Christiane Gebhardt,* Li Li,
Karolina Pajerowska-Mukthar,
Ute Achenbach, Amirali Sattarzadeh,
Christina Bormann, Evgeniya
Ilarionova, and Agim Ballvora

Candidate Gene Approach to Identify Genes Underlying Quantitative Traits and Develop Diagnostic Markers in Potato

PMax-Planck Institute for Plant Breeding Research,
Carl von Linné Weg 10, 50829 Köln, Germany.
*Corresponding author (gebhardt@mpiz-koeln.mpg.de).

Published in Crop Sci. 47(S3) S106–S111 (2007).
doi: 10.2135/cropsci2007.04.0005IPBS

677 S. Segoe Rd., Madison, WI 53711 USA

Potatoes (*Solanum tuberosum* L.) and humans (*Homo sapiens* L.) are both outcrossing species. The phenotypic variation of both is controlled by environmental factors and by natural DNA polymorphisms between individuals. Therefore we adopt similar approaches as used in human population genetics, such as association mapping, to identify loci and their alleles that are causal for complex agronomic characters such as quantitative resistance to pathogens or tuber sugar content. Functional analysis of genes operating in resistance to pests and pathogens or in carbohydrate metabolism, either in potato itself or in other plants including the model species *Arabidopsis thaliana* (L.) Heynh., provides many functional candidates for these complex agronomic traits. In our approach, functional candidate genes are tested for linkage to quantitative trait loci (QTL) for pathogen resistance or tuber quality traits, thereby selecting positional candidates (genes colocalizing with a QTL). DNA polymorphisms in or physically linked to positional candidate genes are then evaluated in populations of tetraploid potato genotypes and tested for association with phenotypes evaluated in the same populations. We have used the candidate gene approach to identify DNA-based markers that are diagnostic for quantitative resistance to late blight caused by *Phytophthora infestans* (Mont.) de Bary, to the root cyst nematode *Globodera pallida* (Stone), and for chip color in tetraploid potato cultivars and in advanced breeding clones.

With more than 300,000 million megagrams produced per year, potato (*Solanum tuberosum* L.) is the fourth most important crop worldwide, after maize (*Zea mays* L.), rice (*Oryza sativa* L.), and wheat (*Triticum aestivum* L.) (Graf, 2002). Potatoes are cultivated for direct consumption, for processed food products such as chips and French fries, for industrial production of starch and starch derivatives, and for animal feed. Potatoes may also be used in the future as bioreactors for producing specific organic molecules (e.g., Börnke et al., 2002). Potato starch may gain higher importance as a biopolymer substitute for oil-based chemicals. The breeding of new potato cultivars aims at optimizing the crop for three main areas: the fresh market, the processing industry, and the starch industry. Approximately 40 different characteristics are considered in the selection process. Most of these characteristics are complex, meaning that they show quantitative variation, which is controlled by an unknown number of genes and by environmental factors. The most relevant characteristics considered in central Europe follow (Ross, 1986).

Tuber Yield, Starch Content, and Starch Yield
Suitability for transport (bruising), storage (cold-sweetening), and processing (chip color)
Tuber shape, eye depth, flesh color
Cooking quality, taste
Maturity type (earliness)
Resistance to pathogens: *Phytophthora infestans* (Mont.) de Bary (late blight), viruses (*Potato virus Y*, *Potato leaf roll virus*), bacteria [*Erwinia carotovora* (Jones) Dye], parasitic nematodes (*Globodera* spp.).

Human Population Genetics as Model for Potato Association Studies

The cultivated potato is tetraploid, non-inbred, and vegetatively propagated by tubers. Polyploidy and inbreeding depression prevent the generation of pure lines. When the ploidy level is reduced from 4*n* to 2*n* (Hermsen and Verdenius, 1973; Powell and Uhrig, 1987), the diploid potatoes become self incompatible. Potato genotypes at all ploidy levels, particularly all advanced breeding materials, are therefore usually heterozygous (Gebhardt, 2004). A feature common between potato and humans is that both species are out-crossing and consequently, heterozygous parents generate segregating F1 progeny. In both species, phenotypic and genotypic variation is abundant and natural. The concepts and approaches developed in human genetics based on the Human Genome Project for improving medical diagnosis and treatment serve therefore as model for developing innovative breeding strategies for the potato crop.

Abbreviations: EST, expressed sequence tag; PCR, polymerase chain reaction; QTL, quantitative trait loci; SNP, single nucleotide polymorphism.

In human population genetics, association studies using DNA-based markers are considered a feasible strategy for elucidating the molecular basis of complex inherited diseases. Individuals related by descent are phenotyped, for example, for disease symptoms. The same individuals are also genotyped with DNA-based markers. Appropriate statistics are then used to identify associations between specific marker alleles and symptom severity. The association between DNA variation and disease phenotype allows genetic risk assessment and, eventually, personalized therapy (Schafer and Hawkins, 1998; Risch, 2000). Similarly, the finding of association between DNA variation and complex agronomic characteristics of crops in general and potato in particular will allow the assessment of genetic potential in populations of cultivars, breeding lines, and wild species that are related by descent. DNA variants associated with agronomic performance will provide excellent diagnostic tools for marker-assisted breeding and facilitate the identification and molecular cloning of genes controlling complex agronomic characters.

The Candidate Gene Approach

There are two approaches to identify DNA-based markers associated with quantitative traits. In the genome-wide approach, populations of individuals related by descent are genotyped with DNA-based markers covering the whole genome, ideally positioned at regular intervals on the physical map. The genome-wide approach essentially requires a sequenced genome and a large number of markers and is therefore expensive and statistically complex (Hirschhorn and Daly 2005). It is unbiased, however, as no hypotheses have to be made about identity and function of the genes underlying the trait. In the candidate gene approach, genotyping is targeted to functional and positional candidate genes (Pflieger et al., 2001). Functional candidates are genes that have been shown or are suspected to have a functional role in the phenotype of interest, for example, quantitative resistance to a pathogen. Allelic variants of such genes may be causal for the observed natural trait variation. In this case, DNA polymorphisms located within the candidate gene or physically close to it will be associated with trait variation. To select the most promising candidates from a large number of functional candidate genes, gene sequences are tested for linkage to QTL for the trait of interest by molecular mapping, thereby identifying positional candidates (genes colocalizing with a QTL) (Pajerowska et al., 2005). The candidate gene approach is greatly facilitated by plant genomics resources such as expressed sequence tag (EST) databases and analysis of gene function in model organisms. The candidate gene approach depends on and takes advantage of the knowledge available in the literature and databases on physiology, biochemistry and molecular genetics of a trait of interest. It is biased, in that it makes hypotheses about the

identity and function of the genes underlying the trait, but does not require a sequenced genome and is feasible with limited human and financial resources.

We therefore adopted the candidate gene approach to explore the possibilities of finding marker–trait associations in populations of tetraploid potato cultivars and breeding clones. Our first examples and experiences are outlined below.

Markers Associated with Resistance to Late Blight

A gene bank collection of 600 potato cultivars bred between 1850 and 1990 in different countries was genotyped with four DNA-based markers from a singular genomic region on potato chromosome V, which has been previously shown to harbor QTL for resistance to late blight caused by the oomycete *Phytophthora infestans*. The same region also contains major genes (*R* genes) for pathogen resistance, among those *R1* for resistance to late blight (reviewed in Gebhardt and Valkonen 2001). *R1* is one member of a clustered family of nucleotide binding–leucine rich repeats type genes (Ballvora et al., 2002), a gene class to which most known plant resistance genes belong. By position as well as function, the *R1* gene family is a candidate for the quantitative resistance factors in the "hot spot" for pathogen resistance on potato chromosome V. Specific DNA fragments from this region were tested for association with quantitative resistance to late blight based on passport evaluation data that were available for approximately 400 genotypes. Highly significant association with quantitative resistance of foliage and tubers to late blight was detected with a polymerase chain reaction (PCR) marker specific for *R1* and with two anonymous PCR markers flanking the *R1* locus at 0.2-cM genetic distance. The marker alleles associated with increased resistance were traced to an introgression from the wild species *S. demissum* Lindl. These DNA markers are the first that have some general diagnostic value for quantitative resistance in a large collection of cultivars (Gebhardt et al., 2004).

Allele Specific Marker Diagnostic for Nematode Resistance

Globodera pallida (Stone) is a parasitic root cyst nematode that causes reduction of crop yield and quality in infested potato fields. Genes for resistance to *G. pallida* have been introgressed into the cultivated potato gene pool from the wild, tuber bearing *Solanum* species *S. spegazzinii* Bitt. and *S. vernei* Bitt. and Wittm. Resistance to this nematode is quantitative. Selection of resistant genotypes in breeding programs is hampered by the fact that the phenotypic evaluation of resistance to *G. pallida* is time consuming, costly, and often ambiguous. We identified single nucleotide polymorphisms (SNPs) that were linked to a major QTL for nematode resistance located in the same resistance "hot spot" on potato chromosome V as QTL for late blight resistance. One SNP marker was converted into an allele specific PCR assay and was tested in 34 resistant potato cultivars bred in the Netherlands and Germany and in 22 susceptible cultivars. The marker was only detected in cultivars with high levels of nematode resistance and was absent in susceptible cultivars. It has therefore high predictive value for resistance to *G. pallida* present in advanced breeding materials (Sattarzadeh et al., 2006).

Invertase Alleles Associated with Chip Color and Tuber Starch Content

Starch and sugar content of potato tubers are models for the candidate gene approach to identify the molecular basis of QTL, because the biochemistry and molecular genetics of plant carbohydrate metabolism is well studied. Sugar content is important for the quality of processed products such as chips and French fries. High content of reducing sugars glucose and fructose results in inferior chip quality. Tuber starch content affects nutritional quality. On one hand, mapping genes known to play a role in carbohydrate metabolism and transport (Chen et al., 2001) and, on the other hand, QTL for tuber starch and sugar content (Schäfer-Pregl et al., 1998, Menendez et al., 2002) resulted in 20 to 30 candidate loci for tuber quality traits (Figure 1). One of them is a locus on potato chromosome IX, which encodes two tandem duplicated invertase genes (Maddison et al., 1999). The *invGE*/*GF* locus colocalizes with a cold-sweetening QTL (Menendez et al., 2002). DNA variation at *invGE*/*GF* was analyzed in 188 tetraploid potato cultivars obtained from the potato breeding companies Saka-Ragis Pflanzenzucht (Windeby, Germany), Böhm-Nordkartoffel Agrarproduktion (Ebstorf, Germany), and NORIKA (Groß Lüsewitz, Germany). These cultivars have been previously assessed for chip quality and tuber starch content. Two closely correlated invertase alleles, *invGE-f* and *invGF-d*, were associated with better chip quality in the breeding populations. Allele *invGF-b* was associated with lower tuber starch content (Li et al., 2005). Interestingly, the potato invertase gene *invGE* is orthologous to the tomato (*Solanum lycopersicum* L.) invertase gene *Lin5,* which is causal for the fruit sugar yield QTL *Brix9-2-5* (Fridman et al., 2000), suggesting that natural variation of sugar yield of tomato fruits and sugar content of potato tubers is controlled by functional variants of orthologous invertase genes. This demonstrates that the same well-known gene can be responsible for functionally related but physiologically different QTL in related plant species.

Conclusion and Outlook

Marker-assisted breeding was so far of limited relevance in commercial potato breeding programs due to the following handicaps. By reason of feasibility, linkage analysis of quantitative and qualitative agronomic traits in potato using

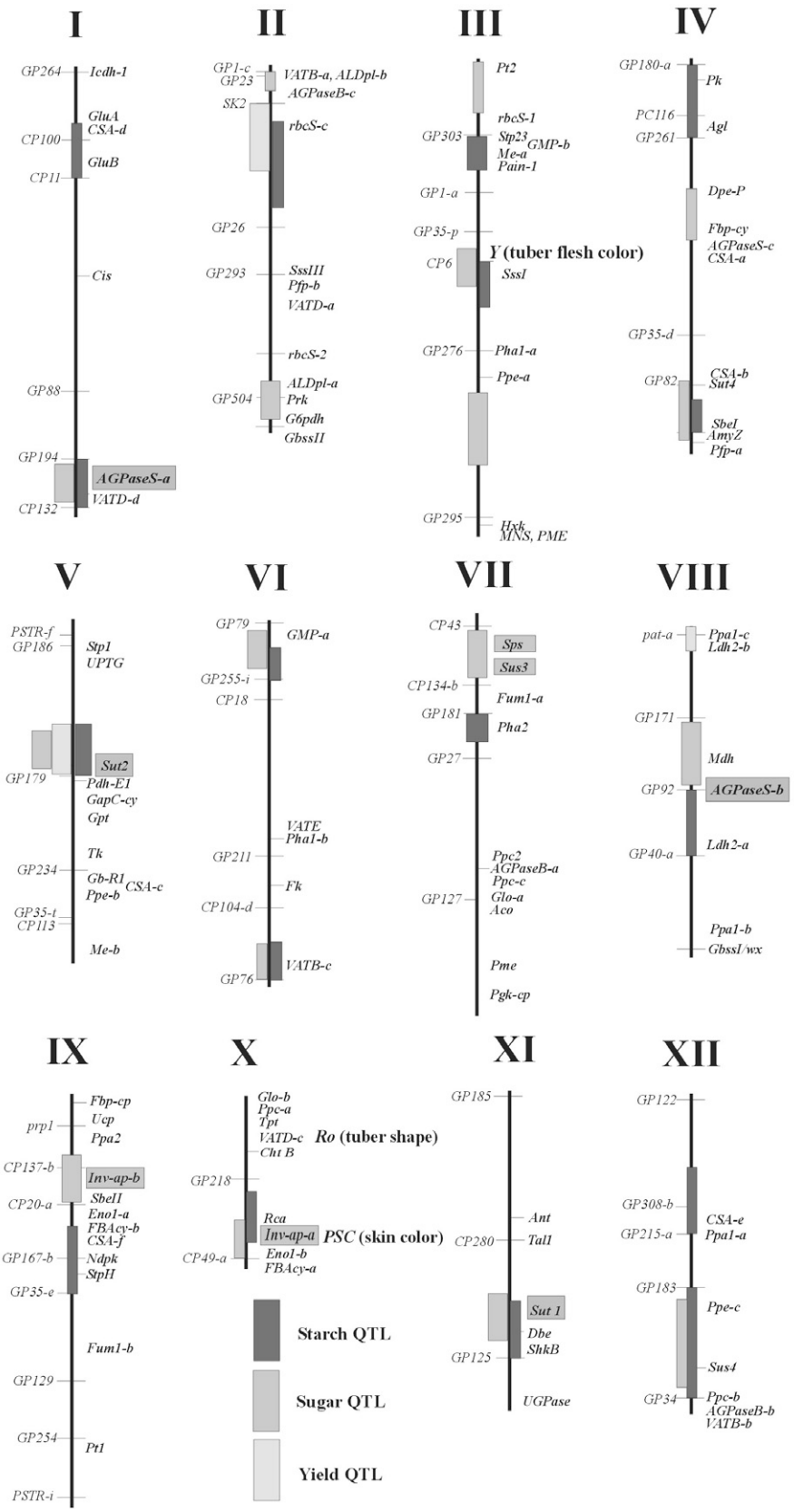

Figure 1. Coming together: QTL and candidate genes. The figure shows a potato function map for tuber quality traits. It compiles published positions of QTL for tuber starch and sugar content, tuber yield and of candidate genes (genes functional in carbohydrate metabolism and transport). Details in Schäfer-Pregl et al. (1998), Chen et al. (2001) and Menendez et al. (2002).

DNA-based markers is mostly being performed in experimental, diploid, often interspecific mapping populations. This material is not well adapted and its general agronomic qualities are inferior to tetraploid, advanced breeding clones and cultivars. DNA markers linked to an agronomic trait in an experimental mapping population are not immediately diagnostic for the same trait in breeding schemes that consist of intercrossing multiple heterozygous parents instead of backcross and pure line breeding. Association genetics is the solution to this problem. As our examples show, it can be done directly in populations of advanced, tetraploid breeding materials, and marker alleles associated with a quantitative trait evaluated in such populations have direct diagnostic value. To date, there are only a handful of such markers described in potato, all identified based on a candidate gene approach (Gebhardt et al., 2004; Simko et al., 2004; Li et al., 2005; Sattarzadeh et al., 2006). This number will undoubtedly increase in the coming years and this will have a hopefully positive impact on the way new cultivars are developed in this "difficult" polyploid crop. The ultimate markers will be diagnostic for the functional variants of the genes that underlie quantitative agronomic traits, allowing "breeding by design." This is where plant genome research can play an important role. Truly complementary collaboration between breeders, the experts in assessing the phenotype, and molecular geneticists, the experts in assessing genes at the molecular level, is required to achieve this goal.

References

Ballvora, A., M.R. Ercolano, J. Weiß, K. Meksem, C. Bormann, P. Oberhagemann, F. Salamini, and C. Gebhardt. 2002. The *R1* gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. Plant J. 30:361–371.

Börnke, F.M., M. Hajirezaei, and U. Sonnewald. 2002. Potato tubers as bioreactors for palatinose production. J. Biotechnol. 96:119–124.

Chen, X., F. Salamini, and C. Gebhardt. 2001. A potato molecular function map for carbohydrate metabolism and transport. Theor. Appl. Genet. 102:284–295.

Fridman, E., T. Pleban, and D. Zamir. 2000. A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proc. Natl. Acad. Sci. USA 97:4718–4723.

Gebhardt, C. 2004. Potato genetics: Molecular maps and more in biotechnology in agriculture and forestry. Vol. 55. p. 215–227. *In* H. Lörz and G. Wenzel (ed.) Molecular marker systems. Springer-Verlag, Berlin, Heidelberg.

Gebhardt, C., A. Ballvora, B. Walkemeier, P. Oberhagemann, and K. Schüler. 2004. Assessing genetic potential in germ plasm collections of crop plants by marker-trait association: A case study for potatoes with quantitative variation of resistance to late blight and maturity type. Mol. Breed. 13:93–102.

Gebhardt, C., and J.P.T. Valkonen. 2001. Organization of genes controlling disease resistance in the potato genome. Annu. Rev. Phytopathol. 39:79–102.

Graf, G. 2002. Die Weltkartoffelernte 2001. (In German.) Kartoffelbau 11:440–442.

Hermsen, J.G.T.H., and J. Verdenius. 1973. Selection from *Solanum tuberosum* group phureja of genotypes combining high-frequency haploid induction with homozygosity for embryo spot. Euphytica 22:244–259.

Hirschhorn, J.N., and K.J. Daly. 2005. Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. 6:95–108.

Li, L., J. Strahwald, H.-R. Hofferbert, J. Lübeck, E. Tacke, H. Junghans, J. Wunder, and C. Gebhardt. 2005. DNA variation at the invertase locus *invGE/GF* is associated with tuber quality traits in populations of potato breeding clones. Genetics 170:813–821.

Maddison, A.L., P.E. Hedley, R.C. Meyer, N. Aziz, D. Davidson, and G.C. Machray. 1999. Expression of tandem invertase genes associated with sexual and vegetative growth cycles in potato. Plant Mol. Biol. 41:741–751.

Menendez, C.M., E. Ritter, R. Schäfer-Pregl, B. Walkemeier, A. Kalde, F. Salamini, and C. Gebhardt. 2002. Cold-sweetening in diploid potato. Mapping QTL and candidate genes. Genetics 162:1423–1434.

Pajerowska, K.M., J.E. Parker, and C. Gebhardt. 2005. Potato homologues of *Arabidopsis thaliana* genes functional in defense signaling—Identification, genetic mapping and molecular cloning. Mol. Plant. Microbe Interact. 18:1107–1119.

Pflieger, S., V. Lefebvre, and M. Causse. 2001. The candidate gene approach in plant genetics: A review. Mol. Breed. 7:275–291.

Powell, W., and H. Uhrig. 1987. Anther culture of *Solanum* genotypes. Plant Cell Tissue Organ Cult. 11:13–24.

Risch, N.J. 2000. Searching for genetic determinants in the new millennium. Nature 405:847–856.

Ross, H. 1986. Potato breeding—Problems and perspectives. J. Plant Breeding Suppl. 13. Parey, Berlin and Hamburg.

Sattarzadeh, A., U. Achenbach, J. Lübeck, J. Strahwald, E. Tacke, H.-R. Hofferbert, T. Rothsteyn, and C. Gebhardt. 2006. Single nucleotide polymorphism (SNP) genotyping as basis for developing a PCR-based marker highly diagnostic for potato varieties with high resistance to *Globodera pallida* pathotype Pa2/3. Mol. Breed. 18:301–312.

Schäfer-Pregl, R., E. Ritter, L. Concilio, J. Hesselbach, L. Lovatti, B. Walkemeier, H. Thelen, F. Salamini, and C. Gebhardt. 1998. Analysis of quantitative trait loci (QTL) and quantitative trait alleles (QTA) for potato tuber yield and starch content. Theor. Appl. Genet. 97:834–846.

Schafer, A.J., and J.R. Hawkins. 1998. DNA variation and the future of human genetics. Nat. Biotechnol. 16:33–39.

Simko, I., S. Costanzo, K.G. Haynes, B.J. Christ, and R.W. Jones. 2004. Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. Theor. Appl. Genet. 108:217–224.