Applied Genomics - An Innovative Tool to Improve Quality in Chains

Predicting mealiness in apples – a case study

Monique van Wordragen[1], Peter Balk, Robert Hall[2], Mariska Nijenhuis,
Hetty van den Broeck[2], Oscar Vorst[2], and Astrid Poelman.
Agrotechnological Research Institute, PO-box 17, 6700 AA Wageningen, The Netherlands.
[1]corresponding author: m.f.vanwordragen@ato.dlo.nl
[2]Plant Research International, PO-box 16, 6700 AA Wageningen, The Netherlands.

Keywords: *Malus ×domestica*, cDNA arrays, quality change models, storage, batch identification, initial quality.

Abstract

Genomics technologies can be applied for unravelling the biological processes that are important for pre- and post-harvest quality. Such knowledge will allow precise monitoring of physiological condition and can be used as a tool for optimising production chains. A combination with quality change models will allow prediction of future product quality with much more precision and certainty than is currently possible. To demonstrate the validity of this approach we set up a project to identify, via cDNA microarray analysis, gene expression patterns that are correlated to the development of mealiness during storage of Cox apples (*Malus ×domestica* 'Cox's Orange Pippin'). The experimental set-up encompassed, next to mRNA profiling, both instrumental and sensory analyses of apples at various time points before and during storage. We were able to select a subset of genes the expression of which is indicative of reduced sensory quality (mealiness and off-flavour). Moreover we detected batch differences in initial profiles that could be related to storage quality. Such indicators can be used as decision support tools for growers and traders.

INTRODUCTION

The advent of genomics technologies has triggered a revolution in the medical and life sciences that can only be compared with the introduction of the microprocessor in information technology. Compared to standard methodologies, genomics represents a tremendous acceleration in knowledge development. It enables the formulation of research goals that were far beyond reach only 5 years ago. In the near future, this innovative technology will inevitably have a large impact on the agribusiness sector as well. Initial developments are already noticeable (Lemieux et al., 1998; Robinson, 1999; Menrad, 2000). Basically, genomics is the large-scale, fast and reliable analysis of the various groups of bio-molecules (DNA, RNA, proteins, metabolites) present in each living tissue. Most progress has been made in the analysis of DNA including e.g. large-scale sequencing projects such as the Human Genome project and the Arabidopsis and Rice sequencing projects and high throughput screening of mRNA (cDNA microarrays or DNA chips). Physiological changes in living material are preceded by a shift in gene expression, followed by alterations in protein and/or metabolite content. Therefore, mRNA or protein profiles may be used as early indicators for imminent physiological developments. This is not a new idea as molecular biologists have been coupling gene expression with functional changes for decades. However, since most quality characteristics are complex traits, it is insufficient to take into account just one or a few genes. Genomics technology now offers the opportunity to monitor thousands of genes or gene products simultaneously. This allows us to gain an integrated view of the processes that constitute a living organism and it enables us to zoom in on those events that seem most linked to the trait of interest. Thus far, genomics research related to agroproduct quality has mainly been focused on fundamental topics, such as unravelling biological processes involved in quality characteristics, QTL-mapping and marker assisted breeding (Pan, 1994; Buchanan-Wollaston, 1997; Sharma et al, 2002). The diagnostic potential of

genomics, which is broadly exploited in e.g. the medical sector (Carrico, 2000; Vesell, 2000; Ivanov et al. 2000), has been slow to take off in plant science. In this paper we will focus on the opportunities that open up when genomics technologies are used as a diagnostic instrument in agroproduction and distribution chains.

Fresh products that are generated in food production chains often show inter-batch variation in intrinsic quality. This initial batch quality has a strong influence on the type of application that the product can have in the (fresh) food chain and consequently, on its market value (Hertog et al., 1999). In addition, batch quality is the main parameter in decisions concerning (international) market choice. At present, these batch-to-batch differences in quality are only marginally determined and consequently are hardly exploited.

Genomics technology potentially offers a complete new spectrum of possibilities to assess the quality of fresh products and can be used as a support tool for decisions concerning applications, treatments or destinations for specific batches. An extremely powerful predictive tool can be generated by the combination of genomics data to quality change models. At present, one of the main drawbacks in mathematical models designed to predict quality changes in agro-products is the lack of reliable, high-density data input (Wilkinson 2001; Schouten, 2002). Ideally, this input should give a full and detailed description of the initial physiological state of the product. Data generated by e.g. cDNA microarrays exactly match these requirements and will therefore enable more accurate predictions of remaining quality or shelf life. The reliable information will support traders in making market choices and will expand their export range. Other applications of genomics that are envisaged are e.g.:

- (semi-)continuous monitoring of product condition during storage or ripening for dynamic control of environmental conditions. Genomics profiling determines the minimal requirements of the batch.
- Precise indication of the need for protective treatments – detection of initial infections or enhanced susceptibility to infection
- Enable an economically viable choice for cheaper means of transportation (sea instead of air transport) by a reliable indication of remaining quality after transport.

As an initial proof of concept for these ideas we designed an experiment in which genomics expertise is coupled to plant physiology and sensory science. The main goal was to determine whether a complex and hard to define post-harvest quality trait could be correlated to gene expression profiles. As a test case we chose development of ‘mealiness’ during storage of Cox’s ‘Orange Pippin’ apples.

MATERIALS AND METHODS

Plant material

For the trial we used two local provenances, designated ‘P39’ and ‘S39’ of *Malus ×domestica* ‘Cox’s Orange Pippin’. Both provenances were harvested in week 39 in 2002 and stored until use in the same controlled environment room at 4 ºC, 90% humidity in the dark.

Experimental set up

Apples from both provenances were stored at 2 temperatures: 4 ºC (control) and 18 ºC (mealiness induction), both at 90% humidity. Samples were taken at time point 0 and after 1, 2 and 3 weeks for morphological, physiological, sensory and molecular analyses. The experiment was repeated starting 1 week later using new batches of apples.

Sensory analysis

An independent sensory panel, consisting of 9 men and 5 women was trained according to the Quantitative Descriptive Analysis (QDA) method (Stone and Sidel, 1993). The fruit samples were provided at the same temperature, freshly cut and peeled. Only the non-blushed site of the apple was used. In one day, 12 samples (6 in duplicate)

were randomly presented to each panel member in 2 sessions of 6. Results were analysed using standard statistical methods such as ANOVA and PCA analysis.

Morphological and texture assays

At each sample moment, apples were scored visually for external disorders. Affected apples were excluded from the analyses. Texture of fruit tissue was determined at each sample date by firmness measurements. Apples (N=10) were peeled and inserted into an automated Fruit Texture Analyser (Güss). Firmness was determined by the level of resistance offered against penetration of a metal rod (Veltman et al., 2003).

cDNA library and RNA isolation

To generate cDNA clones for constructing the microarray, three cDNA libraries were produced. One full-length cDNA library derived from pooled tissue samples of apples in all physiological stages, made using the ZAP cDNA synthesis and Gigapack III GOLD cloning kit (Stratagene) conforming to the instructions of the manufacturer. Two subtracted libraries, Forward and Reverse (F and R), enriched for cDNAs related to stored or fresh apples respectively. For this the Clontech PCR Select (SSH) cDNA subtraction kit was used according to the instructions of the manufacturer. Gene fragments selected for the array were PCR-amplified and purified using a QIAquick 96 PCR purification kit. RNA was isolated using a CTAB procedure, essentially as described by Chang et al. (1993).

Array construction and hybridisation

In total 1588 PCR fragments derived from the three libraries were sequenced, 480 came from the Full library, 724 from Forward and 384 from Reverse. The 1588 fragments were organised in 1013 contigs, each potentially representing a different gene. For construction of the array 985 of these DNA fragments, representing 892 different contigs were used. Spotting of the array, labelling of the test samples and hybridisation was performed as described before (Aharoni et al., 2000). Results were analysed using statistical methods such as PCA and hierarchical clustering using GeneMaths 2.0 (Applied Maths).

RESULTS AND DISCUSSION

A large difference in quality was observed between the two provenances. After storage at 18 ºC, P39 apples did not develop major damage and less than 5% had to be discarded. The other provenance, S39, however was severely affected by pathogens and rot. More than 50% had to be discarded and this prevented analyses at later time points. However, physiological and morphological screens of both provenances at time-point 0 did not reveal this difference in quality.

Texture evaluation

Texture measurements showed a rapid decrease in firmness in S39 apples, stored at 18 ºC. The decrease in P39 apples at 18 ºC is significant ($p<0.05$) but much less pronounced. At 4 ºC significant reduction in firmness was only observed for S39. There was no significant difference between initial values for S39 and P39. Instrumental and visual inspection of the apples did not reveal any initial quality difference between provenances P39 and S39. Variation in quality was only detected by firmness measurements of 18 ºC-stored batches, in which quality differences were already evident on visual inspection.

Sensory analysis

The taste panel training resulted in 19 defined sensory attributes, subdivided into the categories odour, first bite, taste, mouth feel and after taste. No significant differences were identified during the experiment for the attributes related to odour. Table 1 lists the remaining attributes and shows the significant differences between the samples tested. No

differences in initial sensory quality were found between S39 and P39. However, the results after storage show that for almost all attributes, samples stored at 4 ºC, have higher scores compared to samples stored at 18 ºC. The only exceptions are off-flavour and mealiness. The only attribute that clearly increased over time is mealiness. There is a strong effect from storage at 18 ºC, but few differences were observed between week 0, 1, 2 and 3 for samples stored at 4 ºC.

Gene expression profiling

The mRNA samples used for hybridising to the array are listed in Table 2. All samples were co-hybridised with a common reference sample to allow for comparisons to be made between independent hybridisations. PCA analysis (Figure 1) revealed that samples D and H (1 week at 18 ºC) and samples E and I (2 weeks at 18 ºC) formed two clearly distinct groups. All other samples (storage at 4 ºC) clustered together in a third group. As shown in Figure 1 relatively few genes are responsible for this result. The individual expression profiles of these most discriminating genes are shown in Figure 2. The expression profiles of these genes form an accurate indicator of the physiological condition of the apple. In addition 20 genes were identified that showed significantly different gene-expression in initial samples of P39 and S39 (Figure 3). Remarkably, the expression profile for these genes in P39-2 week storage samples has an intermediate pattern. This suggests that the physiological age of the S39 samples at week 0 is comparable to that of P39 apples after 2 weeks storage at 4 ºC.

Combining data

In Figure 4 the sensory and gene expression data are plotted together in a PCA graph. In this figure two main clusters are formed that represent genes correlated to storage at 18 ºC on the left, and genes correlated to cold storage on the right. The sensory attributes are distributed unevenly in this plot. Most attributes correlate with the right cluster, which is related to freshness. Only mealiness and off-flavour correlate with the left cluster. This is in agreement with the results found in the sensory analyses, revealing that only off-flavour and mealiness had higher scores in 18 ºC stored fruit. In agreement with the PCA in Figure 1, the sample plot for this PCA shows a close correlation between samples E and I, and between samples D and H. The remaining group, though more correlated to each other than to E, I, D or H, is more dispersed in this second PCA. Remarkably, sample J (S39, 0 weeks storage) correlates closely to samples C and G (2 weeks stored, 4 ºC). This supports the conclusion drawn from Figure 2, that sample J is physiologically comparable to two week-stored P39 apples.

CONCLUSIONS

The described experiments show that it is feasible to correlate complex physiological conditions to a manageable set of gene expression profiles. Robust selection and a sound experimental set up ensure that the selected genes are reproducible indicators and can thus be implemented in a practical assay. Such assays allow sensory evaluation without the need for a trained sensory panel.

In addition, it was shown that expressional profiling is able to discriminate between batches that appear identical in other screens, but perform differentially downstream in the chain. This type of results can have major implications in agro-production. Reliable batch typing will enable the development of precision chains in which stakeholders will be able to select the most appropriate post harvest chain route for each batch. For the apple test case described here, this means that traders will be able to select batches with high post-storage quality. However, though the data presented are sufficient as a proof-of-concept, validation is needed before implementation of the results in agroproduction practice. The next step in proving the concept will be linking gene expression information to quality change models. This should allow for an accurate prediction of storage performance and development of sensory attributes. This challenging work is now in progress in our lab.

ACKNOWLEDGEMENTS

The authors want to acknowledge Pol Tijskens and Herman Peppelenbos for critical and inspiring discussions. This work was funded by a grant from the Dutch Ministry of Agriculture as part of DWK programme number 355.

Literature Cited

Aharoni, A., Keizer, L.C.P. Bouwmeester, H.J., Sun, Z., Alvarez-Huerta, M., Verhoeven, H.A., Blaas, J., Van Houwelingen, A.M.M.L., De Vos, R.C.H., Van der Voet, H., Jansen, R.C., Guis, M., Mol, J., Davis, R.W., Schena, M., Van Tunen, A.J. and O'Connell, A.P. 2000. Identification of the SAAT Gene Involved in Strawberry Flavor Biogenesis by Use of DNA Microarrays. Plant Cell 12:647-662.

Buchanan-Wollaston, V. 1997. The molecular biology of leaf senescence. J. Exp. Bot. 48;181-199.

Carrico, J.M . 2000. Human Genome Project and pharmacogenomics-implications for pharmacy. J. Am. Pharm.Assoc. 40:115-116

Chang, S., Puryear J. and Cairney, J. 1993. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. 11:113-116.

Hertog, M.L.A.T.M.; Boerrigter, H.A.M.; van den Boogaard, G.J.P.M.; Tijskens, L.M.M.; van Schaik, A.C.R. 1999. Predicting keeping quality of strawberries (cv. "Elsanta") packed under modified atmospheres: a integrated model approach. Postharvest Biol and Technol.15:1-12

Ivanov, I., Schaab, C., Planitzer, S., Teichmann, U., Machl, A., Theml, S., Meier-Ewert, S., Seizinger, B., Loferer, H. 2000. DNA microarray technology and antimicrobial drug discovery. Pharmacogenomics. 1:169-178.

Lemieux, B., Aharoni, A., Schena, M. 1998. Overview of DNA chip technology. Molecular Breeding 4:277 – 289

Menrad, K. 2000. Economic implications of Agro-Food biotechnology. J. Sci. Food. Agric. 80:539-546.

Pan, A, Hayes, P.M., Chen, F., Chen, T.H.H., Blake, T., Wright, S., Karsai, I., Bedo, Z. 1994. Genetic analysis of the components of winterhardiness in barley (Hordeum vulgare L.). Theor.Appl.Genet. 89:900-910.

Robinson, C. 1999. Making forest biotechnology a commercial reality. Nature Biotechnology 17:27-30.

Schouten, R.E., Tijskens, L.M.M., Van Kooten, O. 2002. Predicting keeping quality of batches of numbers of cucumber fruit based on a physiological mechanism. Postharvest Biol.Technol. 26:209 – 220.

Sharma, H.C., Crouch, J.H., Sharma, K.K., Seetharama, N., Hash, C.T. 2002. Plant Science 163:381-395.

Stone, H. and Sidel, J.L. 1993. Sensory evaluation practices. Academic Press Inc. San Diego.

Veltman, R.H., Verschoor, J.A. and Ruijsch van Dugteren, J.H. 2003. Dynamic control system (DCS) for apples (*Malus domestica* Borkh. cv 'Elstar'): optimal quality through storage based on product response. Post Harvest Biol. Techn. 27: 79-86.

Vesell, E.S. 2000. Advances in pharmacogenetics and pharmacogenomics. J. Clin. Pharmacol. 40:930-938

Wilkinson, E.C. and Tijskens, L.M.M. 2001. Modelling food quality. In: L.M.M. Tijskens, M.L.A.T.M. Hertog and B.M. Nicolai. (eds.) Food process modelling, Cambridge, UK.

Tables

Table 1. Sensory analysis of apple samples from two provenances (P), stored at two temperatures (T) for 0, 1, 2 or 3 weeks. Only attributes revealing significant differences between the samples are shown.

SAMPLE			First Bite		Taste					Mouth feeling			After taste		
P	T °C	wk	Crispness	Juicy	Taste sour	Taste fresh	Taste intensity	Taste apple	Off-flavour	Mealiness	Firmness	Juiciness	Ease to swallow	After taste intensity	Aftertaste duration
S39	4	0	abc[1]	ab	ab	ab	a	a	f	cd	ab	ab	abcd	a	ab
P39-I[2]	4	0	abc	a	ab	a	a	a	ef	d	ab	ab	abc	ab	a
P39-II	4	0	ab	abc	ab	a	a	ab	cef	d	abc	a	ab	abc	ab
P39-I	4	1	abc	abcd	abc	abcd	bcd	abc	cef	bcd	abcd	abc	d	bcde	abcd
P39-II	4	1	abc	abc	bcde	ab	abc	abc	abc	abc	abcd	bcde	abcd	abc	abcd
P39-I	4	2	abcd	abc	abc	abcd	ab	abc	abcdef	bcd	cd	ab	a	abcd	bcde
P39-II	4	2	abc	abcde	ab	ab	abc	abc	bcdef	bcd	abc	abc	abcd	abc	abc
P39-I	4	3	abc	abc	ab	a	a	a	ab	bcd	a	abc	abc	abc	abc
P39-II	4	3	a	abcd	a	ab	a	ab	abcd	bcd	abcd	cdef	ab	ab	a
P39-I	18	1	cd	cdef	abcd	bcde	bcd	abcd	cdef	abc	e	cdef	d	bcdef	abcd
P39-II	18	1	abc	abc	abcd	abc	abc	abc	abcdef	abc	abc	abcd	abcd	cdef	bcde
P39-I	18	2	abcd	bcde	def	cde	d	d	abcde	abcd	cd	cdef	abcd	ef	cde
P39-II	18	2	abc	ef	cdef	cde	d	cd	ab	bc	abcd	ef	bcd	f	cde
P39-I	18	3	bcd	def	ef	e	cd	cd	a	a	bcd	def	cd	def	de
P39-II	18	3	d	f	f	de	d	bcd	abd	ab	d	f	d	f	e

[1]Identical letters means: no significant difference. An 'a' represents a higher score than a 'b' and so on.
[2]I and II refer to duplicates

Table 2. Samples used for hybridising the cDNA array

Provenance	Storage Temp.	Weeks of storage	Trial nr.	Sample code
P39	4 °C	0	I	A
	4 °C	1	I	B
	4 °C	2	I	C
	18 °C	1	I	D
	18 °C	2	I	E
	4 °C	0	II	F
	4 °C	2	II	G
	18 °C	1	II	H
	18 °C	2	II	I
S39	4	0	II	J

Figures

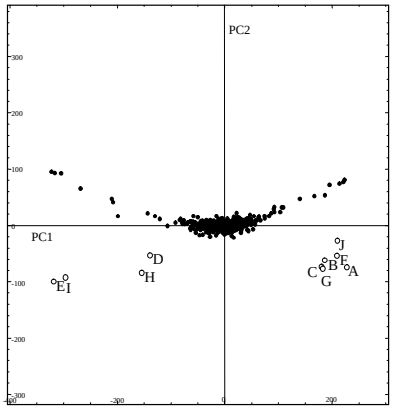

Fig. 1. PCA plot of gene expression profiles showing cDNA clones (dots) and samples (open circles)

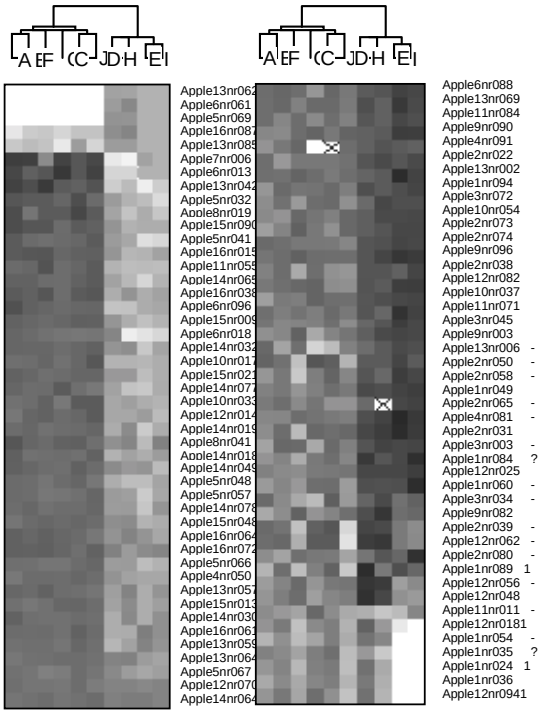

Fig. 2. Gene expression profiles most discriminating genes. White means low expression, dark means high expression. The left panel shows genes up-regulated and the right panel down-regulated during storage. The dendrogram represents the results of cluster-analysis of the samples.

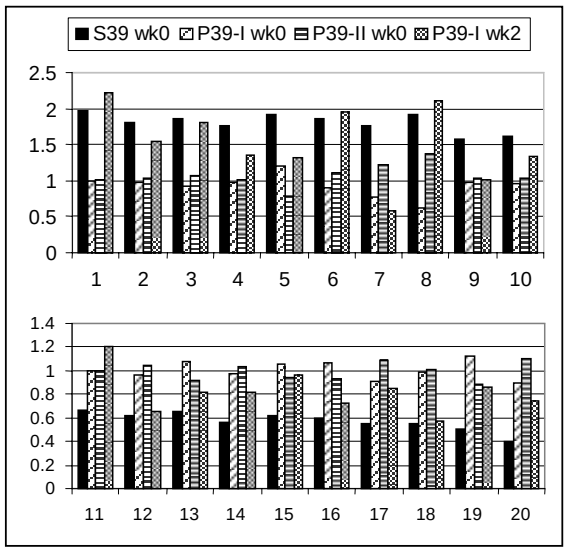

Fig. 3. Selection of 20 genes that are all differentially expressed between the provenances P39 (trials I and II) and S39, at week 0. P39 stored for 2 weeks shows an intermediate pattern. Left panel shows 10 genes of which expression is higher in S39 and the right panel shows 10 genes that are down-regulated in S39. Y-axis shows arbitrary units.

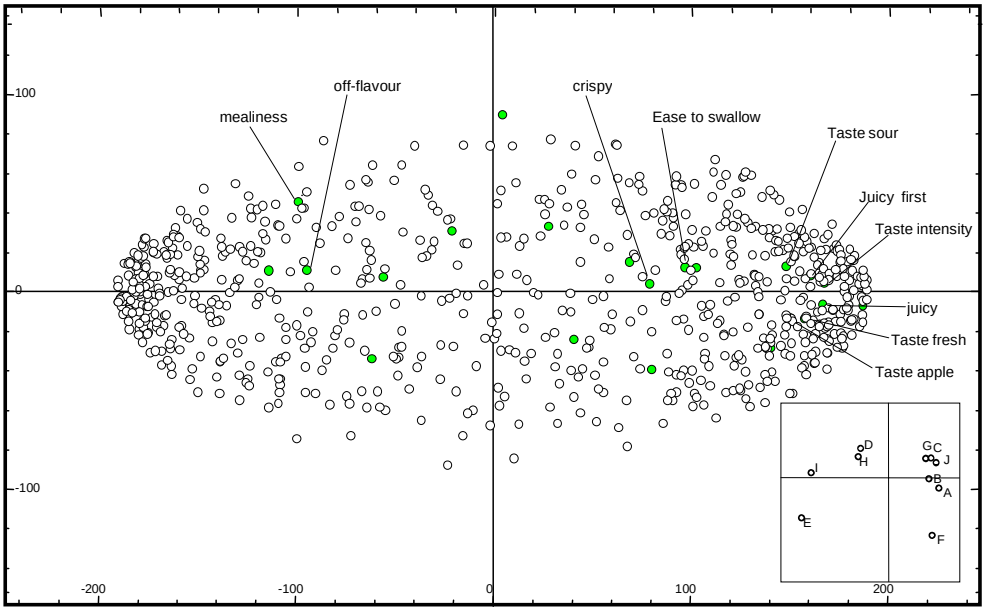

Fig. 4. PCA plot combining gene-expression and sensory data. Data were normalised both for genes and samples, resulting in an equal weighing factor for all genes and samples, irrespective of their absolute value. The cluster on the left is associated with storage at 18 ºC, the cluster on the right is associated with storage at 4 ºC.