Journal of Allergy and Clinical Immunology 105: 286-291

Cadot P, Lejoly M, Vanhoeyveld EW, Stevens EAM (1995) Influence of the pH of the extraction medium on the composition of birch (*Betula verrucosa*) pollen extracts. Allergy 50: 431-437

Carnés J, Ferrer A, Fernández-Caldas E (2006) Allergenicity of 10 different apple varieties. Annals of Allergy, Asthma and Immunology 96: 564-570

Castro AJ, Alche JD, Cuevas J, Romero PJ, Alche V, Rodriguez-Garcia MI (2003) Pollen from different olive tree cultivars contains varying amounts of the major allergen ole e 1. International Archives of Allergy and Immunology 131: 164-173

Cervera MT, Storme V, Soto A, Ivens B, Van Montagu M, Rajora OP, Boerjan W (2005) Intraspecific and interspecific genetic and phylogenetic relationships in the genus *Populus* based on AFLP markers. Theoretical and Applied Genetics 111: 1440-1456

Chen ZD, Manchester SR, Sun HY (1999) Phylogeny and evolution of the Betulaceae as inferred from DNA sequences, morphology, and paleobotany. American Journal of Botany 86: 1168-1181

Chinn S, Jarvis D, Luczynska C, Burney P (1998) Individual allergens as risk factors for bronchial responsiveness in young adults. Thorax 53: 662-667

Cline J, Braman JC, Hogrefe HH (1996) PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. Nucleic Acids Research 24: 3546-3551

Clot B (2001) Airborne birch pollen in Neuchâtel (Switzerland): onset, peak and daily patterns. Aerobiologia 17: 25-29

Cook C, Heath F, Thompson RL (2000) A meta-analysis of response rates in Web- or internet-based surveys. Educational and Psychological Measurement 60: 821-836

Cornelisse-Vermaat JR, Voordouw J, Yiakoumaki V, Theodoridis G, Frewer LJ (2007) Food-allergic consumers' labelling preferences: a cross-cultural comparison. Eur J Public Health: ckm032

Corsico R, Falagiani P, Ariano R, Berra D, Biale C, Bonifazi F, Campi P, Feliziani V, Frenguelli G, Galimberti M, Gallesio MT, Liccardi G, Loreti A, Marcer G, Marcucci F, Meriggi A, Minelli M, Nardelli R, Nardi G, Negrini CA, Papa G, Piu G, Pozzan M, Purello D'Ambrosio F, Riva G (2000) An epidemiological survey on the allergological importance of some emerging pollens in Italy. Journal of Investigational Allergology and Clinical Immunology 10: 155-161

CPRO-DLO (1999) 19e Rassenlijst voor fruitgewassen. Deel grootfruit, Wageningen

Crimi E, Voltolini S, Gianiorio P, Orengo G, Troise C, Brusasco V, Crimi P, Negrini AC (1990) Effect of seasonal exposure to pollen on specific bronchial sensitivity in allergic patients. Journal of Allergy and Clinical Immunology 85: 1014-1019

Crimi P, Boidi M, Minale P, Tazzer C, Zanrdi S, Ciprandi G (1999) Differences in prevalence of allergic sensitization in urban and rural school children. Annals of Allergy, Asthma and Immunology 83: 252-256

D'Amato G, Spieksma FTM, Liccardi G, Jäger S, Russo M, Kontou-Fili K, Nikkels H, Wüthrich B, Bonini S (1998) Pollen-related allergy in Europe. Allergy 53: 567-578

De Groot JCW, Fiers MWEJ, Van Ham RCHJ, America AHP (2008) Post alignment clustering procedure for comparative quantitative proteomics LC-MS Data. Proteomics 8: 32-36

De Jong PC (1993) An introduction to *Betula*: its morphology, evolution classification and distribution with a survey of recent work. In: The IDS *Betula* Symposium, International Dendrology Society, Sussex, U.K.

Deliza R, Rosenthal A, Hedderley D, MacFie HJH, Frewer LJ (1999) The Importance of Brand, Product Information and Manufacturing Process in the Development of Novel Environmentally Friendly Vegetable Oils. Journal of International Food and Agribusiness Marketing 10: 67-77

Detandt M, Nolard N (2000) The fluctuations of the allergenic pollen content of the air in Brussels (1982 to 1997). Aerobiologia 16: 55-61

Dodo HW, Konan KN, Egnin M, Viquez OM (2007) Alleviating peanut allergy using genetic engineering: the silencing of the immunodominant allergen Ara h 2 leads to its significant reduction and a decrease in peanut allergenicity. Plant Biotechnology Journal 6: 135-145

Durham SR, Till SJ (1998) Immunologic changes associated with allergen immunotherapy. Journal of Allergy and Clinical Immunology 102: 157-164

Ebner C, Birkner T, Valenta R, Rumpold H, Breitenbach M, Scheiner O, Kraft D (1991) Common epitopes of birch pollen and apples - Studies by Western and Northern blot. Journal of Allergy and Clinical Immunology 88: 588-594

Ekramoddoullah AKM, Yu XS, Sturrock R, Zamani A, Taylor D (2000) Detection and seasonal expression pattern of a pathogenesis-related protein (PR-10) in Douglas-fir (*Pseudotsuga menziesii*) tissues. Physiologia Plantarum 110: 240-247

Emberlin J, Detandt M, Gehrig R, Jaeger S, Nolard N, Rantio-Lehtimäki A (2002) Responses in the start of *Betula* (birch) pollen seasons to recent changes in spring temperatures across Europe. International Journal of Biometeorology 46: 159-170

Eriksson NE, Holmén A, Möller C, Wihl JÅ (1998) Sensitization according to skin prick testings in atopic patients with asthma or rhinitis at 24 allergy clinica in Northern Europe and Asia. Allergology International 47: 187-196

Fagan JK, Scheft PA, Hryhorczuk D, Ramakrishnan V, Ross M, Persky V (2001) Prevalence o f asthma and other allergic diseases in an adolescent population: Association with gender and race. Annals of Allergy, Asthma and Immunology 86: 177-184

Fahlbusch B, Hornung D, Heinrich J, Dahse HM, Jager L (2000) Quantification of group 5 grass pollen allergens in house dust. Clinical and Experimental Allergy 30: 1645-1652

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 164: 1567-1587

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: Dominant markers and null alleles. Molecular Ecology Notes 7: 574-578

Farris JS (1989) The Retention Index and the Rescaled Consistency Index. Cladistics-the International Journal of the Willi Hennig Society 5: 417-419

Fernández-Caldas E, Carnés J, Iraola V, Casanovas M (2007) Comparison of the allergenicity and Ole e 1 content of 6 varieties of *Olea europaea* pollen collected during 5 consecutive years. Annals of Allergy, Asthma and Immunology 98: 464-470

Fernández-Rivas M, Bolhaar S, González-Mancebo E, Asero R, van Leeuwen A, Bohle B, Ma Y, Ebner C, Rigby N, Sancho AI, Miles S, Zuidmeer L, Knulst A, Breiteneder H, Mills C,

Hoffmann-Sommergruber K, van Ree R (2006) Apple allergy across Europe: How allergen sensitization profiles determine the clinical expression of allergies to plant foods. Journal of Allergy and Clinical Immunology 118: 481-488

Fernández-Rivas M, González-Mancebo E, Rodríguez-Pérez R, Benito C, Sánchez-Monge R, Salcedo G, Alonso MD, Rosado A, Tejedor MA, Vila C, Casas ML (2003) Clinically relevant peach allergy is related to peach lipid transfer protein, Pru p 3, in the Spanish population. Journal of Allergy and Clinical Immunology 112: 789-795

Fernandez-Rivas M, Miles S (2004) Food allergies: Clinical and psychosocial perspectives. In: MIlls ENC, Shewry PR (eds) Plant Food Allergies. Blackwell Science Ltd., Oxford, pp 1-23

Fernández-Rivas M, Van Ree R, Cuevas M (1997) Allergy to Rosaceae fruits without related pollinosis. Journal of Allergy and Clinical Immunology 100: 728-733

Ferreira F, Ebner C, Kramer B, Casari G, Briza P, Kungl AJ, Grimm R, Jahn-Schmid B, Breiteneder H, Kraft D, Breitenbach M, Rheinberger H-J, Scheiner O (1998) Modulation of IgE reactivity of allergens by site-directed mutagenesis: potential use of hypoallergenic variants for immunotherapy. Faseb Journal 12: 231-242

Ferreira F, Hawranek T, Gruber P, Wopfner N, Mari A (2004) Allergic cross-reactivity: from gene to the clinic. Allergy 59: 243-267

Ferreira F, Hirtenlehner K, Jilek A, Godnik-Cvar J, Breiteneder H, Grimm R, Hoffmann-Sommergruber K, Scheiner O, Kraft D, Breitenbach M, Rheinberger H-J, Ebner C (1996) Dissection of immunoglobulin E and T lymphocyte reactivity of isoforms of the major birch pollen allergen Bet v 1: Potential use of hypoallergenic isoforms for immunotherapy. Journal of Experimental Medicine 183: 599-609

Ferreira F, Hirthenlehner K, Briza P, Breiteneder H, Scheiner O, Kraft D, Breitenbach M, Ebner C (1997) Isoforms of atopic allergens with reduced allergenicity but conserved T cell antigenicity: Possible use for specific immunotherapy. International Archives of Allergy and Immunology 113: 125-127

Feugey L, Strullu DG, Poupard P, Simoneau P (1999) Induced defence responses limit Hartig net formation in ectomycorrhizal birch roots. New Phytologist 144: 541-547

Finucane ML, Slovic P, Mertz CK, Flynn J, Satterfield TA (2000) Gender, race, and perceived risk: The 'white male' effect. Health, Risk and Society 2: 170-172

Flores T, Alape-Girón A, Flores-Díaz M, Flores HE (2002) Ocatin. A novel tuber storage protein from the Andean tuber crop oca with antibacterial and antifungal activities. Plant Physiology 128: 1291-1302

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151: 1531-1545

Frei T, Leuschner RM (2000) A change from grass pollen induced allergy to tree pollen induced allergy: 30 years of pollen observation in Switzerland. Aerobiologia 16: 407-416

Frewer L (2000) Risk perception and risk communication about food safety issues. Nutrition Bulletin 25: 31-33

Frewer L (2003) Societal issues and public attitudes towards genetically modified foods. Trends In Food Science & Technology 14: 319-332

Frewer L, Howard C, Shepherd R (1997) Public concerns in the United Kingdom and specific applications of genetic engineering: risk, benefit, and ethics. Science, Technology & Human Values 22: 98-124

Frewer L, Lassen J, Kettlitz B, Scholderer J, Beekman V, Berdal KG (2004) Societal aspects of genetically modified foods. Food And Chemical Toxicology 42: 1181-1193

Frewer LJ, Miles S, Marsh R (2002) The media and genetically modified foods: Evidence in support of social amplification of risk. Risk Analysis 22: 701-711

Friedl-Hajek R, Radauer C, O'Riordain G, Hoffmann-Sommergruber K, Leberl K, Scheiner O, Breiteneder H (1999) New Bet v 1 isoforms including a naturally occurring truncated form of the protein derived from Austrian birch pollen. Molecular Immunology 36: 639-645

Fritsch R, Bohle B, Vollmann U, Wiedermann U, Jahn-Schmid B, Krebitz M, Breiteneder H, Kraft D, Ebner C (1998) Bet v 1, the major birch pollen allergen, and Mal d 1, the major apple allergen, cross-react at the level of allergen-specific T helper cells. Journal of Allergy and Clinical Immunology 102: 679-686

Fujimoto Y, Nagata R, Fukasawa H, Yano K, Azuma M, Iida A, Sugimoto S, Shudo K, Hashimoto Y (1998) Purification and cDNA cloning of cytokinin-specific binding protein from mung bean (*Vigna radiata*). European Journal of Biochemistry 258: 794-802

Furlow JJ (1990) The genera of Betulaceae in the southeastern United States. Journal of the Arnold Arboretum 71: 1-67

Gafvelin G, Thunberg S, Kronqvist M, Gronlund H, Gronneberg R, Troye-Blomberg M, Akdis M, Fiebig H, Purohit A, Horak F, Reisinger J, Niederberger V, Akdis CA, Cromwell O, Pauli G, Valenta R, van Hage M (2005) Cytokine and antibody responses in birch-pollen-allergic patients treated with genetically modified derivatives of the major birch pollen allergen Bet v 1. International Archives Of Allergy And Immunology 138: 59-66

Gajhede M, Osmark P, Poulsen FM, Ipsen H, Larsen JN, Van Neerven RJJ, Schou C, Løwenstein H, Spangfort MD (1996) X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. Nature Structural Biology 3: 1040-1045

Ganiere P, Chern WS, Hahn D (2006) A continuum of consumer attitudes toward genetically modified foods in the United States. Journal of Agricultural and Resource Economics 31: 129-149

Gao ZS, van de Weg WE, Schaart JG, Schouten HJ, Tran DH, Kodde LP, van der Meer IM, van der Geest AHM, Kodde J, Breiteneder H, Hoffmann-Sommergruber K, Bosch D, Gilissen L (2005a) Genomic cloning and linkage mapping of the Mal d 1 (PR-10) gene family in apple (*Malus domestica*). Theoretical And Applied Genetics 111: 171-183

Gao ZS, Van de Weg WE, Schaart JG, Van Arkel G, Breiteneder H, Hoffmann-Sommergruber K, Gilissen LJWJ (2005b) Genomic characterization and linkage mapping of the apple allergen genes Mal d 2 (thaumatin-like protein) and Mal d 4 (profilin). Theoretical and Applied Genetics 111: 1087-1097

Gaskell G, Allansdottir A, Allum N, Corchero C, Fiscghler C, Hampel J (2006) Eurobarometer 64.3: Europeans and Biotechnology in 2005: Patterns and Trends. INRA

Ghosh D, Gupta-Bhattacharya S (2008) Structural insight into protein T1, the non-allergenic member of the Bet v 1 allergen family-An in silico analysis. Molecular Immunology 45: 456-462

Gilissen LJWJ, Bolhaar STH, Matos CI, Rouwendal GJA, Boone MJ, Krens FA, Zuidmeer L, van Leeuwen A, Akkerdaas J, Hoffmann-Sommergruber K, Knulst AC, Bosch D, van de Weg WE, van Ree R (2005) Silencing the major apple allergen Mal d 1 by using the RNA interference approach. Journal of Allergy and Clinical Immunology 115: 364-369

Gioulekas D, Papakosta D, Damialis A, Spieksma F, Giouleka P, Patakas D (2004) Allergenic pollen records (15 years) and sensitization in patients with respiratory allergy in Thessaloniki, Greece. Allergy 59: 174-184

Gislason D, Björnsson E, Gislason T, Janson C, Sjöberg O, Elfman L, Boman G (1999) Sensitization to airborne and food allergens in Reykjavík (Iceland) and Uppsala (Sweden) - a comparative study. Allergy 54: 1160-1167

Goodman RE, Hefle SL, Taylor SL, Van Ree R (2005) Assessing genetically modified crops to minimize the risk of increased food allergy: A review. International Archives of Allergy and Immunology 137: 153-166

Gowland MH (2001) Food allergen avoidance - the patient's viewpoint. Allergy 56: 117-120

Grote M, Valenta R, Reichelt R (2003) Abortive pollen germination: A mechanism of allergen release in birch, alder, and hazel revealed by immunogold electron microscopy. Journal of Allergy and Clinical Immunology 111: 1017-1023

Gruber P, Vieths S, Wangorsch A, Nerkamp J, Hofmann T (2004) Maillard reaction and enzymatic browning affect the allergenicity of Pru av 1, the major allergen from cherry (Prunus avium). Journal of Agricultural and Food Chemistry 52: 4002-4007

Guarner F, Bourdet-Sicard R, Brandtzaeg P, Gill HS, McGuirk P, van Eden W, Versalovic J, Weinstock JV, Rook GAW (2006) Mechanisms of disease: The hygiene hypothesis revisited. Nature Clinical Practice Gastroenterology and Hepatology 3: 275-284

Heiss S, Fischer S, Müller WB, Weber B, Hirschwehr R, Spitzauer S, Kraft D, Valenta R (1996) Identification of a 60 kd cross-reactive allergen in pollen and plant-derived food. Journal of Allergy and Clinical Immunology 98: 938-947

Helsper J, Gilissen L, van Ree R, America AHP, Cordewener JHG, Bosch D (2002) Quadrupole time-of-flight mass spectrometry: A method to study the actual expression of allergen isoforms identified by PCR cloning. Journal of Allergy And Clinical Immunology 110: 131-138

Herman EM, Helm RM, Jung R, Kinney AJ (2003) Genetic modification removes an immunodominant allergen from soybean. Plant Physiology 132: 36-43

Hillis DM, Huelsenbeck JP (1992) Signal, noise, and reliability in molecular phylogenetic analyses. Journal of Heredity 83: 189-195

Ho P, Vermeer EB, Zhao JH (2006) Biotechnology and food safety in China: Consumers' acceptance or resistance? Development and Change 37: 227-254

Hoffmann-Sommergruber K, O'Riordain G, Ahorn H, Ebner C, Laimer da Camara Machado M, Pühringer H, Scheiner O, Breiteneder H (1999) Molecular characterization of Dau c 1, the Bet v 1 homologous protein from carrot and its cross-reactivity with Bet v 1 and Api g 1. Clinical and Experimental Allergy 29: 840-847

Hoffmann-Sommergruber K, Vanek-Krebitz M, Radauer C, Wen J, Ferreira F, Scheiner O, Breiteneder H (1997) Genomic characterization of members of the Bet v 1 family: genes coding for allergens and pathogenesis-related proteins share intron positions. Gene 197: 91-100

Horak F, Studnicka M, Gartner C, Veiter A, Tauber E, Urbanek R, Frischer T (2002) Parental farming protects children against atopy: longitudinal evidence involving skin prick tests. Clinical and Experimental Allergy 32: 1155-1159

Hourihane JOB (2001) The threshold concept in food safety and its applicability to food allergy. Allergy S67: 86-90

Huang JC, Chang FC, Wang CS (1997) Characterization of a lily tapetal transcript that shares sequence similarity with a class of intracellular pathogenesis-related (IPR) proteins. Plant Molecular Biology 34: 681-686

Huelsenbeck JP, Larget B, Miller RE, Ronquist F (2002) Potential applications and pitfalls of Bayesian inference of phylogeny. Systematic Biology 51: 673-688

Ipsen H, Hansen OC (1991) The Nh2-Terminal amino-acid-sequence of the immunochemically partial identical major allergens of alder (*Alnus glutinosa*) Aln g I, Birch (*Betula verrucosa*) Bet v I, Hornbeam (*Carpinus betulus*) Car b I and Oak (*Quercus alba*) Que a I pollens. Molecular Immunology 28: 1279-1288

Jacobsen E, Schouten HJ (2007) Cisgenesis strongly improves introgression breeding and induced translocation breeding of plants. Trends in Biotechnology 25: 219-223

Jaen A, Sunyer J, Basagana X, Chinn S, Zock JP, Anto JM, Burney P (2002) Specific sensitization to common allergens and pulmonary function in the European Community Respiratory Health Survey. Clinical and Experimental Allergy 32: 1713-1719

Jahn-Schmid B, Radakovics A, Luttkopf D, Scheurer S, Vieths S, Ebner C, Bohle B (2005) Bet v 1142-156 is the dominant T-cell epitope of the major birch pollen allergen and important for cross-reactivity with Bet v 1-related food allergens. Journal Of Allergy And Clinical Immunology 116: 213-219

Jan Ms, Fu TT, Huang CL (2007) A conjoint/logit analysis of consumers' responses to genetically modified Tofu in Taiwan. Journal of Agricultural Economics 58: 330-347

Jarolim E, Rumpold H, Endler AT, Ebner H, Breitenbach M, Scheiner O, Kraft D (1989) IgE and IgG antibodies of patients with allergy to birch pollen as tools to define the allergen profile of Betula verrucosa. Allergy 44: 385-395

Järvinen P, Palmé A, Morales LO, Lännenpää M, Keinänen M, Sopanen T, Lascoux M (2004) Phylogenetic relationships of *Betula* species (*Betulaceae*) based on nuclear ADH and chloroplast matK sequences. American Journal of Botany 91: 1834-1845

Jauhar PP (2003) Formation of 2n gametes in durum wheat haploids: Sexual polyploidization. Euphytica 133: 81-94

Jõgi R, Janson C, Björnsson E, Boman G, Björkstén B (1998) Atopy and allergic disorders among adults in Tartu, Estonia compared with Uppsala, Sweden. Clinical and Experimental Allergy 28: 1072-1080

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, Van De Wiel C, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevschi A, Marmiroli N, Aert R, Volckaert G, Rueda J, Linacero R, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Molecular Breeding 3: 381-390

Judo MSB, Wedel AB, Wilson C (1998) Stimulation and suppression of PCR-mediated recombination. Nucleic Acids Research 26: 1819-1825

Juniper EF, Guyatt GH (1991) Development and testing of a new measure of health-status for clinical-trials in Rhinoconjunctivitis. Clinical and Experimental Allergy 21: 77-83

Juniper EF, Thompson AK, Ferrie PJ, Roberts JN (2000) Development and validation of the Mini Rhinoconjunctivitis Quality of Life Questionnaire. Clinical and Experimental Allergy 30: 132-140

Karamloo F, Scheurer S, Wangorsch A, May S, Haustein D, Vieths S (2001a) Pyr c 1, the major allergen from pear (*Pyrus communis*), is a new member of the Bet v 1 allergen family. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences 756: 281-293

Karamloo F, Wangorsch A, Kasahara H, Davin LB, Haustein D, Lewis NG, Vieths S (2001b) Phenylcoumaran benzylic ether and isoflavonoid reductases are a new class of cross-reactive allergens in birch pollen, fruits and vegetables. European Journal of Biochemistry 268: 5310-5320

Karlsson AL, Alm R, Ekstrand B, Fjelkner-Modig S, Schiött A, Bengtsson U, Björk L, Hjernø K, Roepstorff P, Emanuelsson CS (2004) Bet v 1 homologues in strawberry identified as IgE-binding proteins and presumptive allergens. Allergy 59: 1277-1284

Keinänen M, Julkunen-Tiitto R, Rousi M, Tahvanianen J (1999) Taxonomic implications of phenolic variation in leaves of birch (*Betula* L.) species. Biochemical Systematics and Ecology 27: 243-254

Kelso JM, Jones RT, Tellez R, Yunginger JW (1995) Oral allergy syndrome successfully treated with pollen immunotherapy. Annals of Allergy, Asthma and Immunology 74: 391-396

Kerkhof M, Postma DS, Schouten JP, de Monchy JGR (2003) Allergic sensitization to indoor and outdoor allergens and relevance to bronchial hyperresponsiveness in younger and older subjects. Allergy 58: 1261-1267

Kihlström A, Lilja G, Pershagen G, Hedlin G (2002) Exposure to birch pollen in infancy and development of atopic disease in childhood. Journal of Allergy and Clinical Immunology 110: 78-84

Kihlström A, Lilja G, Pershagen G, Hedlin G (2003) Exposure to high doses of birch pollen during pregnancy, and risk of sensitization and atopic disease in the child. Allergy 58: 871-877

Kilpeläinen M, Terho EO, Helenius H, Koskenvuo M (2002) Childhood farm environment and asthma and sensitization in young adulthood. Allergy 57: 1130-1135

Kleine-Tebbe J, Wangorsch A, Vogel L, Crowell DN, Haustein UF, Vieths S (2002) Severe oral allergy syndrome and anaphylactic reactions caused by a Bet v 1-related PR-10 protein in soybean, SAM22. Journal of Allergy and Clinical Immunology 110: 797-804

Kluge AG, Farris JS (1969) Quantitative Phyletics and Evolution of Anurans. Systematic Zoology 18: 1-&

Knibb RC, Armstrong A, Booth DA, Platts RG, Booth IW, MacDonald A (1999) Psychological characteristics of people with perceived food intolerance in a community sample. Journal of Psychosomatic Research 47: 545-554

Koistinen KM, Hassinen VH, Gynther PAM, Lehesranta SJ, Keinänen SI, Kokko HI, Oksanen EJ, Tervahauta AI, Auriola S, Kärenlampi SO (2002) Birch PR-10c is induced by factors causing oxidative stress but appears not to confer tolerance to these agents. New Phytologist 155: 381-391

Koivikko A, Kupias R, Makinen Y, Pohjola A (1986) Pollen seasons: Forecasts of the most important allergenic plants in Finland. Allergy: European Journal of Allergy and Clinical Immunology 41: 233-242

Koopman WJM (2000) Identifying lettuce species (*Lactuca* subsect. *Lactuca*, Asteraceae): A practical application of flow cytometry. Euphytica 116: 151-159

Koopman WJM (2005) Phylogenetic signal in AFLP data sets. Systematic Biology 54: 197-217

Koopman WJM, Wissemann V, Cock Kd, Huylenbroeck Jv, Riek JD, Sabatino GJH, Visser D, Vosman B, Ritz CM, Maes B, Werlemark G, Nybom H, Debener T, Linde M, Smulders MJM (2008) AFLP markers as a tool to reconstruct complex relationships in the genus *Rosa* (Rosaceae). American Journal of Botany: in press

Kootstra HS, Vlieg-Boerstra BJ, Dubois AEJ (2008) Assessment of the reduced allergenic properties of the Santana apple. Annals of Allergy, Asthma and Immunology: in press

Koppelman SJ, Vlooswijk RAA, Knippels LMJ, Hessing M, Knol EF, van Reijsen FC, Bruijnzeel-Koomen CAFM (2001) Quantification of major peanut allergens Ara h 1 and Ara h 2 in the peanut varieties Runner, Spanish, Virginia, and Valencia, bred in different parts of the world. Allergy 56: 132-137

Kosunen TU, Höök-Nikanne J, Salomaa A, Sarna S, Aromaa A, Haahtela T (2002) Increase of allergen-specific immunoglobulin E antibodies from 1973 to 1994 in a Finnish population and a possible relationship to Helicobacter pylori infections. Clinical and Experimental Allergy 32: 373-378

Krause TG, Koch A, Poulsen LK, Kristensen B, Olsen OR, Melbye M (2002) Atopic sensitization among children in an Arctic environment. Clinical and Experimental Allergy 32: 367-372

Kwaasi AAA, Harfi HA, Parhar RS, Collison KS, Al-Sedairy ST, Al-Mohanna FA (2000) Cultivar-specific IgE-epitopes in date (*Phoenix dactylifera* L.) fruit allergy. Correlation of skin test reactivity and IgE-binding properties in selecting date cultivars for allergen standardization. International Archives of Allergy and Immunology 123: 137-144

Lahti A, Bjorksten F, Hannuksela M (1980) Allergy to birch pollen and apple, and cross-reactivity of the allergens studied with the RAST. Allergy: European Journal of Allergy and Clinical Immunology 35: 297-300

Lake AA, Hyland RM, Mathers JC, Rugg-Gunn AJ, Wood CE, Adamson AJ (2006) Food shopping and preparation among the 30-somethings: Whose job is it? (The ASH30 study). British Food Journal 108: 475-486

Larsen JN (1995) Isoallergens - significance in allergen exposure and response. ACI News 7: 141-160

Larsen JN, Stroman P, Ipsen H (1992) PCR based cloning and sequencing of isogenes encoding the tree pollen major allergen Car b I from *Carpinus betulus*, hornbeam. Molecular Immunology 29: 703-711

Latalowa M, Mietus M, Uruska A (2002) Seasonal variations in the atmospheric Betula pollen count in Gdánsk (southern Baltic coast) in relation to meteorological parameters. Aerobiologia 18: 33-43

Le LQ, Mahler V, Lorenz Y, Scheurer S, Biemelt S, Vieths S, Sonnewald U (2006) Reduced allergenicity of tomato fruits harvested from Lyc e 1-silenced transgenic tomato plants. Journal of Allergy and Clinical Immunology 118: 1176-1183

Lemmetyinen J, Keinonen K, Sopanen T (2004) Prevention of the flowering of a tree, silver birch. Molecular Breeding 13: 243-249

Li J, Shoup S, Chen Z (2005) Phylogenetics of *Betula* (Betulaceae) inferred from sequences of nuclear ribosomal DNA. Rhodora 107: 69-86

Linneberg A, Nielsen NH, Madsen F, Frølund L, Dirksen A, Jørgensen T (2000) Increasing prevalence of specific IgE to aeroallergens in an adult population: Two cross-sectional surveys 8 years apart - The Copenhagen Allergy Study. Journal of Allergy and Clinical Immunology 106: 247-252

Liu JJ, Ekramoddoullah AKM (2003a) Root-specific expression of a western white pine PR10 gene is mediated by different promoter regions in transgenic tobacco. Plant Molecular Biology 52: 103-120

Liu JJ, Ekramoddoullah AKM (2004) Characterization, expression and evolution of two novel subfamilies of Pinus monticola cDNAs encoding pathogenesis-related (PR)-10 proteins. Tree Physiology 24: 1377-1385

Liu JJ, Ekramoddoullah AKM, Yu XS (2003b) Differential expression of multiple PR10 proteins in western white pine following wounding, fungal infection and cold-hardening. Physiologia Plantarum 119: 544-553

Lødrup Carlsen KC, Carlsen KH, Buchmann MS, Wikstrøm J, Mehl R, GAIN (2002) Cockroach sensitivity in Norway: a previously unidentified problem? Allergy 57: 529-533

Lüttkopf D, Müller U, Skov PS, Ballmer-Weber BK, Wüthrich B, Hansen KS, Poulsen LK, Kästner M, Haustein D, Vieths S (2002) Comparison of four variants of a major allergen in hazelnut (*Corylus avellana*) Cor a 1.04 with the major hazel pollen allergen Cor a 1.01. Molecular Immunology 38: 515-525

Maas CJM, Snijders TAB (2003) The Multilevel Approach to Repeated Measures for Complete and Incomplete Data. Quality and Quantity 37: 71-89

Mahler V, Fischer S, Heiss S, Duchêne M, Kraft D, Valenta R (2001) cDNA cloning and characterization of a cross-reactive birch pollen allergen: identification as a pectin esterase. International Archives of Allergy and Applied Immunology: 64-66

Maleki SJ, Chung SY, Champagne ET, Raufman JP (2000) The effects of roasting on the allergenic properties of peanut proteins. Journal of Allergy and Clinical Immunology 106: 763-768

Maliouchenko O, Palme AE, Buonamici A, Vendramin GG, Lascoux M (2007) Comparative phylogeography and population structure of European Betula species, with particular focus on B. pendula and B. pubescens. Journal of Biogeography 34: 1601-1610

Mari A, Wallner M, Ferreira F (2003) Fagales pollen sensitization in a birch-free area: a respiratory cohort survey using Fagales pollen extracts and birch recombinant allergens (rBet v 1, rBet v 2, rBet v 4). Clinical and Experimental Allergy 33: 1419-1428

Marklund B, Ahlstedt S, Nordström G (2004) Health-related quality of life among adolescents with allergy-like conditions - With emphasis on food hypersensitivity. Health and Quality of Life Outcomes 2

Markovic-Housley Z, Degano M, Lamba D, von Roepenack-Lahaye E, Clemens S, Susani M, Ferreira F, Scheiner O, Breiteneder H (2003) Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier. Journal of Molecular Biology 325: 123-133

Marzban G, Puehringer H, Dey R, Brynda S, Ma Y, Martinelli A, Zaccarini M, van der Weg E, Housley Z, Kolarich D, Altmann F, Laimer M (2005) Localisation and distribution of the major allergens in apple fruits. Plant Science 169: 387-394

McCabe M, Lyons RA (2001) Management of peanut allergy. The Lancet 357: 1531-1532

McFadden HG, Chapple R, de Feyter R, Dennis E (2001) Expression of pathogenesis-related genes in cotton stems in response to infection by *Verticillium dahliae*. Physiological and Molecular Plant Pathology 58: 119-131

McGregor CE, Lambert CA, Greyling MM, Louw JH, Warnich L (2000) A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. Euphytica 113: 135-144

Meltzer EO (2001) Quality of life in adults and children with allergic rhinitis. Journal of Allergy and Clinical Immunology 108: S45-S53

Miles S, Bolhaar S, Gonzáez-Mancebo E, Hafner C, Hoffmann-Sommergruber K, Fernández-Rias M, Knulst A (2005) Attitudes towards low-allergen food in food allergic consumers. Nutrition and Food Science 35: 220-228

Miles S, Frewer LJ (2001) Investigating specific concerns about different food hazards. Food Quality and Preference 12: 47-61

Mills ENC, Mackie AR, Burney P, Beyer K, Frewer L, Madsen C, Botjes E, Crevel RWR, Van Ree R (2007) The prevalence, cost and basis of food allergy across Europe. Allergy: European Journal of Allergy and Clinical Immunology 62: 717-722

Mills ENC, Valovirta E, Madsen C, Taylor SL, Vieths S, Anklam E, Baumgartner S, Koch P, Crevel RWR, Frewer L (2004) Information provision for allergic consumers - where are we going with food allergen labelling? Allergy 59: 1262-1268

Miralles JC, Caravaca F, Guillén F, Lombardero M, Negro JM (2002) Cross-reactivity between *Platanus* pollen and vegetables. Allergy 57: 146-149

Mittag D, Akkerdaas J, Ballmer-Weber BK, Vogel L, Wenising M, Becker WM, Koppelman SJ, Knulst AC, Helbling A, Hefle SL, van Ree R, Vieths S (2004) Ara h 8, a Bet v 1-homologous allergen from peanut, is a major allergen in patients with combined birch pollen and peanut allergy. Journal Of Allergy And Clinical Immunology 114: 1410-1417

Mittag D, Vieths S, Vogel L, Wagner-Loew D, Starke A, Hunziker P, Becker WM, Ballmer-Weber BK (2005) Birch pollen-related food allergy to legumes: identification and characterization of the Bet v 1 homologue in mungbean (*Vigna radiata*), Vig r 1. Clinical And Experimental Allergy 35: 1049-1055

Mogensen JE, Ferreras M, Wimmer R, Petersen SV, Enghild JJ, Otzen DE (2007) The major allergen from birch tree pollen, Bet v 1, binds and permeabilizes membranes. Biochemistry 46: 3356-3365

Mogensen JE, Wimmer R, Larsen JN, Spangfort MD, Otzen DE (2002) The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands. Journal of Biological Chemistry 277: 23684-23692

Moiseyev GP, Fedoreyeva LI, Zhuravlev YN, Yasnetskaya E, Jekel PA, Beintema JJ (1997) Primary structures of two ribonucleases from ginseng calluses - New members of the PR-10 family of intracellular pathogenesis-related plant proteins. FEBS Letters 407: 207-210

Moons A, Prinsen E, Bauw G, Van Montagu M (1997) Antagonistic effects of abscisic acid and jasmonates on salt stress-inducible transcripts in rice roots. Plant Cell 9: 2243-2259

Moverare R, Westritschnig K, Svensson M, Hayek B, Bende M, Pauli G, Sorva R, Haahtela T, Valenta R, Elfman L (2002) Different IgE reactivity profiles in birch pollen-sensitive patients from six European populations revealed by recombinant allergens: An imprint of local sensitization. International Archives of Allergy and Immunology 128: 325-335

Musidlowska-Persson A, Alm R, Emanuelsson C (2007) Cloning and sequencing of the Bet v 1-homologous allergen Fra a 1 in strawberry (*Fragaria ananassa*) shows the presence of an intron and little variability in amino acid sequence. Molecular Immunology 44: 1245-1252

Nagamitsu T, Kawahara T, Kanazashi A (2006) Endemic dwarf birch *Betula apoiensis* (Betulaceae) is a hybrid that originated from *Betula ermanii* and *Betula ovalifolia*. Plant Species Biology 21: 19-29

Nayga Jr RM, Fisher MG, Onyango B (2006) Acceptance of genetically modified food: Comparing consumer perspectives in the United States and South Korea. Agricultural Economics 34: 331-341

Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. Annual Review of Genetics 39: 121-152

Neudecker P, Lehmann K, Nerkamp J, Haase T, Wangorsch A, Fötisch K, Hoffmann S, Rösch P, Vieths S, Scheurer S (2003) Mutational epitope analysis of Pru av 1 and Api g 1, the major allergens of cherry (*Prunus avium*) and celery (*Apium graveolens*): correlating IgE reactivity with three-dimensional structure. Biochemical Journal 376: 97-107

Neudecker P, Schweimer K, Nerkamp J, Scheurer S, Vieths S, Sticht H, Rösch P (2001) Allergic cross-reactivity made visible - Solution structure of the major cherry allergen Pru av 1. Journal of Biological Chemistry 276: 22756-22763

Niederberger V, Pauli G, Gronlund H, Froschl R, Rumpold H, Kraft D, Valenta R, Spitzauer S (1998) Recombinant birch pollen allergens (rBet v 1 and rBet v 2) contain most of the IgE epitopes present in birch, alder, hornbeam, hazel, and oak pollen: A quantitative IgE inhibition study with sera from different populations. Journal of Allergy and Clinical Immunology 102: 579-591

Nilsson L, Björkstén B, Hattevig G, Kjellman B, Sigurs N, Kjellman NIM (1997) Season of birth as predictor of atopic manifestations. Archives of Disease in Childhood 76: 341-344

Nowak D, Heinrich J, Jörres R, Wassmer G, Berger J, Beck E, Boczor S, Claussen M, Wichmann HE, Magnussen H (1996) Prevalence of respiratory symptoms, bronchial hyperresponsiveness and atopy among adults: West and East Germany. European Respiratory Journal 9: 2541-2552

Odelberg SJ, Weiss RB, Hata A, White R (1995) Template-switching during DNA synthesis by Thermus aquaticus DNA polymerase I. Nucleic Acids Research 23: 2049-2057

Ohno S (1970) Evolution by gene duplication. Springer-Verlag, Heidelberg, Germany

Ohta T (2000) Evolution of gene families. Gene 259: 45-52

Ortolani C, Bruijnzeel-Koomen C, Bengtsson U, Bindslev-Jensen C, Björkstèn B, Høst A, Ispano M, Jarish R, Madsen C, Nekam K, Paganelli R, Poulsen LK, Wüthrich B (1999) Controversial aspects of adverse reactions to food. Allergy: European Journal of Allergy and Clinical Immunology 54: 27-45

Ortolani C, Ispano M, Pastorello E, Bigi A, Ansaloni R (1988) The oral allergy syndrome. Annals of Allergy 61: 47-52

Palme AE, Su Q, Palsson S, Lascoux M (2004) Extensive sharing of chloroplast haplotypes among European birches indicates hybridization among *Betula pendula*, *B. pubescens* and *B. nana*. Molecular Ecology 13: 167-178

Park CJ, Kim KJ, Shin R, Park JM, Shin YC, Paek KH (2004) Pathogenesis-related protein 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway. Plant Journal 37: 186-198

Pastorello EA, Pravettoni V, Farioli L, Ispano M, Fortunato D, Monza M, Giuffrida MG, Rivolta F, Scibola E, Ansaloni R, Incorvaia C, Conti A, Ortolani C (1999) Clinical role of a lipid transfer protein that acts as a new apple- specific allergen. Journal of Allergy and Clinical Immunology 104: 1099-1106

Pauli G, Oster JP, Deviller P, Heiss S, Bessot JC, Susani M, Ferreira F, Kraft D, Valenta R (1996) Skin testing with recombinant allergens rBet v 1 and birch profilin, rBet v 2: Diagnostic value for birch pollen and associated allergies. Journal of Allergy and Clinical Immunology 97: 1100-1109

Pekkinen M, Varvio S, Kulju KKM, Kärkkäinen H, Smolander S, Viherä-Aarnio A, Koski V, Sillanpää MJ (2005) Linkage map of birch, *Betula pendula* Roth, based on microsatellites and amplified fragemnt length polymorphisms. Genome 48: 619-625

Penás E, Restani P, Ballabio C, Préstamo G, Fiocchi A, Gómez R (2006) Assessment of the residual immunoreactivity of soybean whey hydrolysates obtained by combined enzymatic proteolysis and high pressure. European Food Research and Technology 222: 286-290

Petrovska N, Wu X, Donato R, Wang Z, Ong EK, Jones E, Forster J, Emmerling M, Sidoli A, O'Hehir R, Spangenberg G (2005) Transgenic ryegrasses (*Lolium* spp.) with down-regulation of main pollen allergens. Molecular Breeding 14: 489-501

Pinto MP, Ribeiro A, Regalado AP, Rodrigues-Pousada C, Ricardo CPP (2005) Expression of *Lupinus albus* PR-10 proteins during root and leaf development. Biologia Plantarum 49: 187-193

Plaschke PP, Janson C, Norrman E, Björnsson E, Ellbjär S, Järvholm B (2000) Onset and remission of allergic rhinitis and asthma and the relationship with atopic sensitization and smoking. American Journal of Respiratory and Critical Care Medicine 162: 920-924

Poupard P, Brunel N, Leduc N, Viémont JD, Strullu D-G, Simoneau P (2001) Expression of a Bet v 1 homologue gene encoding a PR 10 protein in birch roots: induction by auxin and localization of the transcripts by in situ hybridization. Australian Journal of Plant Physiology 28: 57-63

Poupard P, Parisi L, Campion C, Ziadi S, Simoneau P (2003) A wound- and ethephon-inducible PR-10 gene subclass from apple is differentially expressed during infection with a compatible and an incompatible race of *Venturia inaequalis*. Physiological and Molecular Plant Pathology 62: 3-12

Poupard P, Strullu DG, Simoneau P (1998) Two members of the Bet v 1 gene family encoding birch pathogenesis-related proteins display different patterns of root expression and wound-inducibility. Australian Journal of Plant Physiology 25: 459-464

Primavesi L, Brenna OV, Pompei C, Pravettoni V, Farioli L, Pastorello EA (2006) Influence of cultivar and processing on cherry (*Prunus avium*) allergenicity. Journal of Agricultural and Food Chemistry 54: 9930-9935

Prioult G, Nagler-Anderson C (2005) Mucosal immunity and allergic responses: Lack of regulation and/or lack of microbial stimulation? Immunological Reviews 206: 204-218

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959

Pühringer H, Moll D, Hoffmann-Sommergruber K, Watillon B, Katinger H, Machado MLD (2000) The promoter of an apple Ypr10 gene, encoding the major allergen Mal d 1, is stress- and pathogen-inducible. Plant Science 152: 35-50

Rak S, Werner S, De Vos C (1992) Influence of the priming effect with alder pollen on nonspecific bronchial hyperreactivity of asthmatics specifically allergic to birch pollen. International Archives of Allergy and Immunology 99: 449-451

Ramanna MS, Jacobsen E (2003) Relevance of sexual polyploidization for crop improvement - A review. Euphytica 133: 3-18

Rasmussen A (2002) The effects of climate change on the birch pollen season in Denmark. Aerobiologia 18: 253-265

Raukas-Kivioja A, Raukas E, Loit HM, Kiviloog J, Rönmark E, Larsson K, Lundbäck B (2003) Allergic sensitization among adults in Tallinn, Estonia. Clinical and Experimental Allergy 33: 1342-1348

Regel E (1865) Bemerkungen über die Gattungen Betula und Alnus nebst Beschreibung einiger neuer Arten. Bull. Soc. Natural. Moscou 38: 388-434

Richard-Molard C, Brugière N, Moille M, Carrayol E, Limami AM (2004) Molecular characterization of a gene encoding a vegetative storage protein (CiVSP) from *Cichorium intybus* and its expression in the root and shoot in relation to nitrogen status and pathogen resistance. Physiologia Plantarum 121: 568-577

Riedler J, Eder W, Oberfeld G, Schreuer M (2000) Austrian children living on a farm have less hay fever, asthma and allergic sensitization. Clinical and Experimental Allergy 30: 194-200

Robert N, Ferran J, Breda C, Coutos-Thévenot P, Boulay M, Buffard D, Esnault R (2001) Molecular characterization of the incompatible interaction of *Vitis vinifera* leaves with *Pseudomonas syringae* pv. pisi: Expression of genes coding for stilbene synthase and class 10 PR protein. European Journal of Plant Pathology 107: 249-261

Rokas A, Carroll SB (2006) Bushes in the tree of life. PLoS Biology 4: 1899-1904

Rona RJ, Keil T, Summers C, Gislason D, Zuidmeer L, Sodergren E, Sigurdardottir ST, Lindner T, Goldhahn K, Dahlstrom J, McBride D, Madsen C (2007) The prevalence of food allergy: A meta-analysis. Journal of Allergy and Clinical Immunology 120: 638-646

Rönmark E, Perzanowski M, Platts-Mills T, Lundbäck B (2003) Four-year incidence of allergic sensitization among schoolchildren in a community where allergy to cat and dog dominates sensitization: Report from the Obstructive Lung Disease in Northern Sweden Study Group. Journal of Allergy and Clinical Immunology 112: 747-754

Rossi RE, Monasterolo G, Monasterolo S (2003) Detection of specific IgE antibodies in the sera of patients allergic to birch pollen using recombinant allergens Bet v 1, Bet v 2, Bet v 4: evaluation of different IgE reactivity profiles. Allergy 58: 929-932

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497

Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor A, Powell W, Waugh R (1997) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theoretical and Applied Genetics 95: 714-722

Saarinen UM, Kajosaari M (1995) Breastfeeding as prophylaxis against atopic disease: prospective follow-up study until 17 years old. Lancet 346: 1065-1069

Sànchez-Monge R, Lombardero M, Garcìa-Sellès FJ, Barber D, Salcedo G (1999) Lipid-transfer proteins are relevant allergens in fruit allergy. Journal of Allergy and Clinical Immunology 103: 514-519

Schäppi GF, Suphioglu C, Taylor PE, Knox RB (1997) Concentrations of the major birch tree allergen Bet v 1 in pollen and respirable fine particles in the atmosphere. Journal of Allergy and Clinical Immunology 100: 656-661

Schaub B, Lauener R, von Mutius E (2006) The many faces of the hygiene hypothesis. Journal of Allergy and Clinical Immunology 117: 969-977

Schenk MF, Cordewener JHG, America T, Westende WPCvt, Smulders MJM, Gilissen LJWJ (in preparation) Characterization of PR-10 genes from eight Betula species and detection of PR-10 (Bet v 1) isoforms in birch pollen.

Schenk MF, Fischer ARH, Frewer LJ, Gilissen LJWJ, E. J, Smulders MJM (in press) The influence of perceived benefits on acceptance of GM applications for allergy prevention. Health, Risk and Society

Schenk MF, Gilissen LJWJ, Esselink GD, Smulders MJM (2006) Seven different genes encode a diverse mixture of isoforms of Bet v I, the major birch pollen allergen. BMC Genomics 7: 168

Schenk MF, Tienpoint C-N, Koopman WJM, Gilissen LJWJ, Smulders MJM (submitted) Phylogenetic relationships in *Betula* (Betulaceae) based on AFLP markers.

Scheurer S, Son DY, Boehm M, Karamloo F, Franke S, Hoffmann A, Haustein D, Vieths S (1999) Cross-reactivity and epitope analysis of Pru a 1, the major cherry allergen. Molecular Immunology 36: 155-167

Scheurer S, Wangorsch A, Nerkamp J, Stahl Skov P, Ballmer-Weber B, Wüthrich B, Haustein D, Vieths S (2001) Cross-reactivity within the profilin panallergen family investigated by comparision of recombinant profilins from pear (Pyr c 4), cherry (Pru av 4) and celery (Api g 4) with birch pollen profilin Bet v 2. Journal of Chromatography B: Biomedical Sciences and Applications 756: 315-325

Schouten HJ, Krens FA, Jacobsen E (2006) Do cisgenic plants warrant less stringent oversight? Nature Biotechnology 24: 753

Seiberler S, Scheiner O, Kraft D, Lonsdale D, Valenta R (1994) Characterization of a birch pollen allergen, Bet v III, representing a novel class of Ca2+ binding proteins; specific expression in mature pollen and dependence of patients' IgE binding on protein-bound Ca2+. EMBO Journal 13: 3481-3486

Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Analytical Chemistry 68: 850-858

Sicherer SH, Noone SA, Munõz-Furlong A (2001) The impact of childhood food allergy on quality of life. Annals of Allergy, Asthma and Immunology 87: 461-464

Siegrist M (1998) Belief in gene technology: The influence of environmental attitudes and gender. Personality and Individual Differences 24: 861-866

Siegrist M (2000) The influence of trust and perceptions of risks and benefits on the acceptance of gene technology. Risk Analysis 20: 195-203

Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li GZ, McKenna T, Nold MJ, Richardson K, Young P, Geromanos S (2005) Quantitative proteomic analysis by accurate mass retention time pairs. Analytical Chemistry 77: 2187-2200

Silvestri M, Oddera S, Rossi GA, Crimi P (1996) Sensitization to airborne allergens in children with respiratory symptoms. Annals of Allergy, Asthma and Immunology 76: 239-244

Skamstrup Hansen K, Vieths S, Vestergaard H, Skov PS, Bindslev-Jensen C, Poulsen LK (2001) Seasonal variation in food allergy to apple. Journal of Chromatography B: Biomedical Sciences and Applications 756: 19-32

Skjøth CA, Sommer J, Stach A, Smith M, Brandt J (2007) The long-range transport of birch (*Betula*) pollen from Poland and Germany causes significant pre-season concentrations in Denmark. Clinical and Experimental Allergy 37: 1204-1212

Skvortsov AK (2002) A new system of the genus *Betula* L. Bull. Mosc. Soc. Natur. Biological Series. 107: 73-76

Slovic P (1987) Perception of Risk. Science 236: 280-285

Sofiev M, Siljamo P, Ranta H, Rantio-Lehtimäki A (2006) Towards numerical forecasting of long-range air transport of birch pollen: Theoretical considerations and a feasibility study. International Journal of Biometeorology 50: 392-402

Son DY, Scheurer S, Hoffmann A, Haustein D, Vieths S (1999) Pollen-related food allergy: cloning and immunological analysis of isoforms and mutants of Mal d 1, the major apple allergen, and Bet v 1, the major birch pollen allergen. European Journal of Nutrition 38: 201-215

Soriano JB, Antó JM, Sunyer J, Tobías A, Kogevinas M, Almar E, Muniozguren N, Sánchez JL, Palenciano L, Burney P (1999) Risk of asthma in the general Spanish population attributable to specific immunoresponse. International Journal of Epidemiology 28: 728-734

Spangfort MD, Mirza O, Ipsen H, van Neerven RJJ, Gajhede M, Larsen JN (2003) Dominating IgE-binding epitope of Bet v 1, the major allergen of birch pollen, characterized by X-ray crystallography and site-directed mutagenesis. Journal of Immunology 171: 3084-3090

Spellerberg IF, Eriksson NE, Crump VSA (2006) Silver birch (*Betula pendula*) pollen and human health: Problems for an exotic tree in New Zealand. Arboriculture and Urban Forestry 32: 133-137

Spieksma FTM, Corden JM, Detandt M, Millington WM, Nikkels H, Nolard N, Schoenmakers CHH, Wachter R, De Weger LA, Willems R, Emberlin J (2003) Quantitative trends in annual totals

of five common airborne pollen types (*Betula, Quercus*, Poaceae, *Urtica*, and *Artemisia*), at five pollen-monitoring stations in western Europe. Aerobiologia 19: 171-184

Srivastava S, Fristensky B, Kav NNV (2004) Constitutive expression of a PR10 protein enhances the germination of Brassica napus under saline conditions. Plant and Cell Physiology 45: 1320-1324

Stevens WJ, Ebo DG, Hagendorens MM, Bridts CH, De Clerck LS (2003) Is the prevalence of specific IgE to classical inhalant aeroallergens among patients with respiratory allergy changing? Evidence from two surveys 15 years apart. Acta Clinica Belgica 58: 178-182

Strachan DP (1989) Hay-fever, hygiene, and household size. British Medical Journal 299: 1259-1260

Sunyer J, Soriano J, Antó JM, Burgos F, Pereira A, Payo F, Martínez-Moratalla J, Ramos J (2000) Sensitization to individual allergens as risk factors for lower FEV1 in young adults. International Journal of Epidemiology 29: 125-130

Swoboda I, Dang TCH, Heberle-Bors E, Vicente O (1995a) Expression of Bet v 1, the major birch pollen allergen, during anther development - An in situ hybridization study. Protoplasma 187: 103-110

Swoboda I, Hoffmann-Sommergruber K, O'Riordáin G, Scheiner O, Heberle-Bors E, Vicente O (1996) Bet v 1 proteins, the major birch pollen allergens and members of a family of conserved pathogenesis-related proteins, show ribonuclease activity in vitro. Physiologia Plantarum 96: 433-438

Swoboda I, Jilek A, Ferreira F, Engel E, Hoffmann-Sommergruber K, Scheiner O, Kraft D, Breiteneder H, Pittenauer E, Schmid E, Vicente O, Heberle-Bors E, Ahorn H, Breitenbach M (1995b) Isoforms of Bet v 1, the major birch pollen allergen, analyzed by liquid-chromatography, mass-spectrometry, and cDNA cloning. Journal of Biological Chemistry 270: 2607-2613

Swoboda I, Scheiner O, Heberle-Bors E, Vicente O (1995c) cDNA cloning and characterization of three genes in the Bet v 1 gene family that encode pathogenesis-related proteins. Plant Cell and Environment 18: 865-874

Takagi H, Hiroi T, Yang LJ, Tada Y, Yuki Y, Takamura K, Ishimitsu R, Kawauchi H, Kiyono H, Takaiwa F (2005) A rice-based edible vaccine expressing multiple T cell epitopes induces oral tolerance for inhibition of Th2-mediated IgE responses. Proceedings of the National Academy of Sciences of the United States of America 102: 17525-17530

Taylor SL (2003) Safety Assessment of Foods Produced through Agricultural Biotechnology. Nutrition Reviews 61: S135-S140

Taylor SL, Hefle SL, Bindslev-Jensen C, Bock SA, Burks Jr AW, Christie L, Hill DJ, Host A, Hourihane JOB, Lack G, Metcalfe DD, Moneret-Vautrin DA, Vadas PA, Rance F, Skrypec DJ, Trautman TA, Yman IM, Zeiger RS (2002) Factors affecting the determination of threshold doses for allergenic foods: How much is too much? Journal of Allergy and Clinical Immunology 109: 24-30

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. Nucleic Acids Research 22: 4673-4680

Thórsson ÆT, Salmela E, Anamthawat-Jónsson K (2001) Morphological, cytogenetic, and molecular evidence for introgressive hybridization in birch. Journal of Heredity 92: 404-408

Titchener GD, Sapp SG (2002) A comparison of two approaches to understanding consumer opinions of biotechnology. Social Behavior and Personality 30: 373-381

Twardosz A, Hayek B, Seiberler S, Vangelista L, Elfman L, Grönlund H, Kraft D, Valenta R (1997) Molecular characterization, expression in *Escherichia coli*, and epitope analysis of a two EF-hand calcium-binding birch pollen allergen, Bet v 4. Biochemical and Biophysical Research Communications 239: 197-204

UCB (1997) European Allergy White Paper, Brussels

Ukaji N, Kuwabara C, Takezawa D, Arakawa K, Fujikawa S (2004) Accumulation of pathogenesis-related (PR) 10/Bet v 1 protein homologues in mulberry (*Morus bombycis* Koidz.) tree during winter. Plant Cell and Environment 27: 1112-1121

Utriainen M, Kokko H, Auriola S, Sarrazin O, Karenlampi S (1998) PR-10 protein is induced by copper stress in roots and leaves of a Cu/Zn tolerant clone of birch, *Betula pendula*. Plant Cell and Environment 21: 821-828

Valenta R, Duchene M, Breitenbach M, Pettenburger K, Koller L, Rumpold H, Scheiner O, Kraft D (1991) A low molecular weight allergen of white birch (*Betula verrucosa*) is highly homologous to human profilin. International Archives of Allergy and Applied Immunology 94: 368-370

Valenta R, Duchene M, Ebner C, Valent P, Sillaber C, Deviller P, Ferreira F, Tejkl M, Edelmann H, Kraft D, Scheiner O (1992) Profilins constitute a novel family of functional plant pan-allergens. Journal of Experimental Medicine 175: 377-385

Valenta R, Kraft D (1996) Type I allergic reactions to plant-derived food: A consequence of primary sensitization to pollen allergens. Journal of Allergy and Clinical Immunology 97: 893-895

Valenta R, Niederberger V (2007) Recombinant allergens for immunotherapy. Journal of Allergy and Clinical Immunology 119: 826-830

Valenta R, Vrtala S, Laffer S, Spitzauer S, Kraft D (1998) Allergy reviews series II: Recombinant allergens. Allergy: European Journal of Allergy and Clinical Immunology 53: 552-561

Valjakka M, Luomala EM, Kangasjärvi J, Vapaavuori E (1999) Expression of photosynthesis- and senescence-related genes during leaf development and senescence in silver birch (*Betula pendula*) seedlings. Physiologia Plantarum 106: 302-310

Van Kleef E, van Trijp HCM, Luning P (2005) Consumer research in the early stages of new product development: A critical review of methods and techniques. Food Quality and Preference 16: 181-201

Van Putten MC, Frewer LJ, Gilissen LJWJ, Gremmen B, Peijnenburg AACM, Wichers HJ (2006) Novel foods and food allergies: A review of the issues. Trends in Food Science and Technology 17: 289-299

Van Ree R, Akkerdaas JH, Van Leeuwen WA, Fernandez-Rivas M, Asero R, Knul-Brettlova V, Knulst A, Aalberse RC (2000) New perspectives for the diagnosis of food allergy. Allergy and Clinical Immunology International 12: 7-12

Van Ree R, Fernández-Rivas M, Cuevas M, Van Wijngaarden M, Aalberse RC (1995) Pollen-related allergy to peach and apple: An important role for profilin. Journal of Allergy and Clinical Immunology 95: 726-734

Van Ree R, Van Leeuwen WA, Akkerdaas JH, Aalberse RC (1999) How far can we simplify in vitro diagnostics for Fagales tree pollen allergy? A study with three whole pollen extracts and purified natural and recombinant allergens. Clinical And Experimental Allergy 29: 848-855

Vanek-Krebitz M, Hoffmann-Sommergruber K, Laimer da Camara-Machado M, Susani M, Ebner C, Kraft D, Scheiner O, Breiteneder H (1995) Cloning and sequencing of Mal d 1, the major allergen from apple (*Malus domestica*), and its immunological relationship to Bet v 1, the

major birch pollen allergen. Biochemical and Biophysical Research Communications 214: 538-551

Vartiainen E, Petäys T, Haahtela T, Jousilahti P, Pekkanen J (2002) Allergic diseases, skin prick test responses, and IgE levels in North Karelia, Finland, and the Republic of Karelia, Russia. Journal of Allergy and Clinical Immunology 109: 643-648

Vieths S, Scheurer S, Ballmer-Weber B (2002) Current understanding of cross-reactivity of food allergens and pollen. In: Genetically Engineered Foods Assessing Potential Allergenicity Annals of the New York Acedemy of Sciences 964: 47-68

Von Mutius E, Martinez FD, Fritzsch C, Nicolai T, Reitmeir P, Thiemann H-H (1994) Skin-test reactivity and number of siblings. British Medical Journal 308: 692-695

Von Mutius E, Weiland SK, Fritzsch C, Duhme H, Keil U (1998) Increasing prevalence of hay fever and atopy among children in Leipzig, East Germany. Lancet 351: 862-866

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: A new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414

Vriesendorp B, Bakker FT (2005) Reconstructing patterns of reticulate evolution in angiosperms: What can we do? Taxon 54: 593-604

Wagner S, Radauer C, Bublin M, Hoffmann-Sommergruber K, Kopp T, Greisenegger EK, Vogel L, Vieths S, Scheiner O, Breiteneder H (in press) Naturally occurring hypoallergenic Bet v 1 isoforms fail to induce IgE responses in individuals with birch pollen allergy. Journal of Allergy and Clinical Immunology

Walter MH, Liu JW, Wünn J, Hess D (1996) Bean ribonuclease-like pathogenesis-related protein genes (Ypr10) display complex patterns of developmental, dark-induced and exogenous-stimulus-dependent expression. European Journal of Biochemistry 239: 281-293

Wang CS, Huang JC, Hu JH (1999) Characterization of two subclasses of PR-10 transcripts in lily anthers and induction of their genes through separate signal transduction pathways. Plant Molecular Biology 40: 807-814

Wang GCY, Wang Y (1996) The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. Microbiology 142: 1107-1114

Wangorsch A, Ballmer-Weber BK, Rösch P, Holzhauser T, Vieths S (2007) Mutational epitope analysis and cross-reactivity of two isoforms of Api g 1, the major celery allergen. Molecular Immunology 44: 2518-2527

Weber RW (2003) Patterns of pollen cross-allergenicity. Journal of Allergy and Clinical Immunology 112: 229-239

Weiller GF (1998) Phylogenetic profiles: A graphical method for detecting genetic recombinations in homologous sequences. Molecular Biology and Evolution 15: 326-335

Wen J, Vanek-Krebitz M, Hoffmann-Sommergruber K, Scheiner O, Breiteneder H (1997) The potential of Betv1 homologues, a nuclear multigene family, as phylogenetic markers in flowering plants. Molecular Phylogenetics and Evolution 8: 317-333

Wensing M, Akkerdaas JH, van Leeuwen A, Stapel SO, Bruijnzeel-Koomen CAFM, Aalberse RC, Bast BJEG, Knulst AC, van Ree R (2002) IgE to Bet v 1 and profilin: Cross-reactivity patterns and clinical relevance. Journal of Allergy and Clinical Immunology 110: 435-442

Weryszko-Chmielewska E, Puc M, Rapiejko P (2001) Comparative analysis of pollen counts of *Corylus*, *Alnus* and *Betula* in Szczecin, Warsaw and Lublin (2000-2001). Annals of Agricultural and Environmental Medicine 8: 235-240

Westphal S, Kempf W, Foetisch K, Retzek M, Vieths S, Scheurer S (2004) Tomato profilin Lyc e 1: IgE cross-reactivity and allergenic potency. Allergy 59: 526-532

Whinnett A, Mundy NI (2003) Isolation of novel olfactory receptor genes in marmosets (Callithrix): insights into pseudogene formation and evidence for functional degeneracy in non-human primates. Gene 304: 87-96

Wiche R, Gubesch M, Konig H, Fotisch K, Hoffmann A, Wangorsch A, Scheurer S, Vieths S (2005) Molecular basis of pollen-related food allergy: identification of a second cross-reactive IgE epitope on Pru av 1, the major cherry (*Prunes avium*) allergen. Biochemical Journal 385: 319-327

Wickstrom G, Bendix T (2000) The "Hawthorne effect" - What did the original Hawthorne studies actually show? Scandinavian Journal of Work, Environment and Health 26: 363-367

Williams JH, Jr., Arnold ML (2001) Sources of genetic structure in the woody perennial *Betula occidentalis*. International Journal of Plant Sciences 162: 1097-1109

Winkler H (1904) Betulaceae. Das Pflanzenreich 19: 1-149

Wuthrich B, Schindler C, Leuenberger P, Ackermannliebrich U, Alean P, Blaser K, Bolognini G, Bongard JP, Brandli O, Braun P, Bron C, Brutsche M, Defila C, Domenighetti G, Elsasser S, Grize L, Guldimann P, Hufschmid P, Karrer W, Kellerwossidlo H, Keller R, Kunzli N, Luthy JC, Martin BW, Medici T, Monn C, Peeters AG, Perruchoud AP, Radaelli A, Schwartz J, Solari G, Schoni M, Tschopp JM, Villiger B, Zellweger JP, Zemp E (1995) Prevalence of atopy and pollinosis in the adult-population of Switzerland (SAPALDIA study). International Archives of Allergy and Immunology 106: 149-156

Yazdanbakhsh M, Kremsner PG, Van Ree R (2002) Allergy, parasites, and the hygiene hypothesis. Science 296: 490-494

Zechendorf B (1994) What the public thinks about biotechnology - Better than synthetic food but worse than organ-transplantation - a survey of opinion polls. Biotechnology 12: 870-875

Zeiger RS (2003) Food allergen avoidance in the prevention of food allergy in infants and children. Pediatrics 111: 1662-1671

Zylstra P, Rothenfluh HS, Weiller GF, Blanden RV, Steele EJ (1998) PCR amplification of murine immunoglobulin germline V genes: Strategies for minimization of recombination artefacts. Immunology and Cell Biology 76: 395-405

SUMMARY

Allergic diseases, such as hay fever and food allergy, affect a substantial part of the population in westernized countries. Pollen of the European white birch (*Betula pendula*) is a considerable cause of hay fever (seasonal allergic rhinitis) in northern and central Europe. The major birch pollen allergen is Bet v 1, which is the conventional allergen name for the birch pollen proteins of a large group of proteins otherwise known as PR-10 proteins. Individuals that suffer from birch pollen allergy are particularly prone to develop Oral Allergy Syndrome (OAS) due to the occurrence of an IgE-mediated cross-reaction between Bet v 1 and PR-10 proteins in various plant foods. The high prevalence of apple allergy among hay fever sufferers is a good example of OAS. Allergic diseases can be severe for patients and have considerable medical and economical costs as well. Consequently, prevention of birch pollen allergy and OAS would contribute to an improvement of the quality of life of many patients. This thesis examines the feasibility of strategies that are directed towards development of hypoallergenic (= having a reduced allergenicity) birch trees or plant foods in order to alleviate Bet v 1 [PR-10]- associated complaints.

When the development of hypoallergenic products involves application of genetic modification (GM), societal concerns about this technology should be taken into account. The attitude towards hypoallergenic products developed by GM was examined in the survey that is described in *Chapter 2*. This study differentiated between patients and non-patients. Attitude towards GM was measured for two applications directed against hay fever (hypoallergenic birch and grass) and one application directed against food allergy (hypoallergenic apple). Attitude was described in terms of two constructs that were labelled as 'benefits' and 'rejection factors'. Hay fever sufferers perceived greater 'benefits' associated with application of GM to develop hypoallergenic birch trees as compared to non-sufferers. The perceived 'benefits' increased with a higher self-reported impact of hay fever on quality of life. No attitudinal differences were observed between sufferers and non-sufferers for the attitudinal construct 'rejection factors'. Furthermore, the impact of perceived 'benefits' on acceptance of GM was larger than the impact of 'rejection factors'.

The perceptions of 'benefits' were further explored in *Chapter 3*. Here, the attitude towards hypoallergenic apples was examined in a survey in which consumers rate a set of apple profiles that varied in the breeding method that was applied during development (GM *vs.* traditional breeding), in pesticide usage, and in the degree of allergenicity. Acceptance of hypoallergenic apples was high among all consumers, also when GM was involved in their development. Acceptance of hypoallergenic products was higher among consumers with an apple allergy, presumably because they find a personal 'benefit' associated with these products. Novel GM products that are recognised as beneficial by some consumers may consequently experience an increased acceptance. However, both Chapter 2 and 3 indicated a clear consumer preference for

traditional breeding over breeding by GM for the development of hypoallergenic birch trees or food products. Approaches which focus on traditional breeding should thus be explored first.

Chapter 3 also evaluated appreciation of the hypoallergenic apple cultivar Santana, which was introduced in shops in a large-scale sales pilot labelled as 'suitable for individuals suffering from mild apple allergy' in 2006. A survey among consumers that bought the Santana measured the self-reported response to this apple. Almost half (42%) of the apple allergic consumers had no allergic reaction at all after eating the Santana. Most (96%) consumers who did experience an allergic reaction reported the symptoms as minor. The self-reported severity of the apple allergy, the occurrence of other fruit allergies, and age were associated with the occurrence of an allergic reaction to the Santana. Overall, the Santana was valued positively by the majority of apple allergic consumers, regardless of whether these consumers could eat the apple without experiencing an allergic reaction. Development of hypoallergenic foods such as Santana may therefore contribute to food allergy management, although variation among individual consumers in the allergic response to hypoallergenic products should be taken into account when developing and marketing such products.

A prerequisite for designing strategies for selection and breeding of hypoallergenic birch trees is knowledge on diversity of Bet v 1 genes and on allergenicity of the protein variants (=isoforms) that are encoded by these genes. *Chapter 4* describes a study on the variation of Bet v 1 isoforms in the most common birch species in Europe, *Betula pendula*. *PR-10* sequences from three *B. pendula* cultivars were amplified, cloned and sequenced. Forty-four unique *PR-10* sequences were recovered from *B. pendula* and these were assigned to thirteen putative genes based on sequence identity and intron length. Information on gene transcription was inferred from a comparison with existing mRNA sequences and suggested that seven of these genes are transcribed in pollen. Bet v 1 isoforms are known to vary in IgE-reactivity. The study in Chapter 4 showed that isoforms with high and low IgE-reactivity are encoded by different genes. Thus, one birch pollen grain has the genetic background to produce a mixture of isoforms with varying IgE-reactivity. The sequence of one of the isoforms with a high IgE-reactivity was present in all examined *B. pendula* trees. The search for hypoallergenic birch trees was, therefore, expanded to other *Betula* species.

The selection of a representative set of species which cover variation among *Betula* species requires knowledge on phylogenetic relationships within the genus *Betula*. *Chapter 5* describes a study on the reconstruction of the phylogeny of this genus using multilocus data from AFLP markers. The taxonomy of the genus *Betula* is complicated by the occurrence of parallel evolution of morphological traits, of polyploidisation events, and of extensive hybridisation among species. A large number of polymorphic AFLP markers (321 variable bands) were produced in 107 *Betula* accessions from 23 species and 11 hybrids. The analysis identified four distinct groups within the

genus. These groups are partly in disagreement with the traditional, but disputed, division of the genus. The majority of the species and all hybrids fell within subgenus *Betula* and are thus closely related to *B. pendula*. Subgenus *Chamaebetula* and part of the *Neurobetula* species should be merged with subgenus *Betula*. Apart from subgenus *Betula*, the subgenera *Betulenta*, *Betulaster*, and the remaining part of *Neurobetula* formed distinct and well-supported groups and should thus be maintained. The results from the AFLP study are to a large extent congruent with results from previous studies that made use of molecular (sequence) data.

The allergenic potency of *Betula* species other than *B. pendula* is described in *Chapter 6* in which the *PR-10* genes from eight birch species are cloned and sequenced. These species represent the various groups that were previously identified in the genus *Betula*. In total, 134 unique PR-10 sequences were recovered, including both sequences with a full open reading frame and pseudogenes. Sequences were attributed to putative genes, which could, in turn, be subdivided into seven subfamilies. Five subfamilies were common to all birch species. Q-TOF LC-MS^E was applied to detect peptide fragments of Bet v 1 that are unique for particular isoforms, in order to identify which PR-10 genes are expressed in pollen. The relative abundance of individual isoforms in the pollen proteome was also determined by Q-TOF LC-MS^E. Each of the five examined birch species expressed a mixture of isoforms with at least 4-5 different isoforms. Both isoforms with a high and low IgE-reactivity were abundant in the Bet v 1 mixture of *B. pendula*. The other birch species lacked Bet v 1 isoforms that are similar to known isoforms with a low IgE-reactivity, but isoforms that are similar to known isoforms with a high IgE-reactivity were abundant in all species except *B. lenta*.

In *Chapter 7*, the antigenic and allergenic profiles of pollen extracts from twenty-four different birch trees were determined by SDS-PAGE and Western blotting. Fifteen different *Betula* species were examined, covering all previously identified subgenera/groups in the genus *Betula*. The major birch allergen Bet v 1 was an abundant protein in all examined pollen extracts. Immune-reactivity of the extracts was tested using a pool of sera that were obtained from birch pollen allergic patients. A strong 17 kDa band, representing Bet v 1, was recognized by the serum pool in all pollen extracts. The degree of immune-reactivity correlated well with the total amount of Bet v 1 in the extract, which varied from 44% to 61% of the total protein content. Pollen extracts from different birch trees varied in the total protein content, presumably the result of variation in pollen quality due to pollen rupture and varying hydration during the extraction. Bet v 1 isoforms in the pollen of eleven *Betula* species were subsequently digested with trypsin and the resulting fragments were analyzed and quantified by Q-TOF LC-MS^E. Peptides that contained amino acid residues that are associated with high IgE-reactivity were detected in all examined species, and were abundant as well. Differences between *Betula* species in the relative presence of these amino acid residues were small. As a consequence, differences in allergenicity between birch trees

are probably far too small to have clinical relevance, implying that all examined *Betula* species will be highly allergenic.

The research presented here did not identify any birch trees in which Bet v 1 variants with a high IgE-reactivity are reduced in abundance or are absent. Development of hypoallergenic birch trees thus must rely on other approaches, such as selection or introduction of male sterility (trees that do not produce male catkins were observed during the study), or the application of RNAi to silence the Bet v 1 genes in pollen. Also, the conducted research showed that acceptance of hypoallergenic GM products is quite high, particularly for allergic sufferers who were allergic to the product being modified. This supports the hypothesis that consumer attitude towards GM is partly driven by recognition of specific and personally relevant benefits. The hypoallergenic apple cultivar Santana, which was developed by selection and not by GM, was well received by allergic consumers, indicating that there is a market for hypoallergenic products.

SAMENVATTING

Een aanzienlijk deel van de bevolking in de westerse samenleving lijdt aan allergische aandoeningen als hooikoorts en voedselallergie. Eén van de belangrijkste veroorzakers van hooikoorts in het noordelijke en centrale deel van Europa is het stuifmeel (pollen) van de ruwe berk (*Betula pendula*). De meeste patiënten reageren op een eiwit uit de PR-10 familie, Bet v 1 genaamd. Veel mensen met een allergie tegen berkenstuifmeel ontwikkelen ook een voedselallergie. Dit komt doordat het Bet v 1 eiwit veel lijkt op eiwitten in bepaalde voedingsmiddelen die daardoor ook als allergeen herkend worden door het immuunsysteem. Dit verschijnsel wordt kruisreactiviteit genoemd. Een goed voorbeeld hiervan is het veelvuldig voorkomen van appelallergie onder hooikoortspatiënten met aan allergie tegen berkenstuifmeel. Het type voedselallergie dat hiermee gepaard gaat, wordt aangeduid als het Orale Allergie Syndroom en treedt vooral op in de mond, neus en keel met verschijnselen als jeuk en zwellingen. Mensen die hier last van hebben, mijden deze voedingsproducten, waardoor belangrijke componenten van een gezonde voeding ontbreken. Allergische aandoeningen hebben een negatief effect op de kwaliteit van leven en zijn daarom een relevant doelwit voor onderzoek naar preventie. Preventie in dit onderzoekskader betekent het ontwikkelen van berken en voedingsproducten met een verminderde allergeniciteit (=hypoallergeen), om zo allergische klachten te kunnen verlichten. De belangrijkste strategieën die in dit proefschrift aan de orde komen, berusten op selectie van natuurlijk uitgangsmateriaal en op genetische modificatie (GM).

Indien bij de ontwikkeling van hypoallergene producten GM wordt toegepast, moet men rekening houden met maatschappelijke zorgen rondom deze technologie. *Hoofdstuk 2* onderzoekt de houding van consumenten ten opzichte van hypoallergene producten die ontwikkeld worden met GM. Hiervoor is een enquête opgesteld, waarin mensen zelf aangeven of ze al of niet allergisch zijn. De enquête onderzoekt vervolgens acceptatie van twee toepassingen gericht op hooikoorts (hypoallergene berken en grassen) en één toepassing gericht op voedselallergie (hypoallergene appels). De houding ten opzichte van GM kan herleid worden tot twee zogenaamde constructen (groepen van antwoorden) die vereenvoudigd aangeduid worden als “positieve aspecten” (zoals nuttig, noodzakelijk en belangrijk) en “negatieve aspecten” (zoals riskant, onnatuurlijk en onethisch). Hooikoortspatiënten verbinden meer “positieve aspecten” aan de toepassing van GM om hypoallergene berken te ontwikkelen in vergelijking met niet-patiënten. De veronderstelde “positieve aspecten” zijn groter voor naarmate mensen meer last ondervinden van de hooikoortsklachten. Met betrekking tot de “negatieve aspecten” zijn er geen verschil tussen patiënten en niet-patiënten waargenomen. De “positieve aspecten” en de “negatieve aspecten” bepalen samen de houding en dus acceptatie of verwerping van GM. Het relatieve effect van veronderstelde “positieve aspecten” op de acceptatie van GM groter is in dit onderzoek groter dan het effect van ”negatieve aspecten”.

De perceptie van de “positieve aspecten” van toepassing van GM voedingsproducten is nader onderzocht in *Hoofdstuk 3*. In een tweede enquête is aan consumenten een aantal mogelijk te ontwikkelen appels voorgelegd ter beoordeling. Deze appels verschillen in: (1) de toegepaste veredelingsmethode (GM tegenover traditionele veredeling), (2) het gebruik van pesticiden en (3) de mate van allergeniciteit. Een belangrijk resultaat van dit onderzoek is dat de acceptatie van hypoallergene appels in het algemeen hoog is onder alle consumenten, ook wanneer GM is gebruikt bij de ontwikkeling ervan. Daarbij is acceptatie significant het hoogst is bij consumenten die last hebben van appelallergie. Waarschijnlijk is het directe (persoonlijke) voordeel wat zij hebben een doorslaggevende factor. Echter, zowel *Hoofdstuk 2* als *3* laten een duidelijke consumentenvoorkeur zien voor het gebruik van traditionele veredeling boven toepassing van GM bij de ontwikkeling van hypoallergene producten. Daarom dient de haalbaarheid van benaderingen die gebruik maken van traditionele verdeling het eerst onderzocht te worden.

Hoofdstuk 3 evalueert tevens de waardering van een hypoallergene appel die momenteel al beschikbaar is. Het appelras Santana is op grote schaal geïntroduceerd door een Nederlandse supermarktketen en gelabeld verpakt als “geschikt voor mensen met een milde appelallergie”. Daarnaast is dit ras verkocht in diverse winkels zonder aparte aanduiding. Op grond van een vragenlijst onder kopers van Santana blijkt dat het optreden van een allergische reactie op Santana bij individuele consumenten samenhangt met de (zelfgeregistreerde) ernst van de appelallergie, het hebben van andere fruitallergieën en de leeftijd van de consument. Bijna de helft van de consumenten met appelallergie krijgt geen allergische reactie na het eten van Santana. Bij vrijwel alle consumenten die wel een allergische reactie krijgen, zijn de symptomen gering. De Santana wordt positief gewaardeerd door de meerderheid van de appelallergische consumenten, ongeacht het optreden van een allergische reactie. Dit onderzoek toont aan dat de ontwikkeling van hypoallergene voedingsproducten die verkregen zijn via selectie en veredeling alsook via GM een belangrijke en door de consument geaccepteerde bijdrage kan leveren aan de beheersing van de allergieproblematiek. Daarbij moet men echter rekening houden met verschillen in het al dan niet optreden van een allergische reactie op dergelijke producten tussen individuele consumenten

Hoofdstuk 4 onderzoekt de diversiteit van Bet v 1 eiwitvarianten binnen de meest algemene berkensoort van Europa, namelijk *B. pendula*. Deze kennis is onontbeerlijk voor de mogelijke selectie en veredeling van hypoallergene berken. De genen die voor Bet v 1 coderen behoren tot de PR-10 genfamilie. Om vast te stellen hoeveel PR-10 genen aanwezig zijn in een individuele boom, zijn van elk van drie *B. pendula* cultivars de PR-10 genen geïsoleerd en gekloneerd om de DNA sequentie ervan te kunnen bepalen. In totaal zijn 44 unieke PR-10 sequenties geïsoleerd uit de drie cultivars. Op basis van overeenkomsten tussen de sequenties en de lengte van de

aanwezige introns moeten deze sequenties afkomstig te zijn van tenminste 13 genen. Transcripten van zeven van deze genen zijn aangetoond in mRNA van stuifmeel. Het was al bekend dat Bet v 1 eiwitvarianten kunnen verschillen in allergeniciteit, wat tot uitdrukking komt in variatie in de reactie met IgE antilichamen van hooikoortspatiënten. Het huidige onderzoek toont aan dat eiwitvarianten met een hoge en lage IgE-reactiviteit afkomstig zijn van afzonderlijke genen. Eén stuifmeelkorrel van *B. pendula* zal dus een mengsel bevatten van eiwitvarianten met variatie in de mate van IgE-reactiviteit. De DNA sequentie van een eiwitvariant met een hoge IgE-reactiviteit blijkt echter aanwezig te zijn in alle onderzochte *B. pendula* bomen. Daarom is het onderzoek uitgebreid met een zoektocht naar hypoallergene berken bij andere *Betula* soorten.

Kennis over de verwantschap tussen soorten binnen het genus *Betula* maakt het mogelijk om een collectie van soorten te maken die de aanwezige variatie binnen het genus goed representeert. In *Hoofdstuk 5* reconstrueren we de fylogenie van het genus *Betula* met behulp van AFLP markers. De taxonomie van dit genus is controversieel en de bepaling ervan wordt bemoeilijkt door het optreden van parallelle evolutie van morfologische kenmerken, polyploidisatie en het veelvuldige optreden van hybridisatie tussen berkensoorten. Moleculaire merkers (zoals AFLPs) kunnen hierbij duidelijkheid verschaffen. AFLP onderzoek aan 107 *Betula* accessies, afkomstig van 23 soorten en 11 hybriden, heeft een groot aantal polymorfe merkers (321 variabele banden) opgeleverd. Met behulp hiervan zijn binnen het genus *Betula* nu vier afzonderlijke groepen onderscheiden, waarmee enkele taxonomische problemen zijn opgehelderd. De meerderheid van de soorten en alle hybriden kunnen tot het subgenus *Betula* gerekend worden en zijn sterk verwant aan *B. pendula*. Het subgenus *Chamaebetula* en een deel van de *Neurobetula* soorten blijken nu ook thuis te horen in het subgenus *Betula*. Naast het subgenus *Betula* vormen de subgenera *Betulenta*, *Betulaster* en het resterende deel van de *Neurobetula* ieder een op basis van de AFLP merkers goed ondersteunde groep. De resultaten van het AFLP onderzoek zijn in goede overeenstemming met eerdere studies die gebruik maakten van moleculaire (sequentie) data.

De allergeniciteit van andere berkensoorten naast *B. pendula* is onderzocht in *Hoofdstuk 6*. Van acht soorten, die op basis van de fylogenetische analyse representatief zijn voor de subgenera, zijn de DNA sequenties van de PR-10 genen bepaald. Dit heeft in totaal 134 unieke sequenties opgeleverd waaronder zowel sequenties zijn die voor een volledig eiwit codeerden als ook zogeheten pseudogenen. Alle sequenties zijn toegewezen aan afzonderlijke genen en op basis van de mate van onderlinge overeenkomst onderverdeeld in zeven subgroepen. Vijf daarvan zijn gemeenschappelijk voor alle onderzochte berkensoorten. Vervolgens zijn de Bet v 1 eiwitten uit het stuifmeel (na behandeling met trypsine om de eiwitten te fragmenteren) geanalyseerd en gekwantificeerd met Q-TOF LC-MS^E. Deze analyse laat zien dat de genen van twee subgroepen tot expressie komen in het stuifmeel en dat in stuifmeel van elk van de onderzochte berkensoorten een mengsel aanwezig is van tenminste 4 of 5 eiwitvarianten. Met de Q-TOF LC-MS^E is tevens de

relatieve hoeveelheid van de afzonderlijke eiwitvarianten in het stuifmeel bepaald. Zo is aangetoond dat het Bet v 1 mengsel in stuifmeel van *B. pendula* grote hoeveelheden bevat van zowel eiwitvarianten met een hoge als met een lage IgE-reactiviteit. Eiwitvarianten die lijken op varianten met een lage IgE-reactiviteit ontbreken in de andere soorten, terwijl varianten met een hoge IgE-reactiviteit overvloedig voorkomen in alle soorten met uitzondering van *B. lenta*.

Hoofdstuk 7 beschrijft de karakterisering met SDS-PAGE en Western blotting van stuifmeelextracten afkomstig van 15 berkensoorten die alle subgenera uit de fylogenetische analyse vertegenwoordigen. De SDS-PAGE laat in alle onderzochte extracten een zeer sterke band van 17 kDa zien, wat overeenkomt met de bekende eiwitmassa van Bet v 1. De immunoreactiviteit van de extracten is onderzocht met behulp van sera afkomstig van patiënten met berkenallergie. De sera reageren in alle extracten met de band van 17 kDa. De sterkte van de immunoreactiviteit blijkt gecorreleerd aan de totale hoeveelheid Bet v 1 in het extract. Het totale eiwitgehalte van de afzonderlijke extracten is variabel, waarschijnlijk ten gevolge van verschillen in stuifmeelkwaliteit tussen de diverse soorten door openbarsten van sommige stuifmeelkorrels en door verschillen in watergehalte. De relatieve hoeveelheid Bet v 1 in de verschillende extracten is echter behoorlijk constant en varieert van 44 tot 61% van de totale eiwithoeveelheid. Met Q-TOF LC-MS^E zijn in alle onderzochte soorten eiwitfragmenten aangetoond waarin op specifieke posities aminozuren voorkomen die geassocieerd zijn met een hoge IgE-reactiviteit. Hierbij zijn slechts kleine verschillen waargenomen in het relatieve voorkomen van de met een hoge IgE-reactiviteit geassocieerde aminozuren tussen berkensoorten. Op grond hiervan wordt geconcludeerd dat eventuele verschillen in allergeniciteit waarschijnlijk niet klinisch relevant zijn, wat impliceert dat alle onderzochte berkensoorten een hoge allergeniciteit hebben. Dit onderzoek laat zien dat waar blootstelling aan berkenstuifmeel plaatsvindt, dit stuifmeel een vergelijkbaar mengsel aan Bet v 1 varianten bevat. Dit impliceert dat het sensibilisatie profiel van allergische patiënten die reageren op Bet v 1 wereldwijd grote overeenkomsten zal hebben.

Samenvattend heeft het onderhavige onderzoek geen hypoallergene berkensoorten of cultivars geïdentificeerd die stuifmeel maken waarin eiwitvarianten van Bet v 1 met een hoge IgE-reactiviteit niet of weinig voorkomen. Voor de productie van hypoallergene berken moeten dus andere strategieën gevolgd worden, zoals selectie of ontwikkeling (via GM) van mannelijke steriele berken, of de toepassing van RNAi om expressie van Bet v 1 genen in het stuifmeel uit te schakelen. In de eerste hoofdstukken is aangetoond dat de maatschappelijke acceptatie van GM bij de productie van hypoallergene producten relatief hoog is, vooral onder allergische patiënten die een persoonlijk voordeel hebben bij hypoallergene producten. De positieve reacties van appelallergische consumenten op de hypoallergene appelcultivar Santana - die via selectie en niet met GM is ontwikkeld - tonen bovendien aan dat er daadwerkelijk een markt is voor hypoallergene producten.

DANKWOORD

Wat in 2003 nog ver leek, is inmiddels dan bijna afgesloten; een AIO project van vier jaar gaat voorbij voordat je er goed en wel erg in hebt. Dat deze vier jaar zijn omgevlogen komt deels doordat het onderwerp van het onderzoek me na aan het hart lag, ik had immers zelf ook baat bij onderzoek aan berkenallergie en de bijbehorende fruitallergieën, maar komt ook door de uitstekende begeleiding en de prettige werkomgeving zoals ik die tijdens het onderzoek ervaren heb. Dank dus voor allen die hebben bijgedragen aan het onderzoek, die met ideeën zijn gekomen of hebben gezorgd voor de nodige mentale ondersteuning. Ik ga hieronder een aantal mensen noemen, maar heb niet de illusie dat deze lijst compleet is, dus voor een ieder die ik vergeten ben, bij deze mijn excuses en dank.

Natuurlijk ben ik in de eerste plaats dank verschuldigd aan mijn directe begeleiders. Allereerst Lynn Frewer voor het wegwijs maken in de sociale wetenschappen en de prettige samenwerking op dat terrein. Evert Jacobsen voor het overzien van de grote lijnen en de hulp bij de naderende deadline aan het eind van het project. Voor de dagelijkse begeleiding ben ik veel dank verschuldigd aan Luud Gilissen en René Smulders, van wie ik zeer veel heb kunnen leren en met wie ik een zeer prettige samenwerking heb gehad. Doordat ik bij hen altijd terecht kon voor vragen en het oplossen van knelpunten heb ik dit project ook daadwerkelijk in vier jaar kunnen afronden.

Naast de begeleidingscommissie wil ik drie personen in het bijzonder noemen. Dat zijn mijn kamergenoot Paul Arens, die vaak als eerste aanspreekpunt diende bij problemen of vragen. Ik hoop dat ik ondanks mijn vervuiling van de kamer met stapels papier, berkenstuifmeel en schimmelsporen uit vergeten koffiebekers toch een goede kamergenoot ben geweest. Verder Arnout Fischer, die een zeer grote bijdrage heeft geleverd aan het sociale deel van het project en mijn kennis van statistiek flink heeft bijgespijkerd. En dan Jan Cordewener, zonder wie de analyse van de Bet v 1 eiwitten niet mogelijk was geweest en met wie ik het erg prettig vond om mee samen te werken.

Dan zijn er een heleboel mensen die een bijdrage hebben geleverd aan het tot stand komen van de verschillende hoofdstukken. Zoals Dirk Visser, Simon Ribot en Elisa Börger bij het verzamelen en verwerken van het stuifmeel. Jelle Hiemstra bij het verzamelen van bladmateriaal. Claire-Noëlle Thienpoint voor de uitvoering van de AFLP studie en Wim Koopman voor de begeleiding daarvan. Voor de ondersteuning tijdens heb labwerk ben ik Wendy van ’t Westende, Danny Esselink, Martijn van Kaauwen, Hanneke van der Schoot, Yolanda Noordijk dank verschuldigd. Doordat de labs halverwege mijn AIO periode verhuisden en ik toen druk bezig was aan het enquêtewerk en dus pas ver na de verhuizing weer het lab indook, hebben zij mij zelfs twee keer wegwijs moeten maken in het lab. Mojgan Taskindoust bedank ik voor al het labwerk wat zij aan de PR-10 sequenties heeft gedaan. Twan America voor de hulp bij de analyse van de Q-TOF data en Jan-Willem Oomkes voor de hulp bij het blotten. Tenslotte Mirko Bollen en Harry Wichers voor het regelmatige overleg over PR-10 eiwitten en het tonen van de grafisch bijzonder fraaie, zeer interessante en vaak ook onbegrijpelijke plaatjes van eiwitstructuren die Mirko tijdens dit overleg meenam.

Één van de leukste dingen tijdens een AIO project is het hebben van collega's die in het zelfde schuitje zitten en net als jij aan het worstelen zijn met proeven, analyses en het opschrijven van de resultaten. Het is fijn dat je daar samen over kan klagen (en af-en-toe vertellen dat het zo goed gaat), maar om ook eens over iets anders te kunnen praten. Daarom wil ik mijn PRI-collega-AIO's Mirjam Jacobs, Eveline Stilma, Colette Broekgaarden en Marleen Cobben bedanken voor de dagelijkse gesprekken en het fijne gezelschap tijdens de lunchpauzes. Mijn ACW-collega-AIO's wil ik bedanken voor de leuke borrels in de Vlaam, dus AIO 0 (Evelien van Boxtel), AIO 2 (Mirko Bollen), AIO 3 (Prescilla Jeurink), AIO 4 (Teun van Herpen) en AIO 5 (Margreet van Putten) en Yvonne Vissers, aan wie bij mijn weten nog geen nummer is toegekend. Mirko, Prescilla en Yvonne, mocht het nodig zijn om jullie onderzoek van vers bloed te voorzien, weten jullie me vast wel te vinden. Tijdens de werkdag speelt tenslotte de koffiepauze een belangrijke rol en voor het aangename gezelschap tijdens deze pauzes wil ik Ate van der Burgt bedanken en hem bij deze veel succes toewensen bij zijn verdere werk- en atletiekcarrière.

De overige collega's wil ik bedanken voor de gezelligheid, behulpzaamheid en interesse tijdens het werk, dus Clemens, Gerard, Erik, Henk, Linda, Gerda en Gerda, Gao, Miqia, Alexander, Bjorn en Björn, Adriana, Brigitte, Jeroen en de Franse aio's bedankt (en nee, deze lijst is vast niet compleet). Alhoewel ik misschien te weinig mijn gezicht heb laten zien op de vakgroep Marktkunde en Consumentengedrag en er dus wat betreft de bèta-gamma interactie nog winst te behalen valt, wil ik Timo, Heleen, Janneke, Vladimir en Amber en de anderen bij deze bedanken.

Gelukkig bestaat er naast het werk aan de universiteit ook zoiets als vrije tijd. Er zijn veel vrienden en familieleden die het onderzoek de afgelopen jaren met belangstelling hebben gevolgd, waarvoor mijn dank. Een groot deel van mijn vrije tijd heb ik rondgebracht bij de studentenatletiekvereniging Tartlétos, waar ik al vanaf de studietijd trainde. Hardlopen is wat mij betreft een ideale bezigheid om de dagelijkse stress van je af te zetten, maar ik ben er niet de persoon na om in m'n eentje te gaan rennen, dus daarom wil ik hierbij alle trainingsmaatjes bedanken. In het bijzonder wil ik nog de gezellige "La Citta"-avonden met Mart, Ate, Mark en Olaf noemen. Helaas is de frequentie van deze avondjes steeds verder afgenomen, maar het is een traditie die in ere gehouden moet worden. Toen ik aan het AIO project begon, woonde ik nog op Droevendaal in de Tartlétosbarak "de Trimsalon" en ik heb het daar erg leuk gehad met mijn toenmalige huisgenoten Jochem, Ronald, Co, Allard, Mart, Mieke, Bart en Irene. Eveneens uit de studententijd stamt het contact met de studiegenoten van de biologie-opleiding. Ik ben blij dat we elkaar nog met regelmaat treffen op (verjaardags)feesten, mannenavonden en het jaarlijkse "biologenweekendje".

De laatste alinea van dit proefschrift zou ik graag willen reserveren voor het bedanken van mijn familie. In de eerste plaats Anneke en Bjørne, mijn ouders, die altijd zoveel voor mij gedaan hebben en mij gesteund en gemotiveerd hebben. Mijn broer(tje)s Ronald en Stefan en mijn oma wil ik bij deze bedanken voor hun interesse in mijn werk van de afgelopen jaren. Tenslotte, Fabianne bedankt voor al jouw steun en liefde en ik kijk erg uit naar onze toekomst samen!

Martijn

CURRICULUM VITAE

Martijn Schenk is op 13 oktober 1978 geboren in Bennekom en groeide op in de Betuwe, in het dorpje Ochten. Na het doorlopen van het VWO aan het Christelijk Lyceum in Veenendaal is hij in 1997 begonnen aan de studie Biologie aan de Wageningen Universiteit. In 1998 haalde hij *cum laude* zijn propedeuse en begon binnen de doctoraal opleiding Biologie aan de specialisatie populatiebiologie. Zijn eerste afstudeervak deed hij aan de vakgroep Entomologie en vond deels aan de Universiteit in Riverside (California, USA) plaats. Dit onderzoek draaide om de populatiestructuur van de sluipwesp *Trichogramma kaykai*. Het tweede afstudeervak vond plaats aan de vakgroep Biosystamatiek. In deze studie werd gekeken naar de invloed van natuurontwikkeling op biodiversiteit onder loopkevers. Als stage heeft hij enkele maanden doorgebracht in het Afrikaanse Gabon in het Parc National de la Lopé voor dierecologisch onderzoek aan habitatgebruik van bosbuffels. Tijdens de studie is hij actief geweest binnen de studievereniging BVW Biologica en de studentenatletiekvereniging Tartlétos. In augustus 2003 studeerde hij *cum laude* af aan wat inmiddels de MSc opleiding biologie was geworden. In november 2003 begon hij als AIO bij Plant Research International (WUR) aan een onderzoek waarbij zowel het Laboratorium voor Plantenveredeling als de leerstoelgroep Marktkunde en Consumentengedrag betrokken waren. Dit onderzoek richtte zich op de ontwikkeling van berken en voedingsproducten met een verminderde allergeniciteit. De resultaten van dit onderzoek zijn in dit proefschrift beschreven.

Education Statement of the Graduate School

Experimental Plant Sciences

Issued to: Martijn Frido Schenk
Date: 19 February 2008
Group: Plant Breeding & Marketing and Consumer Behaviour,
Wageningen University and Research Centre

1) Start-up phase	*date*
▶ **First presentation of your project**	
Allergens in the green environment	2004
▶ **Writing or rewriting a project proposal**	
▶ **Writing a review or book chapter**	
Strategies for mitigation and prevention of hay fever (IN: Allergy Matters, Wageningen UR Frontis Series, Vol 20,	2004
Consumer, communication and food allergy (IN: Understanding Consumers of Food Products, Woodhead	2006
▶ **MSc courses**	
Food related Allergies and Intolerances (FCH-90306)	2005
▶ **Laboratory use of isotopes**	
Subtotal Start-up Phase	*13.5 credits**

2) Scientific Exposure	*date*
▶ **EPS PhD student days**	
EPS student day, Vrije Universiteit Amsterdam	Jun 03, 2004
EPS student day, Wageningen University	Sep 19, 2006
EPS student day, Wageningen University	Sep 13, 2007
▶ **EPS theme symposia**	
EPS theme 4 'Genome Plasticity', Radboud University, Nijmegen	Dec 08, 2006
EPS theme 4 'Genome Plasticity', Leiden University	Dec 07, 2007
▶ **NWO Lunteren days and other National Platforms**	
▶ **Seminars (series), workshops and symposia**	
ACW Seminar series, Allergy Consortium Wageningen	2004
EPS Flying seminars	2006-2007
Key-gene seminar	Dec 05, 2006
▶ **Seminar plus**	
▶ **International symposia and congresses**	
Allergy matters! 2004, Wageningen (NL)	Feb 04-06, 2004
1st International Symposium on Molecular Allergology, Roma (Italy)	Mar 31-Apr 01, 2006
▶ **Presentations**	
Oral: Münchner Zentrum Allergie und Umwelt, Munich	2004
Oral: LEI, Den Haag	2005
Oral: 1st International Symposium on Molecular Allergology, Rome (Italy)	Mar 31-Apr 01, 2006
Poster: FAV Health 2007, Texas (USA)	Oct 09-Oct 13 2007
▶ **IAB interview**	Sep 18, 2006
▶ **Excursions**	
Subtotal Scientific Exposure	*7.3 credits**

3) In-Depth Studies	*date*
▶ **EPS courses or other PhD courses**	
Bioinformation Technology-1 (VLAG)	Nov 08-16, 2004
The future Genomics society (VLAG, Mansholt)	Feb 14-25, 2005
Molecular phylogenies: reconstruction and interpretation (EPS)	Oct 17-21, 2005
▶ **Journal club**	
Literature discussions Plant Breeding	2006-2007
▶ **Individual research training**	
Subtotal In-Depth Studies	*8.6 credits**

4) Personal development	*date*
▶ **Skill training courses**	
Scientific Writing (CENTA)	2007
▶ **Organisation of PhD students day, course or conference**	
▶ **Membership of Board, Committee or PhD council**	
Subtotal Personal Development	*1.8 credits**

TOTAL NUMBER OF CREDIT POINTS*	31.2

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 credits

** A credit represents a normative study load of 28 hours of study*