

SLUTRAPPORT

UTVECKLING AV GENETISKA MARKÖRER FÖR RESISTENS, KVALITET OCH ODLINGSVÄRDE I HAVRE (projekt nr H0936280)

Alf Ceplitis
Lantmännen Lantbruk Växtförädling
(tidigare Svalöf Weibull AB)

Bakgrund

Havre är en viktig gröda i de nordiska länderna. När det gäller totalt odlad areal, produktion och andel av jordbruksmarkens användning är Sverige, Norge och Finland bland de främsta havreländerna i världen. Kvaliteten på nordiskt odlad havre har högt anseende och med tydligt hälsobefrämjande egenskaper föreligger också ett ökat intresse bland konsumenter för havre som livsmedel. För att Sverige och övriga nordiska länder skall kunna behålla en livskraftig och uthållig havreodling krävs sortmaterial anpassat till nordiska förhållanden. Det är därför av central betydelse att en inhemsk växtförädling kan förse marknaden med attraktiva sorter.

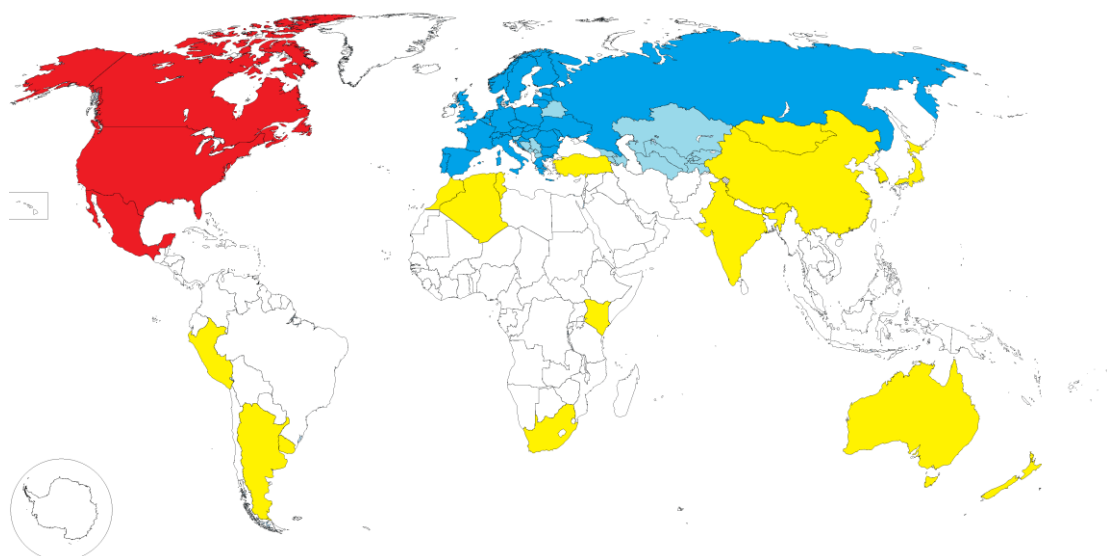
Användandet av genetiska markörer har fått allt större betydelse inom växtförädlingen, särskilt genom de senaste årens snabba utveckling av ”high throughput”-metoder varigenom tusentals markörer kan analyseras i stora mängder individer på kort tid till lågt pris. Havre har länge halkat efter stora grödor som majs, vete och korn på detta område, men nyligen har flera internationella konsortia (där Lantmännen varit delaktiga) arbetet med att ta fram markörer även för havre.

Detta projekts målsättning har varit att använda nya genetiska markörsystem i havre för att identifiera specifika markörer kopplade till egenskaper som representerar viktiga förädlingsmål för att på så sätt effektivisera framställningen av nya, förbättrade sorter för svenskt och nordiskt lantbruk

Material och metoder

Sortkollektion: sammansättning och uppförökning

För att kunna finna kopplingar mellan markörer och egenskaper har vi utgått från en fenotypiskt mycket variationsrik samling havreaccessioner. Totalt har vi insamlat 587 accessioner av hexaploid havre (*Avena sativa*), huvudsakligen moderna sorter med i princip globalt ursprung. Sorterna härstammar från 50 länder med samtliga kontinenter representerade, men tyngdpunkten ligger på europeisk – särskilt nordisk – havre. Sorterna grupperades i tre större regionala grupper: europeiska sorter, nordamerikanska sorter samt sorter från övriga världen (Fig 1). För 52 sorter har vi inga säkra uppgifter om ursprung. Frömaterial från de 587 sorterna har uppförökats i Svalöv 2011 med ytterligare uppförökning i Nya Zeeland 2011-2012 för att erhålla tillräcklig med utsäde för avkastningsförsök. Sortmaterialet har erhållits från Lantmännens egna samlingar, genbanker och andra förädlare. Vi har i görligaste mån försökt undvika vintertyper för att kunna utvärdera samtliga sorter gemensamt i vårsådda försök.



Figur 1. Karta visande ursprungsländerna för de 587 sorterna i projektets havrekollektion. Färgerna indikerar de tre regioner som länderna är indelade i.

Utvärdering av resistens, kvalitet och odlingsvärde

Sorterna i kollektionen har utvärderats med avseende på sjukdomsresistens, kärnkvalitet och inre (biokemisk) kvalitet, samt odlingsvärde. För detta syfte har fältförsök, antingen i småparceller, hill-plots, eller rader, eller i regelrätta avkastningsförsök i större parceller, utförts under åren 2011-2012 på flera platser i Sverige och utomlands. Observationer, mätningar och analyser har gjorts direkt i fält eller i laboratorium. En sammanfattning av de fältförsök och analyser som gjorts för olika egenskaper återfinns i Tabell 1.

Tabell 1. Egenskaper som analyserades i sortkollektionen

Egenskaps-kategori	Egenskap	Försöksplatser (antal)	Försöksår (antal)	Analys/gradering
Resistens	Kronrost	2	2	Visuell
Resistens	Mjöldagg	1	2	Visuell
Resistens	Flygsot	1	2	Visuell
Resistens	Nematoder	1	2	Visuell
Resistens	Fusarium	1	2	Våtkemisk (DON)
Kvalitet	Rymdvikt	5	1	Mekanisk
Kvalitet	Kärnvikt	5	1	Mekanisk
Kvalitet	Kärnstorlek	5	1	Mekanisk
Kvalitet	Kärnhalt	5	1	Mekanisk/NIR
Kvalitet	Skalbarhet	5	1	Mekanisk/NIR
Kvalitet	Betaglukanhalt	5	1	Våtkemisk/NIR
Kvalitet	Fetthalt	5	1	NIR
Kvalitet	Proteinhalt	5	1	NIR
Odlingsvärde	Avkastning	5	1	Mekanisk/NIR ¹
Odlingsvärde	Planthöjd	5	1	Visuell
Odlingsvärde	Blomningstid	5	1	Visuell
Odlingsvärde	Stråstyrka	5	1	Visuell

¹ Avkastning skattad som vikt av parcellskörd korregerad för vattenhalt uppmätt med NIR (NIR = near-infrared spectroscopy)

Marköranalys

Samtliga accessioner såddes i växthus för DNA-extraktion och marköranalys. DNA extraherades med en kloroformbaserad snabbmetod från bladvävnad bulkad från åtta plantor per accession. Extraherat DNA sändes till Diversity Arrays Technology Pty (Australien) för analys med DArT-markörer på en s.k. Medium Resolution Array omfattande maximalt 3072 markörer.

Identifiering av koppling mellan markörer och egen

DArT-markörer kopplade till de studerade egenskaperna identifierades genom s.k. associationskartering. Detta är en statistisk metod som kan påvisa associationer mellan markörer och QTL (Quantitative Trait Loci), dvs. områden i arvmassan där de gener som styr egenskapen i fråga förmodas finnas. Till skillnad från biparental kartering, där man studerar avkomman från korsning mellan två föräldralinjer, utgår man i associationskartering från en samling genotyper (i detta fall sorter) som kan vara mer eller mindre besläktade. Metoden ger bättre möjligheter att lokalisera QTL till en korrekt position i genomet, men kan underskatta effekten av enskilda QTL på det fenotypiska uttrycket. I denna studie har vi använt en generaliserad linjär statistisk modell med markör som fix effekt för analys av association mellan markörer och egenskaper. Bonferonni-korrektion användes för att kompensera individuella signifikansvärden för multipla test.

Vid associationskartering måste hänsyn tas till att den genetiska variationen i materialet kan vara ojämnt fördelat – att det med andra ord finns en populationsstruktur där vissa grupper av sorter inbördes kan vara mer lika än sorter från olika sådana grupper. En inledande analys av eventuell populationsstruktur gjordes därför med två metoder: dels en principalkomponentanalys (PCA) baserad på förekomst av alternativa alleler vid enskilda markörer; dels en skattning av den betingade sannolikheten att materialet är indelat i K antal distinkta grupper eller subpopulationer (där i detta fall $K \in \{1..6\}$).

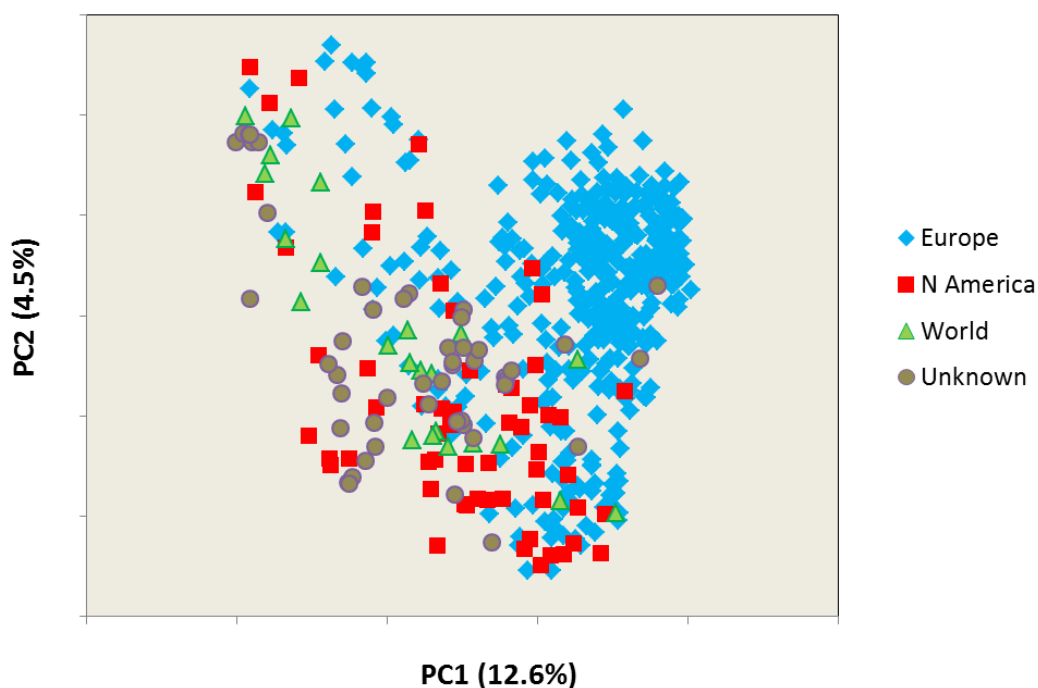
En viktig faktor vid associationskartering är graden av kopplingsojämvikt ("linkage disequilibrium", eller LD) mellan markörer, dvs. i vilken utsträckning alleler vid olika markörloci förekommer tillsammans i en population oftare än vad som förväntas av slumpen. Alleler vid loci som ligger nära varandra på inom samma kromosomavsnitt kan förväntas ha högre LD och hur snabbt LD minskar med ökande avstånd mellan markörer påverkar resultatet av associationskarteringen. Ju snabbare LD minskar med avståndet, desto snävare blir de QTL inom vilka de bakomliggande generna finns, men samtidigt krävs fler markörer för att kunna detektera några QTL. För DArT-markörerna finns positionsangivelser för 1010 markörer från en genetisk karta baserad på avkomman ur en korsning mellan sorterna 'Kanota' och 'Ogle' (den s.k. KO-kartan). Vi har skattat LD som den kvadrerade korrelationskoefficienten (r^2) mellan allelfrekvenserna vid par av markörer och analyserat förhållandet mellan LD och kartavstånd i den aktuella sortkollektionen för att beräkna hur långt LD sträcker sig.

Analys av populationsstruktur, QTL-kartering och LD utfördes med hjälp av datorprogrammen STRUCTURE och TASSEL, samt i programmeringsmiljön R.

Resultat

Utvärderingen av sortkollektionen i fält och laboratorium gav vid handen en mycket stor fenotypisk variation i materialet för samtliga undersökta egenskaper. Ett visst bortfall av data uppstod p.g.a. att vissa accessioner visade sig vara vintertyper (framförallt bland genbanksmaterial från Ryssland/Sovjetunionen), dålig grobarhet, dålig etablering, sen mognad, misslyckad marköranalys och liknande faktorer. Fusarium-försöket i Laberweinting 2012 tröskades oavsiktlig innan provtagning och kunde därför ej användas. Prover för DON-analys fick därför tas från avkastningsförsöken på Bjertorp 2012. Överlag erhöles dock en tillfredsställande mängd data med mindre än 15% "missing values" för varje enskild egenskap. Moderna nordiska och tyska sorter uppvisade den högsta avkastningen, vilket var väntat då de är selekterade för odling i området. Däremot kunde bland äldre och/eller mer exotiskt material finnas sorter med värdefulla egenskaper, framförallt kärn kvalitet och sjukdomsresistens. Sådant material kan vara intressant som genkällor för specifika kvalitetsegenskaper.

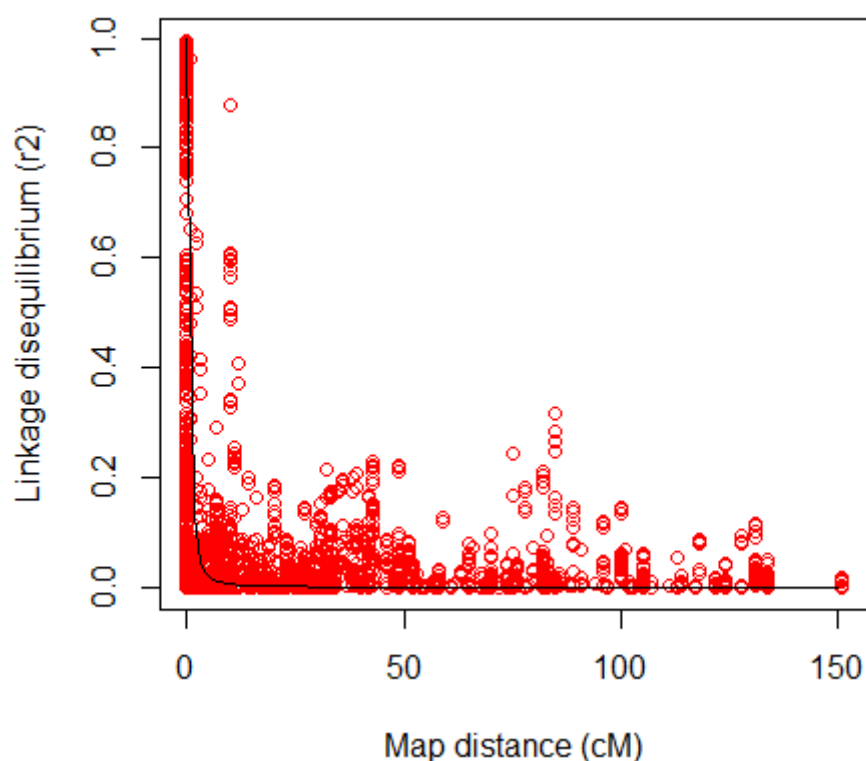
Efter att sorterat bort icke variabla markörer och markörer med mer än 15% "missing values" fanns vi totalt 747 variabla DArT-markörer i sortkollektionen som helhet. De europeiska sorterna utmärktes av den lägsta variationsgraden ($H = 0,26$, där H är den förväntade heterozygotigraden). Därefter kom de nordamerikanska sorterna ($H = 0,35$), medan gruppen av sorter med ursprung utanför Europa och Nordamerika hade den högsta graden av markörvariation ($H = 0,38$). Endast en svag populationsstruktur kunde skönjas i DArT-markördata. Ett diagram över de två första principalkomponenterna, som tillsammans svarar för 17% av den totala variationen i markördata, visar att det finns en viss separation av europeiska och utomeuropeiska sorter, men samtidigt ett påtagligt överlapp (Fig 2)



Figur 2. Diagram över de två första principalkomponenterna baserad på allelförekomst vid 747 DArT-markörer. Varje punkt representerar en sort och färgerna anger ursprungsregion.

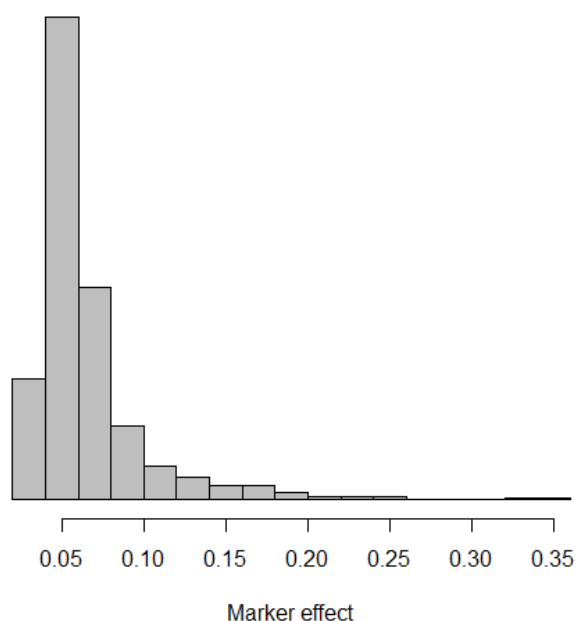
Inte heller vid en probabilistisk analys av populationsstruktur fanns tydligt stöd för ett visst antal subpopulationer. Den betingade sannolikheten för K antal subpopulationer ökade asymptotiskt med stigande K i intervallet $\{1..6\}$, vilket är en indikation på svag eller ingen populationsstruktur.

Av de 747 variabla markörerna återfanns 223 på KO-kartan och hade således en positionsangivelse. Ur korrelationen mellan LD och kartavstånd för varje par av de 223 markörerna skattade vi med hjälp av icke-linjär regression hur snabbt LD avtog med avståndet. Vi fann att LD minskade under 0,2 vid ett kartavstånd $<1,1$ cM (Fig. 3).



Figur 3. Kopplingsjämvikt (r^2) som funktion av kartavstånd (cM) i sortkollektionen, baserat på 233 markörer med angivna positioner i genomet.

Då ingen uppenbar genetisk populationsstruktur kunde upptäckas i sortkollektionen analyserades kopplingen mellan DArT-markörerna och var och en av de studerade egenskaperna under antagandet av avsaknad av sådan struktur. Inför analysen bortsåg vi från markörer med $H < 0,18$ (motsvarande en minsta-allelfrekvens om 0,1). Vi fann markörer associerade till samtliga egenskaper, med undantag av flygsots- och nematodresistens. För enskilda egenskaper fann vi mellan två (för betaglukanhalt) och 57 (för avkastning) associerade markörer. Samtidigt var flera markörer associerade till mer än en egenskap; i de flesta sådana fall gällde det markörer med koppling till avkastning, rymdvikt, kärnstorlek, kärnvikt och i vissa fall även stråstyrka. Markörer med koppling till sjukdomsresistens var oftast inte kopplade till några andra egenskaper. Markörernas effekt på fenotypen (dvs. hur stor andel av den fenotypiska variationen i materialet som förklaras av en given markör) varierade mellan ca 3% och 35%. Fördelningen av marköreffekter hade en tydlig övervikt åt markörer med liten effekt. Endast ett litet antal markörer hade stor effekt ($>30\%$) på fenotypen (Fig. 4).



Figur 4. Fördelning av effekter på fenotypen för de markörer med signifikant association till någon egenskap.

Diskussion

Genetiska markörer kan vara ett viktigt verktyg i växtförädlingen. Markörer kopplade till önskvärda egenskaper kan effektivisera urvalsprocessen under förädlingsgången och ge ledning i valet av föräldrar till nya korsningar. Det finns flera metoder att finna markörer med koppling till egenskaper av intresse. En sådan populär metod är associationskartering som bygger på att man med statistiska metoder finner associationer mellan molekylära markörvarianter och fenotypiska egenskaper i en population av genetiskt olika individer (vilka i sammanhang som detta kan vara lantraser, förädlingslinjer eller sorter).

Associationskartering har använts i ett flertal grödor och för att finna markörer kopplade till en lång rad egenskaper. Denna studie är en av de första av sitt slag i den för nordiskt lantbruk viktiga grödan havre.

Vi har här använt oss av s.k. DArT-markörer och funnit flera sådana markörer kopplade till egenskaper som utgör centrala mål i utvecklingen av förbättrade sorter. En svaghet i detta sammanhang är att de flesta DArT-markörer inte är karterade, dvs. de har inte en väldefinierad position i havrens arvs massa. Detta medför att det i allmänhet inte är möjligt att med t.ex. en jämförande analys av mer välundersökta, besläktade arter (som majs, korn, vete, ris och *Brachypodium*) identifiera de faktiska gener som kontrollerar de undersökta egenskaperna. Å andra sidan behöver det i ett växtförädlings sammanhang inte vara nödvändigt att ha detaljerad information om de bakomliggande generna; det räcker oftast med att ha tillgång till markörer med koppling till en egenskap och vars nedärvning i förädlingsmaterialet man enkelt kan följa.

För att associationskartering skall ge relevant information om kopplingen mellan markör och egenskap krävs att man har kunskap om hur långt kopplingsojämvikten (LD) sträcker sig inom genomet. Associationskartering i sig bygger på att man kan fånga upp LD mellan

markörer och de gener som styr egenskaperna man studerar. Om LD har kort räckvidd krävs många markörer med kort inbördes avstånd; är detta för glest kan man mellanliggande QTL undgå upptäckt. Har LD däremot lång räckvidd behövs färre markörer, men å andra sidan blir de QTL man upptäcker väldigt ”breda”, dvs. de täcker ett stort område i genomet i vilket flera hundra gener kan vara belägna. I den undersökta kollektionen av havresorter sjunker LD till under 0,2 vid ett avstånd på 1,1 cM. Detta betyder att man riskerar att inte upptäcka QTL mellan två markörer på ett inbördes avstånd av $>2,2$ cM. Havrens genom omfattar ca 2000 cM, vilket betyder att det genomsnittliga avståndet mellan de 747 markörerna är ca 2,7 cM. Detta skulle således vara en för låg densitet för att tillåta upptäckt av flertalet QTL. I verkligheten är dock DArT-markörerna ojämnt utspridda i genomet med flera kluster av tätt kopplade markörer. Det är troligt att det framförallt är i sådana områden som vi funnit QTL. Detta kan också förklara varför vi inte hittat några markörer för flygsot- och nematodresistens trots stor variation i materialet: de gener som eventuellt styr dessa egenskaper kan finnas i regioner av arvsmassan som vi inte lyckats täcka in med DArT-markörer.

De markörer som uppvisar signifikanta associationer med någon egenskap har oftast en relativt liten effekt på den fenotypiska variationen hos egenskapen. Det är ett vanligt resultat i associationsstudier; förekomsten av s.k. ”major genes” med stor effekt är ofta mer sällan förekommande. Vidare har vi funnit att vissa markörer är associerade med flera egenskaper. Ett typiskt mönster i denna studie är markörer med associationer till avkastning, kärnhalt, kärnvikt och rymdvikt. Avkastning är en komplex egenskap och det är troligt att den påverkas av egenskaper som kärnhalt, kärnvikt, etc. och att vi därför ser markörer kopplade till flera av dessa egenskaper.

Konventionell användning av enstaka markörer i en förädlingssituation kräver att de egenskaper man vill selektera för är kopplade till en eller ett fåtal markörer med stor effekt. Det är få av de undersökta egenskaperna i denna studie som är associerade med markörer med stor effekt, vilket skulle betyda att få av de undersökta markörerna har någon praktisk användning. Emellertid har nya metoder – s.k. genomic selection – med ursprung i husdjursaveln börjat tillämpas i växtförädlingen. Dessa metoder utgår inte från enstaka markörer/QTL med stor fenotypisk effekt, utan inkluderar alla markörer med någon effekt och summerar över dessa för att konstruera en modell med vilken det är möjligt att prediktera en individs fenotyp utifrån markördatabaser. Vi har nu börjat använda resultaten från detta projekt för att anpassa Lantmännens havreförädlingsprogram till dessa nya metoder.

Publikationer

Vallenback P, Rönnberg-Wästljung AC, Ceplitis A. 2012. Genetic variation, population structure and linkage disequilibrium in a global sample of cultivated oats (*Avena sativa*) using DArT markers. Proceeding of the 15th EWAC-Eucarpia Cereals Section Conference (eds. Börner A, Kobijski B), pp. 95-97.

Därutöver presentationer vid 9th International Oat Conference, Beijing 2012.

Publikationer (2 st) är under beredning.

Slutsatser

I detta projekt har vi genom associationskartering i ett mycket variabelt havrematerial sökt finna genetiska markörer kopplade till egenskaper som representerar viktiga förädlingsmål. Sådana egenskaper är ofta kontrollerade av ett stort antal gener, var och en med relativt liten effekt på fenotypen. I en sådan situation kan det vara svårt att finna enstaka markörer associerade med tydliga fenotypiska skillnader och som skulle kunna användas i ett konventionellt markörstött förädlingsprogram. Däremot visar det sig att ett material som detta kan fungera som en utgångspunkt för s.k. genomic selection, där information från samtliga markörer används. Det är således troligt att en framtida markörstödd förädling kräver ett delvis annorlunda angreppssätt än vad som hittills varit förhärskande.

Resultatförmedling till näringen

Detta projekt är speciellt i den bemärkelsen att det har genomförts av den del av näringen som också skall använda resultaten (dvs. växtförädlingsindustrin). Projektets ursprungliga syfte, att tillhandahålla genetiska markörer för användning i havreförädlingen har infriats, om än inte på det sätt som förutsågs vid projektets start. Resultaten från projektet har i verkligheten inneburit en radikal omläggning av havreförädlingen vid Lantmännen, från ett traditionellt förädlingsupplägg till ett baserat på genomic selection. Detta har fått effekter också på andra förädlingsprogram och lett till ett kraftigt ökat fokus på genom-baserade metoder i växtförädlingen, med flera spin-off projekt i andra grödor. Betydelsen av detta har varit så stor att den post-doc som rekryterades till projektet nu erbjudits en tillsvidareanställning inom Lantmännens växtförädling. Genom detta projekt har Lantmännen dessutom vunnit en ökad visibilitet i forskarvärlden och intressanta samarbeten har inletts inom havreförädlingen, särskilt med forskare i Nordamerika. Utan tvekan har projektet fått en betydelse vars omfattning inte gick att förutse från början, men som nu svårigen kan överskattas.

Svalöv 10 juni 2014

Alf Ceplitis