

Doel bijeenkomst

Het gebruik van onderzoeksresultaten in veredelingsprogramma's

Genen, locatie van genen op chromosomen, moleculaire merkers, genen overbrengen naar een andere achtergrond, mogelijke problemen bij dat overbrengen, het verschil tussen een QTL en een gen

Sjaak van Heusden /Jaap Wolters
Plant Breeding WUR

Inleiding

- Eigenschappen die kunnen veranderen door de aanwezigheid van een of meerdere allelen resp. genen
- Gekloneerde genen (het gen is geïsoleerd en de functie is bevestigd)
 - Bijvoorbeeld door met genetische modificatie het gen in een andere achtergrond te zetten en naar de gevolgen hiervan te kijken
- Niet gekloneerde genen, waar vaak alleen het gebied rondom het gen bekend is
- Kwantitatieve eigenschappen (Quantitatieve Trait Loci, QTLs), een bepaald gen of een bepaalde regio waarvan de aanwezigheid correleert met een kwantitatieve eigenschap



Inleiding

- Filmpjes ter toelichting en vermaak (wel vaak in engels, maar een *picture tells a thousand words*)
 - Een aantal potentiële problemen bij het overbrengen van een gen/eigenschap naar een andere genetische achtergrond zullen besproken worden.
- Als de introgressie (inbouw) van een bepaald gen tot een verandering leidt is validatie (bevestiging) een vereiste als er andere kruisingen gemaakt worden. Nakomelingen met een specifiek gen en zonder dat gen met elkaar vergelijken.



Vervolg

- Per groep wordt een onderzoeksresultaat besproken (o.a. van TKI projecten), het onderzoeksresultaat wordt besproken en er worden een aantal vragen gesteld:
 - Hoe nuttig een eigenschap is en of het de investering waard is om op de eigenschap in een veredelingsprogramma te selecteren
 - Zo ja, hoe kan de eigenschap ingebouwd worden in een veredelingsprogramma
 - Kunnen eventuele problemen opgelost worden
 - De groep presenteert een plan de campagne (veredelingsstactiek) en beantwoordt de vragen.
- De verschillende projecten worden plenair besproken

Inkruisen resistentiegenen

Table 1 *Solanum* R genes that confer resistance to *Phytophthora infestans*

R gene	RGH ^a	<i>Solanum</i> species ^b	Origin ^c	Chr ^d	Reference
R1 family	None				
<i>R1</i>		<i>demissum</i>	Mexico	V	(3, 86)
R2 family	None				
<i>R2</i>		<i>demissum</i>	Mexico	IV	(8, 78, 82)
<i>Rpi-blb3</i>		<i>bulbocastanum</i>	Mexico	IV	(78, 95)
<i>Rpi-abpt</i>		Unknown ^e	Mexico	IV	(49, 78, 97)
<i>R2-like</i>		<i>edinense</i>	Mexico	IV	(17, 78, 96)
<i>Rpi-edn1.1</i>		<i>edinense</i>	Mexico	IV	(17)
<i>Rpi-snk1.1</i>		<i>scbenckii</i>	Mexico	IV	(17)
<i>Rpi-snk1.2</i>		<i>scbenckii</i>	Mexico	IV	(17)
<i>Rpi-bjt1.1</i>		<i>bjertingii</i>	Mexico	IV	(17)
<i>Rpi-bjt1.2</i>		<i>bjertingii</i>	Mexico	IV	(17)
<i>Rpi-bjt1.3</i>		<i>bjertingii</i>	Mexico	IV	(17)
<i>Rpi-mcd1</i>		<i>microdontum</i>	Argentina	IV	(77)
R3a family	I2 (94)				
<i>R3a</i>		<i>demissum</i>	Mexico	XI	(55, 56)
<i>Rpi-sto2</i>		<i>stoloniferum</i>	Mexico	XI	(17)
R4 family	Unknown ^f				
<i>R4</i>		<i>demissum</i>	Mexico	XI	(136, 142)
Rpi-blb1 family	None				
<i>Rpi-blb1, RB</i>		<i>bulbocastanum</i>	Mexico	VIII	(123, 134)
<i>Rpi-sto1</i>		<i>stoloniferum</i>	Mexico	VIII	(145)
<i>Rpi-pta1</i>		<i>stoloniferum</i> ^g	Mexico	VIII	(145)
Rpi-blb2 family	Mi (114)				
<i>Rpi-blb2</i>		<i>bulbocastanum</i>	Mexico	VI	(135)
Rpi-vnt1 family	Tm2 ² (72)				
<i>Rpi-vnt1.1</i>		<i>venturii</i>	Argentina	IX	(100)
<i>Rpi-vnt1.2</i>		<i>venturii</i>	Argentina	IX	(30)
<i>Rpi-vnt1.3</i>		<i>venturii</i>	Argentina	IX	(100)

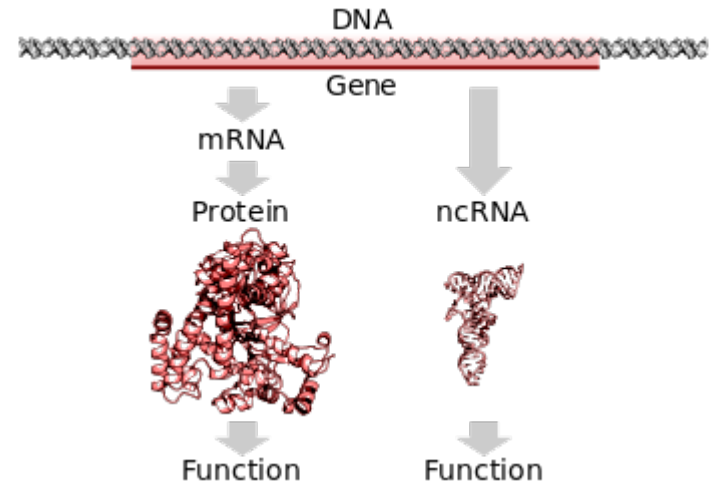
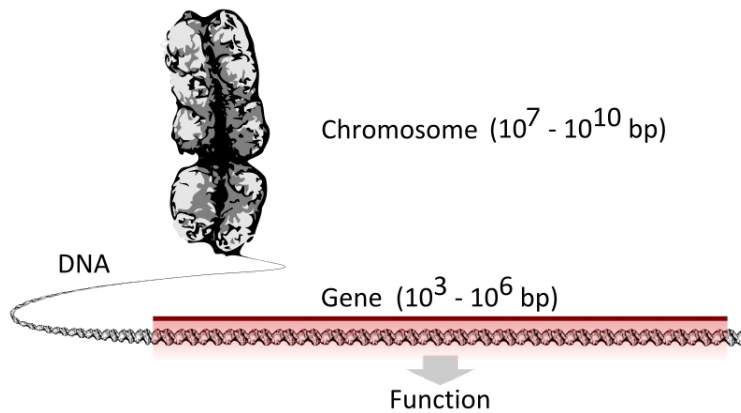
- Er zijn veel resistentiegenen tegen aardappelziekte geïdentificeerd in wilde verwanten van aardappel
- Hoe kun je deze inkruisen in aardappelcultivars?

Enkelgenskenmerken (of QTLs)

- Resistentiegenen
 - Tegen schimmels, virussen, bacteriën, insecten etc.
 - In siergewassen vaak meerdere genen nodig voor een resistentie
- Kwaliteitsaspecten
 - Zuurgraad, Brix, uiterlijk, kleur etc.
- Arbeidsbesparende eigenschappen
 - Minder uitlopers, zaadjes blijven aan plant, geen doornen etc. etc.



Achtergrond



<https://en.wikipedia.org/wiki/Gene>

In tomaat ~ 35.000 genen. Zonder studies is de locatie op de chromosomen onbekend. 6% van het DNA in tomaat zijn genen (verhoudingen verschillen per soort, in tomaat heeft 94% van het DNA geen duidelijke functie). Kunnen gekoppelde merkers ook buiten de genen gelokaliseerd zijn ?

Locatie van genen op de chromosomen

- Genetische studies kunnen laten zien of er één of meerdere genen betrokken zijn bij de verandering van een eigenschap.
- Met behulp van moleculaire merkers kan de koppeling van een gen en een moleculaire merker bepaald worden
 - Merkers: verschillen in DNA volgorde op plekken op een chromosoom (Single Nucleotide Polymorphisms, SNPs)
 - Voorbeelden indels, translocaties



Uitleg SNPs (in mensen)

<https://youtu.be/tJjXpiWKMyA>

Genotypering

- Ieder streepje is een SNP op een bepaald chromosoom: verschil tussen de twee ouders (blauw en rode)
 - Beide ouders homozygoot,
- Door veel SNPs te bepalen kan per plant gezien worden welk gedeelte van een chromosoom uit de ene ouder komt en welke uit de andere
- In een nakomelingenpopulatie kan de volgorde van de merkers bepaald worden (overkruising, koppelingsanalyse)



Homozygotie, heterozygotie



Door meiose ontstaan chromosomen die uit beide ouders komen

<https://youtu.be/4YfDq8rPDdM>

Crossing over (overkruising)

kan plaats vinden:

- * tussen genen,
- * tussen verschillende SNPs
- tussen een gen en een SNP

<https://youtu.be/jd3H0pgVCr8>

- De overkruisingsfrequentie is het aantal recombinaties in 100 nakomelingen

Meiosis

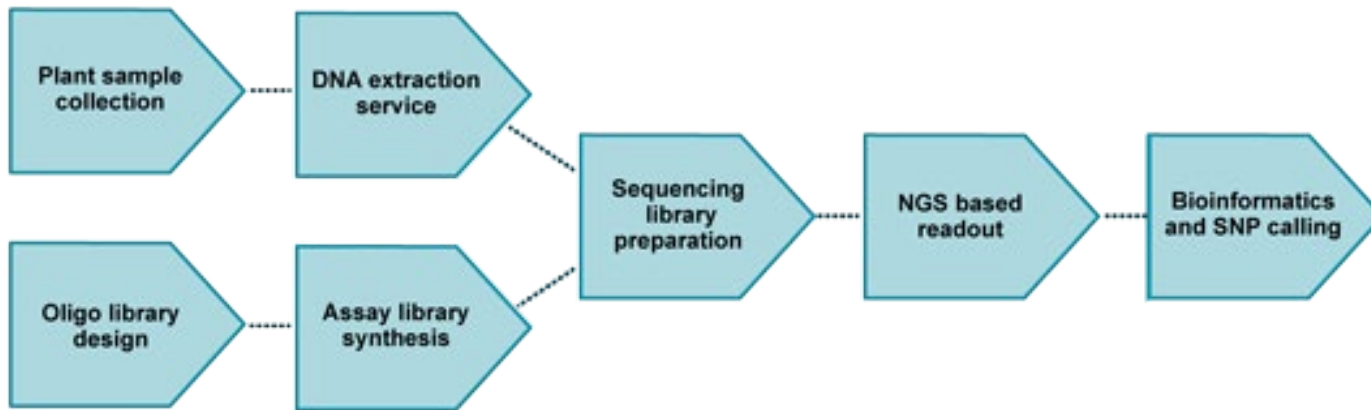
https://youtu.be/D1_mQS_FZ0

Bead array (veel SNPs tegelijkertijd)

- Array die verschillen zichtbaar maakt op vele verschillende posities

<https://youtu.be/IVG04dAAyvY>

Door op grote schaal te sequencen kunnen veel SNPs tegelijk bepaald worden



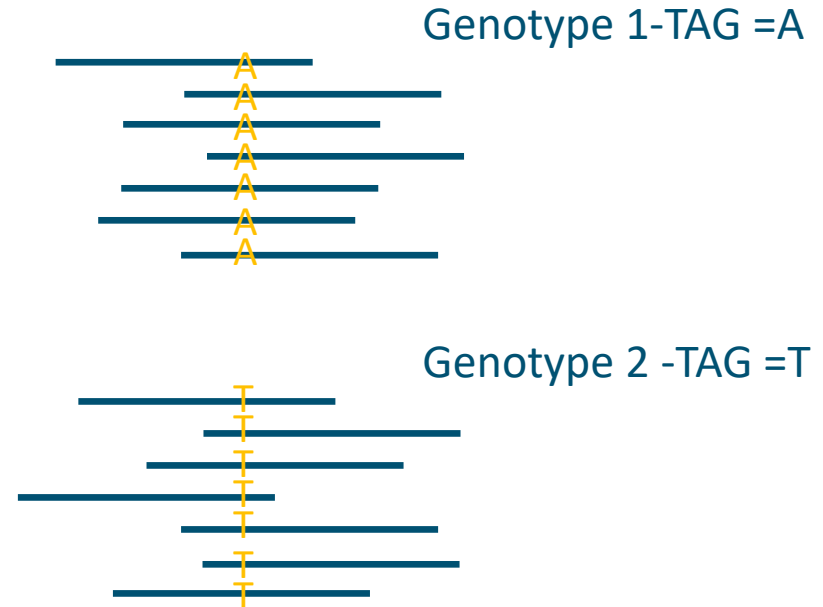
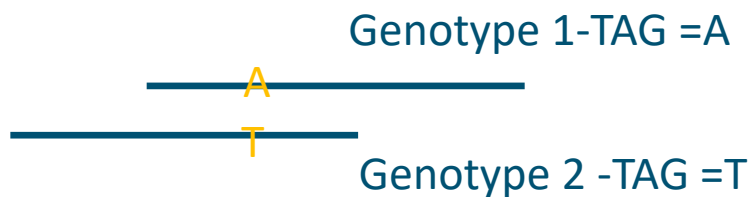
SeqSNP

- Planten waarvan gekozen stukjes van het genoom vermenigvuldigd worden (PCR)
 - Willekeurig gekozen
 - In genen of waar op basis van bekende gegevens (databases) SNPs verwacht worden
 - Dit wordt per plant uitgevoerd, de verschillende planten worden herkend doordat labels toegevoegd worden
- Deze vermenigvuldigde stukjes worden gesequenced (ieder fragment vele keren)
- Scoren van de Single Nucleotide Polymorfismen (sequentieverschillen)



Betrouwbaarheid van bepalingen

- Hoe kan je betrouwbare verschillen vinden
- Bepaalde stukken veel sequencen en vergelijken
- Op deze manier zorgen sequentiefouten niet voor valse SNPs



Onderzoeksprojecten

■ Onderzoek:

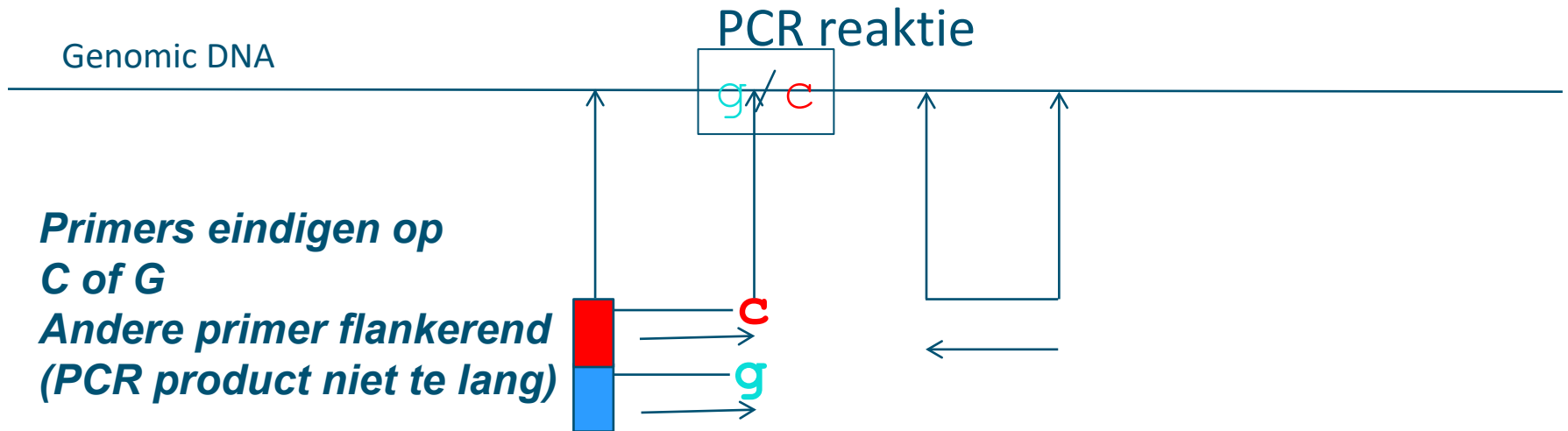
- Veel merkers, willekeurig verspreid over de chromosomen, zoeken naar beste koppeling tussen eigenschap en merkers

■ Gebruik:

- In veredelingsprojecten, met enkele merkers om aanwezigheid gen aan te tonen
- Bij meer eigenschappen zijn er meer merkers nodig!



KASP genotyping (Kbioscience)



PCR reaktie – kort produkt 25 -100 basen

Een SNP/plant per reaktiemengsel

Niet duur – automatische score ~ €0.10 per datapoint

<https://youtu.be/Uq9HhmzOqUQ>

Koppeling (linkage):

Merkers en/of genen op hetzelfde chromosoom

- Hoe dichters twee merkers/genen bij elkaar liggen hoe kleiner is de kans op overkruising
- Genetische afstand wordt bepaald door het aantal overkruisingen te tellen tussen twee verschillende merkers/genen
- Twee merkers die ver van elkaar liggen op hetzelfde chromosoom kunnen niet gekoppeld lijken (oplossing meer merkers op het chromosoom bepalen)
- Merkers op verschillende chromosomen zijn ongekoppeld



Software programma's voor koppelingsanalyses

- Hangt af van het aantal merkers die bepaald zijn
- Programma's zoals JoinMap5 en MapQTL6 zijn in het verleden veel gebruikt maar er zijn veel meer programma's
- Voor analyse polyploïden is nieuwe software gemaakt (TKI)

Bioinformatics, 34(20), 2018, 3496–3502
doi: 10.1093/bioinformatics/bty371
Advance Access Publication Date: 2 May 2018
Original Paper

OXFORD

Genetics and population analysis

polymapR—linkage analysis and genetic map construction from F_1 populations of outcrossing polyploids

Peter M. Bourke¹, Geert van Geest^{1,2}, Roeland E. Voorrips^{1,*}, Johannes Jansen³, Twan Kranenburg¹, Arwa Shahin^{1,4}, Richard G. F. Visser¹, Paul Arens¹, Marinus J. M. Smulders¹ and Chris Maliepaard^{1,*}

¹Plant Breeding, Wageningen University & Research, Wageningen 6708 PB, The Netherlands, ²Deliflor Chrysanten B.V, Maasdijk 2676 BS, The Netherlands, ³Biometris, Wageningen University & Research, Wageningen 6708 PB, The Netherlands and ⁴Van Zanten Breeding B. V, Rijsenhout, 1435 EW, The Netherlands

Hoe groot is het chromosoomgedeelte (locus) waarop je moet selekteren om het gen in te kunnen kruisen

De betrouwbaarheid van de locatie (het interval) hangt af van:

- Hoe betrouwbaar de eigenschap te meten was (betrouwbaarheid fenotypering).
- Hoe betrouwbaar zijn alle merkers gescoord? Met veel merkers zijn incidentele fouten te corrigeren!
- De fysieke grootte van het te beselecteren fragment hangt af van de recombinatiefrequentie rondom de locatie

Recombinatiefrequenties verschillen over het chromosoom heen.

- Als de recombinatiefrequentie hoog is rondom een gen kan de fysieke locatie nauwkeuriger bepaald worden (meer overkruisingen)
- Als de recombinatiefrequentie laag is rondom een gen kan de exacte locatie minder nauwkeurig bepaald worden
- Voorbeeld: genetische lengte van chromosoom 1 tomaat is 133 cM en de fysieke afstand 87×10^6 bp
 - weinig recombinatie (cold spots) $1 \text{ cM} = 11 \times 10^6 \text{ bp}$
 - veel recombinatie (hot spots) $1 \text{ cM} = 0.14 \times 10^6 \text{ bp}$
- Recombinatie frequentie hot spots 80x hoger dan in cold spots (uiteindes chromosomen vs midden chromosome/centromeren)

Gevolg verschillende recombinatiefrequenties

- De fysieke grootte van het te beselecteren fragment
- Grotere inspanning om genen te isoleren (te kloneren)
- Meer kans dat er ongewenste genen in het fragment zitten waarop geselecteerd is
 - In gebieden op het chromosoom met een hoge frequentie zitten relatief meer genen
 - In gebieden op het chromosoom met een lage frequentie zitten relatief minder genen

Vergelijken fysieke en genetische kaarten is mogelijk door NGS (Next Generation Sequences)

Zandraket, Rijst,
Aardappel, Tomaat,
Tomaat, Peper, Kool,
Banaan, Appel, Amandel
Suikerbiet, Komkommer
Etc. etc. etc

Waarom sequencen

- Detectie Single Nucleotide Polymorphisms (SNPs)
 - Door de sequenties van verschillende individuen te vergelijken worden zelfs genoeg SNPs gevonden als er maar heel weinig zijn
 - SNPs hoeven niet in genen te liggen!
- Voorspelling locatie en annotatie waar genen op een chromosoom liggen

Sequencing makes gene annotation possible

- Identifying the locations of genes and all coding regions in a genome
- Hypothesis of what the function of genes is, which must be confirmed
- Identification of gene families
- Compare with similar genes in other species that may have the same function
 - BLAST (Basic Local Alignment Search Tool), compare DNA sequences in various species

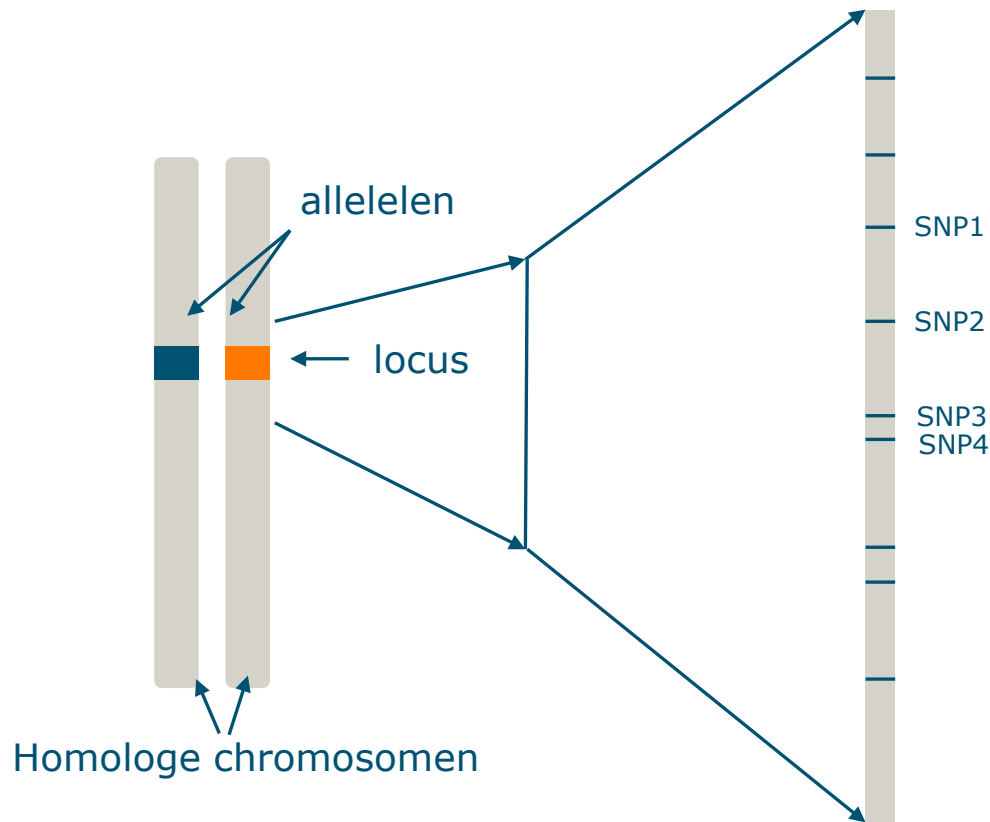
Hoe ver zijn we na een willekeurige studie die succesvol was ...

- Locatie en een idee van de grootte van regio waar gen van interesse ligt
- Overeenkomstige genen liggen in verschillende accessies van een soort meestal op de zelfde locatie op een chromosoom
 - Uitzondering translocaties (een stuk chromosoom is verplaatst/uitgewisseld van naar een andere plek op het genoom) of verwijderingen (deleties) van bepaalde stukken.

Hoe ver zijn we na een willekeurige studie die succesvol was ...

- Single Nucleotide Polymorfismes (SNPs) in de buurt van het gen (SNPs tussen donor van gen en recipient). Weinig overkruisingen tussen gen en SNP.
- Fysieke grootte van de regio alleen bekend als genoom sequentie bekend is!

Exacte locatie gen is onbekend, maar regio is bekend

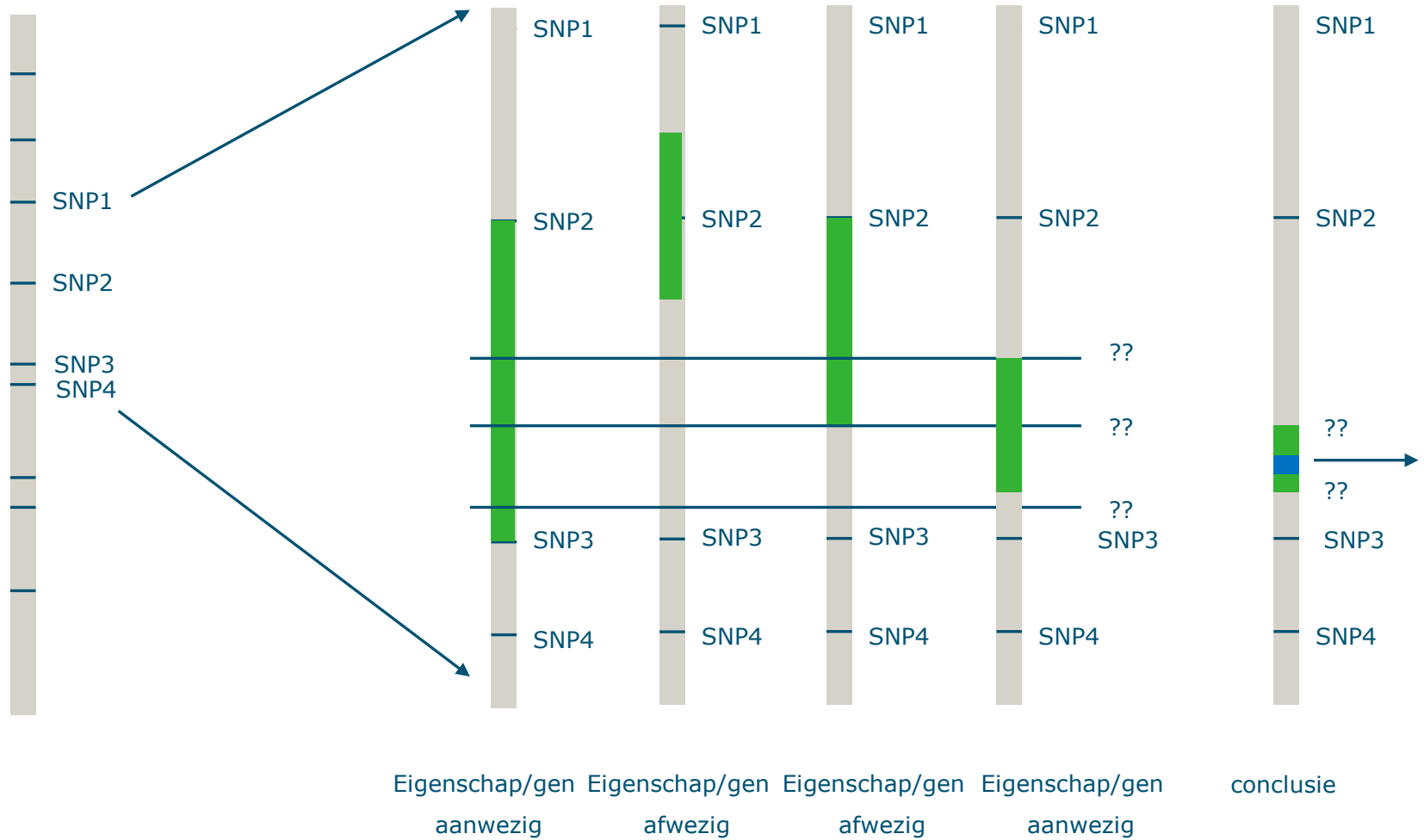


- Resultaat: Alle planten met de regio tussen SNP2 en SNP3 hebben het gen van interesse
- Dit maakt het mogelijk om op het juiste allel te selekteren

Vragen

- Is selectie op SNP2 en SNP3 genoeg om op het juiste allel te selecteren?
 - Ja, als de locatie van het locus goed bepaald is
- Als locatie SNP2 en SNP3 bekend is, is dan ook exacte locatie van het gen bekend
 - Nee, er kunnen meerdere genen tussen SNP2 en SNP 3 liggen
- Wat is nodig om het doelgebied kleiner te maken?
 - Meer merkers en meer recombinatiegebeurtenissen tussen SNP2 en SNP3
- Hoe krijg je meer recombinanten?
 - Populatie uitbreiden

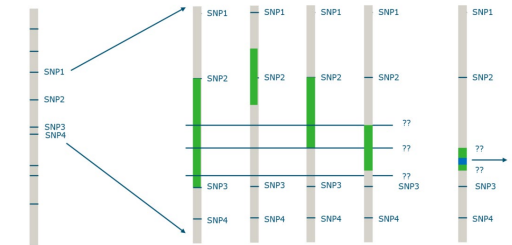
Meer merkers en recombinatie rondom gen

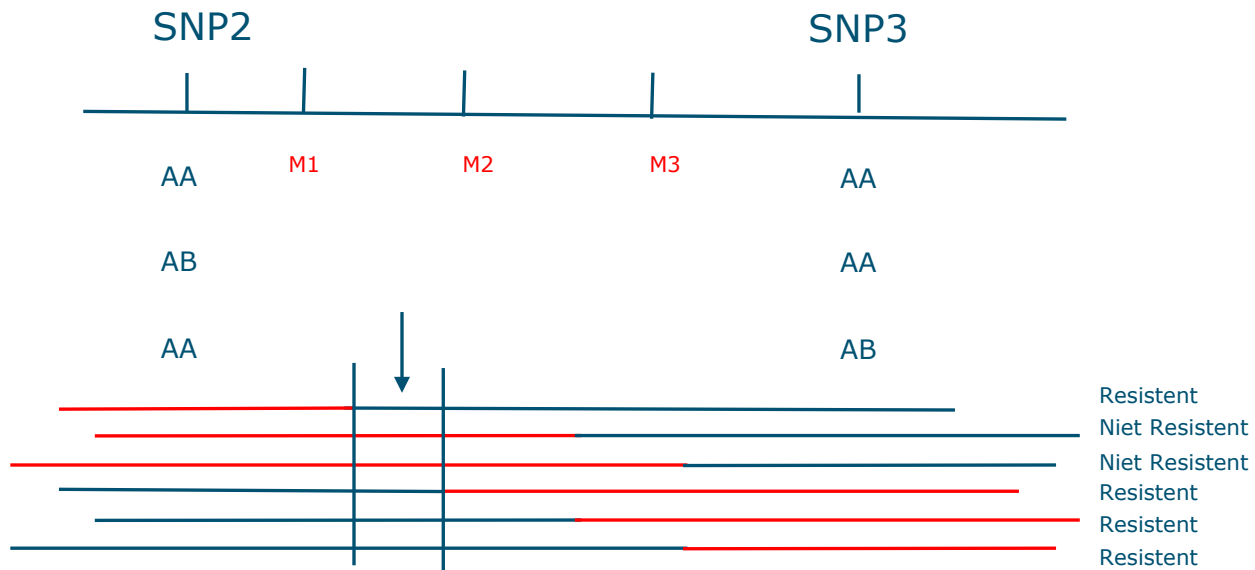


Aanpak om locatie beter te bepalen

- Grotere populatie (> 1000 -2000)
- Eerste test
 - Mbv SNP2 en SNP3 bepalen welke planten van de grote populatie een recombinitie hebben tussen SNP2 en SNP3
- Planten met een recombinitie zelfbestuiven en alle bekende merkers tussen SNP2 en SNP3 bepalen in de nakomelingen.
- Een aantal fenotypen.

Meer merkers en recombinitie rondom gen

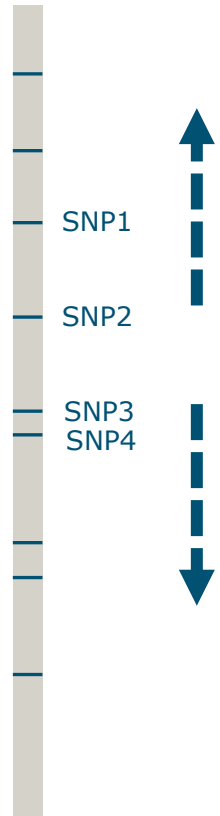




- Donor gen SNP2 AA SNP3 AA : homozygoot —————
- Ontvanger: SNP2 BB SNP3 BB : homozygoot —————
- Testen voor recombinanten: AB AA en AA AB
- Zelfbestuiving
- Markers tussen SNP2 en SNP3 (zoveel mogelijk)
- Fenotypering

Valkuilen / linkage drag

- Bij kruisingen worden vaak meer genen, aanwezig in hetzelfde gebied als het gewenste gen ingekruist (nadelige genen; linkage drag)
 - Tussen SNP2 en SNP3 kunnen ook nadelige genen aanwezig zijn. Deze genen kunnen dmv recombinitie van elkaar gescheiden worden
 - Linkage drag niet verwarren met pleotropie, een gen heft twee functies die niet noodzakelijkerwijze een gelijk fenotype hebben
 - Noordelijk van SNP2 en zuidelijk van SNP3 kunnen nadelige genen aanwezig zijn die niet wegveredeld zijn door gebrek aan overkruisingen resp. merkers



Linkage drag

- In het geval van linkage drag, wat is het directe gevolg voor een veredelingsprogramma?
 - Grotere populatie benodigd om overkruising te vinden tussen de twee genen
 - Eventueel kan in een veredelingsprogramma wel een variëteit gebruikt worden waar het al gelukt is een overkruising te vinden (breeders exemption)



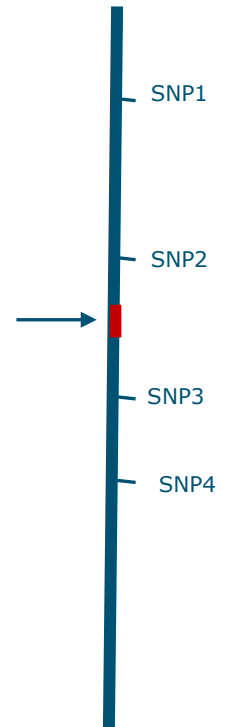
SNP in een bepaalde kruising is geen SNP in een andere kruising (wel dezelfde donor maar ander recipient).

- Het is mogelijk vanuit een bekende sequentie (van een bekende SNP) de dichtsbijzijnde SNP tussen de twee nieuwe ouders te vinden)
 - Mbv sequentie databases
 - Door het gebied naast een SNP te bepalen en de beide ouders met elkaar te vergelijken

Enkelgens kenmerken

■ Gekoppelde merkers

- Is één gekoppelde merker om een gen te selecteren genoeg of moeten er twee zijn of vier?
- Wat is het voordeel van 2 ipv 1 merkers en waar moeten deze merkers aan voldoen?
- Wat zou het grote voordeel van 4 merkers zijn en waar moeten deze merkers aan voldoen?
- Maakt het uit of er een kans op het meeslepen van nadelige genen is?



Waarom kan een enkel gen niet meer werkzaam zijn in een andere kruising?

- Er zijn meerdere genen nodig naast het gen waar je op selecteert.
 - De andere genen splitsten niet uit in de karteringspopulatie waarin het gen geïdentificeerd is.
- In een andere kruising kunnen ook die andere belangrijke of minder belangrijke genen uitsplitsen waardoor het effect van het gen waar je op selecteert minder aanwezig of zelfs afwezig is.
- Welke kennis is nodig om toch nog steeds op de positieve eigenschap te kunnen selecteren. Hoe krijg je die kennis?

QTLs (Quantitative Trait Loci).

- Regio met een of meerdere genen
- Regio die bijdraagt aan de variatie in kwantitatieve phenotypes
- Niet alle genen hebben eenzelfde invloed op de eigenschap.
 - Er kunnen ook op verschillende plekken op het genoom genen zijn die gecorreleerd zijn met dezelfde eigenschap (zoals opbrengst, grootte).
 - Als kwantitatieve eigenschappen door meerdere genen beïnvloedt worden kunnen de metingen van een eigenschap vloeiend in elkaar over gaan. Voorbeelden: lichaamslengte of -gewicht of de zaadopbrengst.



QTLs (Kwantitatieve eigenschappen)

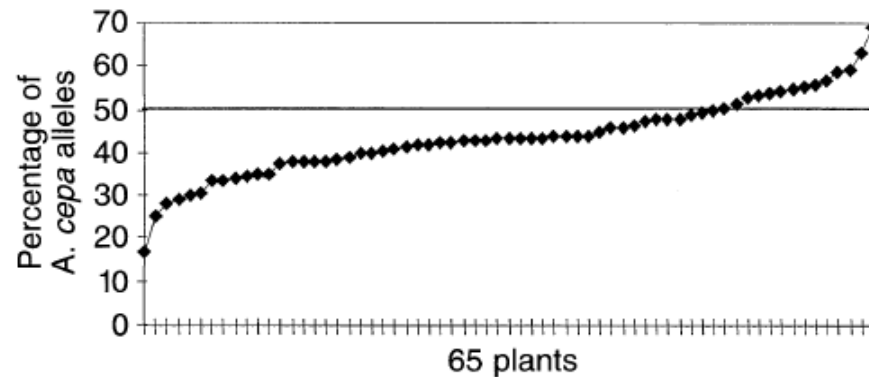
- QTLs of meergenskenmerken niet wezenlijk anders dan enkelgenskenmerken
 - Het nut van het selecteren op een of meerdere QTLs hangt af van de variatie die ermee verklaard wordt
 - Het feit dat er meerdere genen zijn die elkaar kunnen beïnvloeden (zowel positief als negatief) maakt het werken met QTLs lastiger
 - Vaak is de regio van een specifiek QTL ook moeilijker te bepalen. In een regio kunnen ook meerdere genen die elkaar beïnvloeden liggen maar dit hoeft niet noodzakelijkerwijs.



Selectie op de achtergrond in een kruising

- Het percentage van de recurrenente varieert in de nakomelingen van een kruising (zie figuur). Door de plant(en) te kiezen met een hoog percentage recurrenente ouder en natuurlijk de eigenschap kan het aantal terugkruisingen geminimaliseerd worden.

124



<https://wiki.groenkennisnet.nl/>

Fig. 2 The percentage of *A. cepa* alleles in individual F₂ plants sorted from low to high in a population of 65 plants

Validatie

De effect van het inkruisen moet bevestigd worden.

Is het effect uit de karteringspopulatie nog aanwezig in nakomelingen van een nieuwe kruising?

Validatie vooral in kruisingen met dezelfde donor en andere recipiënten

Verschil tussen nakomelingen met het beselecteerde gen

Met homogene achtergrond (alleen gen verschilt)

Met niet-homogene achtergrond (gen verschilt, achtergrond middelt uit)