

Isolated populations and complex disease gene identification

Kati Kristiansson^{*†}, Jussi Naukkarinen^{*‡} and Leena Peltonen^{*†‡}

Addresses: ^{*}National Public Health Institute and FIMM, Institute for Molecular Medicine Finland, Helsinki 00300, Finland. [†]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. [‡]Department of Medical Genetics, University of Helsinki, Helsinki 00014, Finland.

Correspondence: Leena Peltonen. Email: leena.peltonen@sanger.ac.uk

Published: 26 August 2008

Genome Biology 2008, **9**:109 (doi:10.1186/gb-2008-9-8-109)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2008/9/8/109>

© 2008 BioMed Central Ltd

Abstract

The utility of genetically isolated populations (population isolates) in the mapping and identification of genes is not only limited to the study of rare diseases; isolated populations also provide a useful resource for studies aimed at improved understanding of the biology underlying common diseases and their component traits. Well characterized human populations provide excellent study samples for many different genetic investigations, ranging from genome-wide association studies to the characterization of interactions between genes and the environment.

Over the past few years, understanding how genetic variation in individuals and in populations contributes to the biological pathways involved in determining human traits and mechanisms of disease has become a reachable goal for genetic research. Following on from the achievements in molecular studies of monogenic disorders, recent studies have used strategies of hypothesis-free fine mapping of genes and loci to identify underlying factors in common complex diseases with major impacts on public health. These diseases, which include cancers, coronary heart disease, schizophrenia, autism and multiple sclerosis, arise from complex interactions between environmental factors and variation in several different genes. Until recently, detection of the genes underlying these diseases met with only limited success, but the past two years have witnessed the identification of more than 100 well established loci. These successes mainly involved the collection of very large study cohorts for any individual trait and international collaborations on an unprecedented scale [1].

The detection of genes underlying common complex diseases might not always need large global population samples. Samples of individuals from genetically isolated populations, or 'population isolates', have already proved immensely useful in the identification of rare recessive disease genes. Such genes are only detectable in isolated

populations with a limited number of founders, where rare disease alleles are enriched, thus resulting in homozygote individuals affected by the disease. Impressive accomplishments in disease-locus mapping and gene identification using genome-wide scans of only a handful of affected individuals in such populations have been reported, typically based on linkage analyses and homozygosity scanning [2,3]. It is becoming increasingly apparent that studies locating genes underlying complex phenotypes also benefit from the study of samples from homogeneous populations with a limited number of founders - 'founder populations' (Table 1).

Success stories from population isolates

One of the most impressive examples of the resourceful use of known genealogy, large extended families and vast amounts of medical data in genetic studies is provided by the company deCODE genetics in Iceland, where more than 50% of the adult population have volunteered their medical and genetic information to be used in genetic research [4,5]. Although the Icelandic population does not represent a population isolate as conventionally defined, genetic drift over generations has reduced the amount of variation within it relative to the rest of Europe [6]. This, among other benefits of a geographically isolated population, has enabled the identification by means of linkage, and more recently by

Table 1**Recent genetic studies of complex diseases and traits in special populations**

Complex disease/trait	Gene/locus	Population/isolate	Reference
Affective disorders	Several loci	Northern Sweden	[34]
Asthma	IRAK-M (interleukin-1 receptor-associated kinase M)	Sardinia	[74]
Asthma	CHI3L1 (chitinase 3-like 1)	Hutterites	[25]
Asthma	NPSR1 (neuropeptide S receptor 1)	Finland	[24]
Atrial fibrillation	4q25	Iceland	[11]
Bipolar disorder	5q31-34	Antioquia (Colombia), Central Valley of Costa Rica	[29]
Bone mineral density	Several loci	Iceland	[75]
Breast cancer	5p12, 2q35, 16q12	Iceland	[76,77]
Myocardial infarction	9p21	Iceland	[8,78]
Coronary heart disease	8p22	French Canadians	[79]
Crohn's disease	Several loci	French Canadians	[80]
Fasting glucose levels	G6PC2 (glucose-6-phosphatase, catalytic, 2)/ ABCB11 (ATP-binding cassette, sub-family B (MDR/TAP), member 11) region	Finland, Sardinia	[81]
Exfoliation glaucoma	LOXLI (lysyl oxidase-like 1)	Iceland	[82]
Height	Several loci	Iceland	[13]
Height	GDF5 (growth differentiation factor 5)/ UQCC (ubiquinol-cytochrome c reductase complex chaperone) locus	Finland, Sardinia, Amish	[83]
Nicotine dependence and smoking-related diseases	15q24	Iceland	[84]
Obesity	FTO (fat mass and obesity associated), PFKF (phosphofructokinase, platelet)	Sardinia	[28]
Parkinson's disease	GBA (β -glucocerebrosidase)	Ashkenazi Jews	[37]
Pigmentation	Several genes	Iceland	[14,15]
Prostate cancer	Xp11.22, 2p15, 17q	Iceland	[12,85]
Psychotic and bipolar spectrum disorders	TSNAX (translin-associated factor X)/ DISC1 (disrupted in schizophrenia 1) locus	Finland	[86]
Psychosis	TGIF (TGFB-induced factor)	Central Valley of Costa Rica	[87]
Schizophrenia	DRD2 (dopamine D2 receptor)	Basques	[30]
Type 2 diabetes	CDKALI (CDK5 regulatory subunit associated protein 1-like 1)	Iceland	[10]

genome-wide association (GWA) studies, of an impressive number of variants contributing to the development of common/complex disease [5]. Among these are gene loci for myocardial infarction and stroke (*ALOX5AP* and chromosomal region 9p21) [7,8], type 2 diabetes (*TCF7L2* and *CDKALI*) [9,10], atrial fibrillation (4q25) [11] and prostate cancer (2p15 and Xp11.22) [12]. In addition to disease genes, the Icelandic population has revealed genes contributing to a number of complex traits, such as adult stature (several loci,

including *ZBTB38*) [13] as well as skin and hair pigmentation (*SLC24A4*, *KITLG*, *TYR*, *OCA2*, *MC1R* and 6p25.3) [14,15]. The continuing work by deCODE genetics on 50 common diseases is sure to result in a slew of additional gene findings and help to characterize the allelic spectrum of disease-predisposing variants. The wisely designed strategy of fully harvesting the unique population and the combined power of linkage and association has been the basis of the success of genetic research in Iceland.

Another population isolate with proven value in gene mapping is the population of Finland, where genes for 35 monogenic diseases that are more frequent than in other populations have been identified [16]. Features of the Finnish population have also been an advantage in studies of schizophrenia spectrum disorders: a balanced translocation (1;11)(q42.1;q14.3) segregating with schizophrenia was first described in a large Scottish family [17] and evidence for association of the gene *DISC1* with the disorder was subsequently obtained in Finnish families with diagnosed schizophrenia [18,19]. Large pedigrees from the Finnish population were also used successfully in a study of familial combined hyperlipidemia that identified the gene for upstream stimulatory factor 1 (*USF1*) as a risk factor for this complex disease [20]. This association was subsequently replicated in other populations, and evidence of the functional significance of the gene variants and their association with cardiovascular disease and dyslipidemia at the population level has also been obtained [21-23]. Another excellent example from Finland is a gene conferring susceptibility to asthma (*NPSR1*), discovered in Kainuu and North Karelia subpopulations of Eastern Finland representing regions of the late-settlement [24].

The communal lifestyle and genetic isolation of the Hutterites, who live in the northern United States and western Canada, have especially aided studies of asthma and related traits [25]. Recently, the chitinase 3-like 1 gene was identified as an asthma-susceptibility gene in Hutterites, and the finding subsequently replicated in two population cohorts of European descent [25]. Studies of type 2 diabetes and obesity have used Pima Indians [26], as well as other genetic isolates, such as Finland and Sardinia [27,28]. Genes contributing to neuropsychiatric disorders are sought, and previous gene discoveries are confirmed, in studies of special populations, such as people from the Antioquia in Colombia and the Central Valley of Costa Rica [29], Basques from Spain [30], the Micronesian population of the islands of Kosrae [31] and Palau [32], Bulgarian Gypsies [33], and sub-isolates from Sweden [34] and Israel [35]. Other special populations utilized in recent genetic studies of complex diseases include French Canadians [36], Ashkenazi Jews [37], Mennonites [38], Newfoundlanders [39], sub-isolates from the Netherlands [40] and the Amish [41].

The important observation from all these studies is that the genetic variants identified within isolates and/or exceptional families seemingly segregating a common disease in a near-Mendelian fashion are not restricted therein, but are being replicated in large-scale population samples and uncovering new pathways behind these disease processes.

Reduced haplotype complexity

The increasing information in public databases on single nucleotide polymorphisms (SNPs) and their

haplotype-tagging properties [42-44] as well as advances in genome-wide data collection using advanced technology platforms [45] have facilitated the recent deluge of studies utilizing the genome-wide SNP-association strategy to identify loci influencing disease phenotypes. This GWA approach is essentially 'hypothesis free'. It circumvents the necessity of understanding disease pathogenesis, which has previously guided studies of candidate genes selected for their biological relevance. In a GWA study, a dense set of SNPs totaling up to 1 million across the genome is genotyped using a standard platform and tested for association with a disease or quantitative trait. Successful gene identification by GWA studies, which operates very much under the common-disease, common-variant hypothesis, requires that the susceptibility variant itself, or a variant highly correlated with it, is among the markers typed.

As a result of the International HapMap Project [44], the linkage disequilibrium (LD) patterns of most genomic regions are known and SNP genotyping platforms have been designed to detect a restricted number of haplotype-tagging variants with the hypothesis that they should capture most of the common variation within genomic regions [46,47]. Ultimately, the LD structure of each study population determines the number of genotyped SNPs needed for complete coverage in a GWA study.

Several studies have been undertaken to characterize differences in the magnitude and distribution of LD in global populations [48-51]. Even though the density of SNPs required for 100% coverage of the genome in whole-genome genotyping efforts in various global populations remains unknown, on the basis of the size of LD blocks in 'young isolates', populations that are relatively recently (less than 2,000 years ago) inhabited or isolated, it has been concluded that GWA studies in populations such as that of Finland, the Dutch isolate referred to above, Costa Rica, Antioquia, Sardinia or the Ashkenazim require some 30% fewer markers than in more outbred populations, and that the current GWA panels provide excellent genome-wide coverage with a very small number of gaps (Figure 1) [48]. In an isolated population there are a potentially fewer number of haplotypes being segregated through the population and the haplotype-tagging SNPs should also be able to detect those haplotypes that carry more rare alleles. In a more outbred population with considerably higher numbers of haplotypes for a given locus, the causative allele is more likely to be located on several haplotypic backgrounds, thereby diluting its signal to an extent that precludes its identification by genetic means. The value of population isolates and their genomic LD patterns may thus be even greater when lower-frequency (less than 5%) variants are considered [52].

The problem of GWA studies carried out in genetic isolates is that the strong LD that initially helped identify the disease locus may in the end hamper efforts to distinguish the

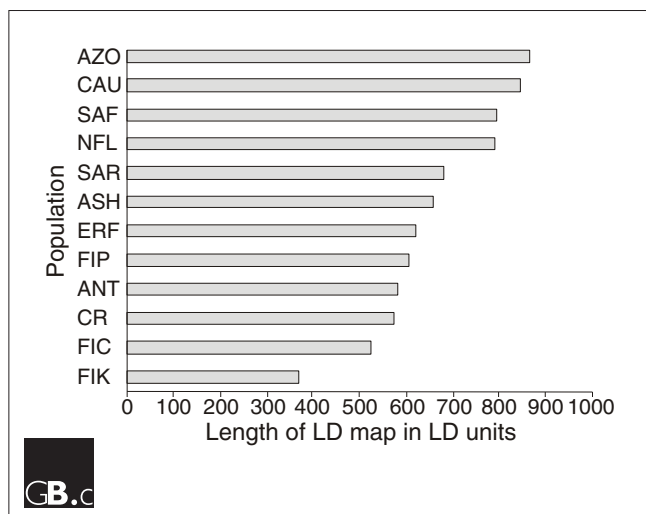


Figure 1

Considerable differences in LD map length across populations. The length of the LD map in LD units (as defined in [88]) in 12 different population samples is depicted in order of decreasing map length. AZO, Azores; CAU, outbred European-derived sample; SAF, Afrikaner; NFL, Newfoundland; SAR, province of Nuoro in Sardinia; ASH, Ashkenazi; ERF, a village in southwestern Netherlands; FIP, Finland nationwide; ANT, Antioquia; CR, Central Valley of Costa Rica; FIC, early-settlement Finland; FIK, Finnish sub-isolate of Kuusamo. Adapted from [48].

biologically relevant variants from insignificant polymorphisms in complete LD with them. Comparing the GWA data across isolates from different populations should help pin down the potential causative variants for functional studies.

Restricted allelic and locus heterogeneity

Extensive allelic and locus heterogeneity, a key feature of common complex diseases, can obscure the association signal within disease-associated genomic regions. This problem is reduced in population isolates. When combined with geographic isolation that prevents the influx of new alleles, genetic drift acts to randomly raise some alleles to fixation and send others to extinction, thus reducing heterogeneity. A representative example of such drift, and of the founder effect, is the enrichment of various recessive diseases in founder populations, such as Ashkenazi Jews [53] and Finns [54], and an exceptionally low prevalence of other diseases in Finns, such as cystic fibrosis or phenylketonuria, which are common in other European populations. In founder populations, these recessive diseases are often characterized by a presence of one founder mutation, whereas numerous mutations in the same genes are identified in the global population [16]. Although allelic heterogeneity is expected to exist behind common diseases even in isolated populations, it is a reasonable expectation that the number of predisposing alleles will be more restricted than in more heterogeneous populations.

Furthermore, isolated populations may facilitate studies of the possible joint actions of associated gene loci as well as studies of the population effect of these associated markers, even before the actual causative variant has been identified. This may be possible as in isolated populations with a high degree of LD, the tagging of specific allele is more reliable than in heterogeneous populations in which broader allelic diversity of associated alleles can obscure these examinations.

In contrast to the 'gene-breaking' mutations underlying most monogenic diseases, variants that affect susceptibility to complex diseases are suggested to be ones that leave gene structure untouched and instead affect the dynamics of gene expression. Such variation can be situated in enhancer elements in the vicinity of the phenotype-causing genes or in the promoters of these genes where various transcription factors bind (*cis*-acting variants). SNPs elsewhere in the genome (*trans*-acting variants) may affect the phenotype via the function of the protein or RNA that the *trans*-acting gene encodes. These *cis*- and *trans*-acting variants account for much of the variation in gene expression between individuals. A good example of the identification of a trait-associated variant in a strong *cis*-regulatory element, using LD and samples from a population isolate, was the finding of the DNA variant behind lactose tolerance/intolerance: the variant was initially found among Finns and later confirmed to represent the common Caucasian mutation. This led to the identification of a regulatory DNA region with enrichment of mutations underlying the trait in numerous global populations [55].

Identification of rare variants

Susceptibility to common complex diseases probably involves the contribution of both common variants and rare mutations [56] and the relative significance of each in particular traits and disease phenotypes will have to be determined by large-scale resequencing studies of associated loci in large study samples. Whereas several common variants are likely to explain a substantial fraction of the heritable variation in complex traits, rare variants probably contribute significantly by having greater effects on the phenotype, as proposed for extreme lipid levels [57,58] (Figure 2). Furthermore, although rare variants are by definition rare by themselves, in a particular population there could exist a myriad of these variants and in combination they might explain a considerable proportion of the variance in a trait of interest [58]. Consequently, in addition to the interrogation of common polymorphisms, the rare variants implicated in many Mendelian diseases along with structural variation in the genome are now studied with increasing interest [59]. Identification of rare high-impact alleles may be of critical importance for our detailed understanding of the biology behind common diseases or traits.

A whole-genome strategy based on common haplotype-tagging SNPs is unlikely to be very successful in detecting

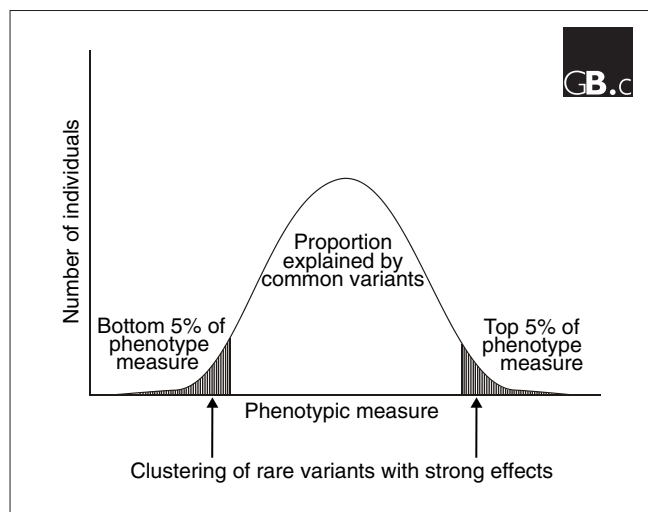


Figure 2

Contribution of rare and common variants to the distribution of a quantitative phenotype. Although common genetic variants explain the majority of the phenotypic variance in the population, the contribution of rare variants with strong effects may be observed at the extreme ends of the phenotypic distribution.

rare variants that increase disease susceptibility [60]. The statistical power to detect susceptibility alleles is positively correlated with the frequency and the penetrance of the allele. Even though detection of rare alleles with high penetrance is essentially as feasible as the detection of common alleles with more modest penetrance, it is unclear how well these rare variants are captured with the GWA arrays designed to tag common SNPs. Thus, while genome-wide association studies are likely to continue to identify the 'low-hanging fruit', study of linkage and association in exceptional families as well as in population isolates may be necessary to identify and define those risk alleles (the majority) that, although significant, are lost in the sea of peaks that fail to reach genome-wide significance in GWA studies as a result of their rarity or population-specific effect [60]. The founder effect, genetic bottlenecks and genetic drift have worked to increase the frequency of certain rare alleles in the population isolate, thus improving the power to detect those in genome-wide studies.

Notably, owing to founder effect and genetic drift, each genetic isolate typically has a unique profile of rare disease alleles [61]. Some rare variants that are readily detected in one population isolate may go unnoticed in others, necessitating the use of multiple isolates to get a picture of the full spectrum of variants with effects on phenotype [62]. Importantly, if the impact of the rare variants on the disease phenotype is really high, measuring them in a clinical setting might turn out to be of critical importance for 'family-specific' or personalized medicine, revealing individuals with the highest genetic risk. The existence of such population-

family-specific alleles is entirely possible - even expected - and personalized medicine just might become more personal than we ever dreamed of.

Population isolates help to minimize the environmental component of disease

In contrast to monogenic diseases, where the genetic composition of an individual often solely determines the disease phenotype, environmental factors are critical risk factors for complex diseases. The incidence and prevalence of many common diseases may vary between founder populations [63], and establishing whether this variation in disease incidence is the result of genetic background or of environmental factors characteristic for the population can be challenging because of complex interactions between genetic risk factors and environmental exposures [63-65]. Natural selection induced by the environment can, for instance, modify allele frequencies and may lead to distinctive disease susceptibilities in different populations [66,67]. Furthermore, inbreeding in founder populations can increase the incidence of some common diseases, for instance via increased homozygosity of rare variants with large recessive effects [68]. In addition to increasing the incidence of the disease in a given population, environmental factors may have an effect on the severity of the disease phenotype.

Data from model animals suggest that the impact of gene-environment interaction on the phenotype may be considerable [69]. Therefore, accurate determination of phenotype, minimally perturbed by differences in environment, is of great importance for GWA studies - arguably even more so than in linkage studies using family data. Although there is variation in environmental exposures between individuals even in the most homogeneous populations, in population isolates the cultural, environmental and phenotypic homogeneity can facilitate disease-gene identification by reducing variance caused by environmental background. More uniform patterns of, for example, nutrition or exposure to pathogens or homogeneous diagnostic standards, more easily obtained for small populations, provide the best human approximation to controlled experiments in uniform conditions in inbred strains of experimental animals.

The importance of knowing the study population

Population isolates with diverse ethnic backgrounds and different degrees of inbreeding have been described from around the world. Each has its unique characteristics, and may have its own advantages and disadvantages in research into complex diseases (Table 2). Such facts should be considered in study design. Several factors, such as the demographic history of the population, age distribution, number of founders, growth pattern, and degree of genetic and cultural isolation since foundation, determine the features of the genetic landscape of a population isolate [70].

Table 2**Use of isolated versus outbred populations**

Benefits of population isolates	Benefits of outbred populations
Higher degree of LD	More affected individuals
Less areas of very low LD ('holes')	More polymorphic markers
Ability to map recessive genes	More opportunity for replication
Fewer number of causative alleles	
Good genealogical records	
More uniform environment	
Less migration	
More standardized phenotyping	
High participation rate in studies	

Relatively young and small founding populations that have experienced population bottleneck events in their history followed by recent expansion in population size should be ideal for initial locus identification using GWA scans. This is because the population history has created a setting in which the genomes are characterized by a high degree of LD and low genetic diversity [48]. Distinguishing the biologically relevant variants at the associated loci would require older isolates with shorter LD intervals. In small, very ancient isolates with limited population growth, such as the Saami of northern Scandinavia, LD is the result of genetic drift, not a founder effect. These old isolates may be very useful for identifying common disease alleles by drift mapping [71]. Population isolates may also contain sub-isolates, which display different LD intervals of disease alleles as well as different mutation frequencies [72]: these sub-isolates may thus be ideal for complex disease gene mapping even when the founder population itself lacks any obvious advantage.

Population isolates have thus earned their place as an indispensable resource for medical genetics through their use in identifying numerous Mendelian disease genes. Their utility is increasingly valued also in complex disease gene mapping. Genetic, environmental and phenotypic homogeneity, good genealogical records, high participation rates in genetic studies, extended LD in the genome, as well as reduced allelic and locus heterogeneity are highly beneficial features for such studies.

Not all genetic isolates are alike: each population has its own advantages and disadvantages for studies of complex diseases, and thus knowing the genetic makeup of the study population is crucial. The choice and design of statistical methods also deserve particular care in studies utilizing population isolates [73] and the study strategy should also differ depending on the allelic architecture of the disease. The global wealth of population isolates with

well established history and carefully phenotyped study samples is paving the way to a more comprehensive understanding of complex disease genetics. The scientific community might observe the resource of population isolates to be harnessed not only in medical genetics but also in public-health genomics.

Acknowledgements

We would like to acknowledge the Center of Excellence in Complex Disease Genetics of the Academy of Finland, the Nordic Center of Excellence in Disease Genetics, Academy of Finland, Biocentrum Helsinki, Finland, and the European Community's Seventh Framework Programme (FP7/2007-2013) ENGAGE project, grant agreement HEALTH-F4-2007-201413.

References

1. Wellcome Trust Case Control Consortium: **Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.** *Nature* 2007, **447**:661-678.
2. Nikali K, Suomalainen A, Terwilliger J, Koskinen T, Weissenbach J, Peltonen L: **Random search for shared chromosomal regions in four affected individuals: the assignment of a new hereditary ataxia locus.** *Am J Hum Genet* 1995, **56**:1088-1095.
3. Neufeld EJ, Mandel H, Raz T, Szargel R, Yandava CN, Stagg A, Faure S, Barrett T, Buist N, Cohen N: **Localization of the gene for thiamine-responsive megaloblastic anemia syndrome, on the long arm of chromosome 1, by homozygosity mapping.** *Am J Hum Genet* 1997, **61**:1335-1341.
4. **Statistics Iceland** [<http://www.statice.is/>]
5. **deCODE genetics** [<http://www.decode.com/>]
6. Helgason A, Nicholson G, Stefansson K, Donnelly P: **A reassessment of genetic diversity in Icelanders: strong evidence from multiple loci for relative homogeneity caused by genetic drift.** *Ann Hum Genet* 2003, **67**:281-297.
7. Helgadóttir A, Manolescu A, Thorleifsson G, Gretarsdóttir S, Jonsdóttir H, Thorsteinsdóttir U, Samani NJ, Gudmundsson G, Grant SF, Thorgeirsson G, Sveinbjornsdóttir S, Valdimarsson EM, Matthiasson SE, Johannsson H, Gudmundsdóttir O, Gurney ME, Sainz J, Thorhallsdóttir M, Andresdóttir M, Frigge ML, Topol EJ, Kong A, Gudnason V, Hakonarson H, Gulcher JR, Stefansson K: **The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke.** *Nat Genet* 2004, **36**:233-239.
8. Helgadóttir A, Manolescu A, Thorleifsson G, Gretarsdóttir S, Jonsdóttir H, Thorsteinsdóttir U, Samani NJ, Gudmundsson G, Grant SF, Thorgeirsson G, Sveinbjornsdóttir S, Valdimarsson EM, Matthiasson SE, Johannsson H, Gudmundsdóttir O, Gurney ME, Sainz J, Thorhallsdóttir M, Andresdóttir M, Frigge ML, Topol EJ, Kong A, Gudnason V, Hakonarson H, Gulcher JR, Stefansson K: **A common variant on chromosome 9p21 affects the risk of myocardial infarction.** *Science* 2007, **316**:1491-1493.
9. Grant SF, Thorleifsson G, Reynisdóttir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadóttir A, Styrkarsdóttir U, Magnusson KP, Walters GB, Palsdóttir E, Jonsdóttir T, Gudmundsdóttir T, Gylfason A, Saemundsdóttir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdóttir U, Gulcher JR, Kong A, Stefansson K: **Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes.** *Nat Genet* 2006, **38**:320-323.
10. Steinthorsdóttir V, Thorleifsson G, Reynisdóttir I, Benediktsson R, Jonsdóttir T, Walters GB, Styrkarsdóttir U, Gretarsdóttir S, Emilsson V, Ghosh S, Baker A, Snorrardóttir S, Bjarnason H, Ng MC, Hansen T, Bagger Y, Wilensky RL, Reilly MP, Adeyemo A, Chen Y, Zhou J, Gudnason V, Chen G, Huang H, Lashley K, Doumatey A, So WY, Ma RC, Andersen G, Borch-Johnsen K, et al: **A variant in CDKALI influences insulin response and risk of type 2 diabetes.** *Nat Genet* 2007, **39**:770-775.
11. Gudbjartsson DF, Arnar DO, Helgadóttir A, Gretarsdóttir S, Holm H, Sigurdsson A, Jonasdóttir A, Baker A, Thorleifsson G, Kristjansson K, Palsdóttir A, Blondal T, Sulem P, Backman VM, Hardarson GA, Palsdóttir E, Helgason A, Sigurjonsdóttir R, Sverrisson JT, Kostulas

- K, Ng MC, Baum L, So WY, Wong KS, Chan JC, Furie KL, Greenberg SM, Sale M, Kelly P, MacRae CA, et al.: **Variants conferring risk of atrial fibrillation on chromosome 4q25.** *Nature* 2007, **448**:353-357.
12. Gudmundsson J, Sulem P, Rafnar T, Bergthorsson JT, Manolescu A, Gudbjartsson D, Agnarsson BA, Sigurdsson A, Benediktsdottir KR, Blondal T, Jakobsdottir M, Stacey SN, Kostic J, Kristinsson KT, Birgisdottir B, Ghosh S, Magnusdottir DN, Thorlacius S, Thorleifsson G, Zheng SL, Sun J, Chang BL, Elmore JB, Breyer JP, McReynolds KM, Bradley KM, Yaspan BL, Wiklund F, Stattin P, Lindström S, et al.: **Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer.** *Nat Genet* 2008, **40**:281-283.
 13. Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, Sulem P, Thorlacius S, Gylfason A, Steinberg S, Helgadóttir A, Ingason A, Steinthorsdóttir V, Olafsdóttir EJ, Olafsdóttir GH, Jonsson T, Borch-Johnsen K, Hansen T, Andersen G, Jorgensen T, Pedersen O, Aben KK, Witjes JA, Swinkels DW, den Heijer M, Franke B, Verbeek AL, Becker DM, Yanek LR, Becker LC, et al.: **Many sequence variants affecting diversity of adult human height.** *Nat Genet* 2008, **40**:609-615.
 14. Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A, Thorleifsson G, Jakobsdottir M, Steinberg S, Pálsson S, Jonasson F, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediktsdottir KR, Aben KK, Kiemeneý LA, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U, Stefansson K: **Genetic determinants of hair, eye and skin pigmentation in Europeans.** *Nat Genet* 2007, **39**:1443-1452.
 15. Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Jakobsdottir M, Steinberg S, Gudjonsson SA, Palsson A, Thorleifsson G, Pálsson S, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediktsdottir KR, Aben KK, Vermeulen SH, Goldstein AM, Tucker MA, Kiemeneý LA, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U, Stefansson K: **Two newly identified genetic determinants of pigmentation in Europeans.** *Nat Genet* 2008, **40**:835-837.
 16. Peltonen L, Jalanko A, Varilo T: **Molecular genetics of the Finnish disease heritage.** *Hum Mol Genet* 1999, **8**:1913-1923.
 17. St Clair D, Blackwood D, Muir W, Carothers A, Walker M, Spowart G, Gosden C, Evans HJ: **Association within a family of a balanced autosomal translocation with major mental illness.** *Lancet* 1990, **336**:13-16.
 18. Ekelund J, Hovatta I, Parker A, Paunio T, Varilo T, Martin R, Suhonen J, Ellonen P, Chan G, Sinsheimer JS, Sobel E, Juvonen H, Arajärvi R, Partonen T, Suvisaari J, Lönnqvist J, Meyer J, Peltonen L: **Chromosome 1 loci in Finnish schizophrenia families.** *Hum Mol Genet* 2001, **10**:1611-1617.
 19. Hennah W, Varilo T, Kestilä M, Paunio T, Arajärvi R, Haukka J, Parker A, Martin R, Levitzky S, Partonen T, Meyer J, Lönnqvist J, Peltonen L, Ekelund J: **Haplotype transmission analysis provides evidence of association for DISC1 to schizophrenia and suggests sex-dependent effects.** *Hum Mol Genet* 2003, **12**:3151-3159.
 20. Pajukanta P, Lilja HE, Sinsheimer JS, Cantor RM, Lusi AJ, Gentile M, Duan XJ, Soro-Paavonen A, Naukkarinen J, Saarela J, Laakso M, Ehnholm C, Taskinen MR, Peltonen L: **Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1).** *Nat Genet* 2004, **36**:371-376.
 21. Komulainen K, Alanne M, Auro K, Kilpikari R, Pajukanta P, Saarela J, Ellonen P, Salminen K, Kulathinal S, Kuulasmaa K, Silander K, Salomaa V, Perola M, Peltonen L: **Risk alleles of USF1 gene predict cardiovascular disease of women in two prospective studies.** *PLoS Genet* 2006, **2**:e69.
 22. Naukkarinen J, Gentile M, Soro-Paavonen A, Saarela J, Koistinen HA, Pajukanta P, Taskinen MR, Peltonen L: **USF1 and dyslipidemias: converging evidence for a functional intronic variant.** *Hum Mol Genet* 2005, **14**:2595-2605.
 23. Reiner AP, Carlson CS, Jenny NS, Durda JP, Siscovick DS, Nickerson DA, Tracy RP: **USF1 gene variants, cardiovascular risk, and mortality in European Americans: analysis of two US cohort studies.** *Arterioscler Thromb Vasc Biol* 2007, **27**:2736-2742.
 24. Laitinen T, Polvi A, Rydman P, Vendelin J, Pulkkinen V, Salmikangas P, Mäkelä S, Rehn M, Pirskanen A, Rautanen A, Zucchelli M, Gullstén H, Leino M, Alenius H, Petäys T, Haahtela T, Laitinen A, Laprise C, Hudson TJ, Laitinen LA, Kere J: **Characterization of a common susceptibility locus for asthma-related traits.** *Science* 2004, **304**:300-304.
 25. Ober C, Tan Z, Sun Y, Pospick JD, Pan L, Nicolae R, Radford S, Parry RR, Heinzmann A, Deichmann KA, Lester LA, Gern JE, Lemanske RF Jr, Nicolae DL, Elias JA, Chupp GL: **Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function.** *N Engl J Med* 2008, **358**:1682-1691.
 26. Baier LJ, Hanson RL: **Genetic studies of the etiology of type 2 diabetes in Pima Indians: hunting for pieces to a complicated puzzle.** *Diabetes* 2004, **53**:1181-1186.
 27. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, et al.: **A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants.** *Science* 2007, **316**:1341-1345.
 28. Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, Najjar S, Nagaraja R, Orrú M, Usala G, Dei M, Lai S, Maschio A, Busonero F, Mulas A, Ehret GB, Fink AA, Weder AB, Cooper RS, Galan P, Chakravarti A, Schlessinger D, Cao A, Lakatta E, Abecasis GR: **Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits.** *PLoS Genet* 2007, **3**:e115.
 29. Herzberg I, Jasinska A, García J, Jawaheer D, Service S, Kremeyer B, Duque C, Parra MV, Vega J, Ortiz D, Carvajal L, Polanco G, Restrepo GJ, López C, Palacio C, Levinson M, Aldana I, Mathews C, Davanzo P, Molina J, Fournier E, Bejarano J, Ramirez M, Ortiz CA, Araya X, Sabatti C, Reus V, Macaya G, Bedoya G, Ospina J, et al.: **Convergent linkage evidence from two Latin-American population isolates supports the presence of a susceptibility locus for bipolar disorder in 5q31-34.** *Hum Mol Genet* 2006, **15**:3146-3153.
 30. Parsons MJ, Mata I, Beperet M, Iribarren-Iriso F, Arroyo B, Sainz R, Arranz MJ, Kerwin R: **A dopamine D2 receptor gene-related polymorphism is associated with schizophrenia in a Spanish population isolate.** *Psychiatr Genet* 2007, **17**:159-163.
 31. Wijsman EM, Rosenthal EA, Hall D, Blundell ML, Sobin C, Heath SC, Williams R, Brownstein MJ, Gogos JA, Karayiorgou M: **Genome-wide scan in a large complex pedigree with predominantly male schizophrenics from the island of Kosrae: evidence for linkage to chromosome 2q.** *Mol Psychiatry* 2003, **8**:695-705, 643.
 32. Devlin B, Bacanu SA, Roeder K, Reimherr F, Wender P, Galke B, Novasad D, Chu A, Cuenco KT, Tiobek S, Otto C, Byerley W: **Genome-wide multipoint linkage analyses of multiplex schizophrenia pedigrees from the oceanic nation of Palau.** *Mol Psychiatry* 2002, **7**:689-694.
 33. Kaneva R, Milanova V, Angelicheva D, Macgregor S, Kostov C, Vladimirova R, Aleksiev S, Angelova M, Stoyanova V, Loh A, Hallmayer J, Kalaydjieva L, Jablensky A: **Bipolar disorder in the Bulgarian Gypsies: Genetic heterogeneity in a young founder population.** *Am J Med Genet B Neuropsychiatr Genet* 2008, doi: 10.1002/ajmg.b.30775.
 34. Venken T, Claes S, Sluijs S, Paterson AD, van Duijn C, Adolfsson R, Del-Favero J, Van Broeckhoven C: **Genomewide scan for affective disorder susceptibility loci in families of a northern Swedish isolated population.** *Am J Hum Genet* 2005, **76**:237-248.
 35. Kohn Y, Danilovich E, Filon D, Oppenheim A, Karni O, Kanyas K, Turetsky N, Korner M, Lerer B: **Linkage disequilibrium in the DTNBP1 (dysbindin) gene region and on chromosome 1p36 among psychotic patients from a genetic isolate in Israel: findings from identity by descent haplotype sharing analysis.** *Am J Med Genet B Neuropsychiatr Genet* 2004, **128B**:65-70.
 36. Seda O, Tremblay J, Gaudet D, Brunelle PL, Gurau A, Merlo E, Pilote L, Orlov SN, Bouvra F, Petrovich M, Kotchen TA, Cowley AW Jr, Hamet P: **Systematic, genome-wide, sex-specific linkage of cardiovascular traits in French Canadians.** *Hypertension* 2008, **51**:1156-1162.
 37. Gan-Or Z, Giladi N, Rozovski U, Shifrin C, Rosner S, Gurevich T, Bar-Shira A, Orr-Urtreger A: **Genotype-phenotype correlations between GBA mutations and Parkinson disease risk and onset.** *Neurology* 2008, **70**:2277-2283.
 38. Orton NC, Innes AM, Chudley AE, Bech-Hansen NT: **Unique disease heritage of the Dutch-German Mennonite population.** *Am J Med Genet A* 2008, **146A**:1072-1087.
 39. Martin GR, Loredó JC, Sun G: **Lack of association of ghrelin precursor gene variants and percentage body fat or serum lipid profiles.** *Obesity (Silver Spring)* 2008, **16**:908-912.
 40. Henneman P, Aulchenko YS, Frants RR, Willems van Dijk K, Oostra BA, van Duijn CM: **Prevalence and heritability of the metabolic syndrome and its individual components in a Dutch isolate: The Erasmus Rucphen Family (ERF) study.** *J Med Genet* 2008, doi:10.1136/jmg.2008.058388
 41. Hsueh WC, Silver KD, Pollin TI, Bell CJ, O'Connell JR, Mitchell BD, Shuldiner AR: **A genome-wide linkage scan of insulin level derived**

- traits: the Amish Family Diabetes Study. *Diabetes* 2007, **56**:2643-2648.
42. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, et al: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409**:860-921.
 43. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: **dbSNP: the NCBI database of genetic variation.** *Nucleic Acids Res* 2001, **29**:308-311.
 44. **The International HapMap Project.** *Nature* 2003, **426**:789-796.
 45. Pe'er I, de Bakker PI, Maller J, Yelensky R, Altshuler D, Daly MJ: **Evaluating and improving power in whole-genome association studies using fixed marker sets.** *Nat Genet* 2006, **38**:663-667.
 46. Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG, Todd JA: **Haplotype tagging for the identification of common disease genes.** *Nat Genet* 2001, **29**:233-237.
 47. Cardon LR, Abecasis GR: **Using haplotype blocks to map human complex trait loci.** *Trends Genet* 2003, **19**:135-140.
 48. Service S, DeYoung J, Karayiorgou M, Roos JL, Pretorius H, Bedoya G, Ospina J, Ruiz-Linares A, Macedo A, Palha JA, Heutink P, Aulchenko Y, Oostra B, van Duijn C, Jarvelin MR, Varilo T, Peddle L, Rahman P, Piras G, Monne M, Murray S, Galver L, Peltonen L, Sabatti C, Collins A, Freimer N: **Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies.** *Nat Genet* 2006, **38**:556-560.
 49. Gu S, Pakstis AJ, Li H, Speed WC, Kidd JR, Kidd KK: **Significant variation in haplotype block structure but conservation in tagSNP patterns among global populations.** *Eur J Hum Genet* 2007, **15**:302-312.
 50. Shifman S, Kuypers J, Kokoris M, Yakir B, Darvasi A: **Linkage disequilibrium patterns of the human genome across populations.** *Hum Mol Genet* 2003, **12**:771-776.
 51. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D: **The structure of haplotype blocks in the human genome.** *Science* 2002, **296**:2225-2229.
 52. Kruglyak L: **Prospects for whole-genome linkage disequilibrium mapping of common disease genes.** *Nat Genet* 1999, **22**:139-144.
 53. Charrow J: **Ashkenazi Jewish genetic disorders.** *Fam Cancer* 2004, **3**:201-206.
 54. Norio R: **Finnish disease heritage I: characteristics, causes, background.** *Hum Genet* 2003, **112**:441-456.
 55. Enattah NS, Trudeau A, Pimenoff V, Maiuri L, Auricchio S, Greco L, Rossi M, Lentze M, Seo JK, Rahgozar S, Khalil I, Alifrangis M, Natah S, Groop L, Shaat N, Kozlov A, Vershchubskaya G, Comas D, Bulayeva K, Mehdi SQ, Terwilliger JD, Sahi T, Savilahti E, Perola M, Sajantila A, Jarvelä I, Peltonen L: **Evidence of still-ongoing convergence evolution of the lactase persistence T-13910 alleles in humans.** *Am J Hum Genet* 2007, **81**:615-625.
 56. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM, Carrick T, Evans AT, Liao H, Zhao Y, Campbell LE, Schmutz M, Gruber R, Janecke AR, Elias PM, van Steensel MA, Nagtzaam I, van Geel M, Steijnen PM, Munro CS, Bradley DG, Palmer CN, Smith FJ, McLean WH, Irvine AD: **Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema.** *Nat Genet* 2007, **39**:650-654.
 57. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN: **Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease.** *Nat Genet* 2003, **33**:177-182.
 58. Cohen JC, Kiss RS, Pertsemliadis A, Marcel YL, McPherson R, Hobbs HH: **Multiple rare alleles contribute to low plasma levels of HDL cholesterol.** *Science* 2004, **305**:869-872.
 59. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavaré S, Deloukas P, Hurler ME, Dermizakis ET: **Relative impact of nucleotide and copy number variation on gene expression phenotypes.** *Science* 2007, **315**:848-853.
 60. Zeggini E, Rayner W, Morris AP, Hattersley AT, Walker M, Hitman GA, Deloukas P, Cardon LR, McCarthy MI: **An evaluation of HapMap sample size and tagging SNP performance in large-scale empirical and simulated data sets.** *Nat Genet* 2005, **37**:1320-1322.
 61. Arcos-Burgos M, Muenke M: **Genetics of population isolates.** *Clin Genet* 2002, **61**:233-247.
 62. Romeo S, Pennacchio LA, Fu Y, Boerwinkle E, Tybjaerg-Hansen A, Hobbs HH, Cohen JC: **Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL.** *Nat Genet* 2007, **39**:513-516.
 63. Schulz LO, Bennett PH, Ravussin E, Kidd JR, Kidd KK, Esparza J, Valencia ME: **Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the U.S.** *Diabetes Care* 2006, **29**:1866-1871.
 64. Yepiskoposyan L, Harutyunyan A: **Population genetics of familial Mediterranean fever: a review.** *Eur J Hum Genet* 2007, **15**:911-916.
 65. Richards EJ: **Inherited epigenetic variation - revisiting soft inheritance.** *Nat Rev Genet* 2006, **7**:395-401.
 66. Barreiro LB, Laval G, Quach H, Patin E, Quintana-Murci L: **Natural selection has driven population differentiation in modern humans.** *Nat Genet* 2008, **40**:340-345.
 67. Le Souef PN, Candelaria P, Goldblatt J: **Evolution and respiratory genetics.** *Eur Respir J* 2006, **28**:1258-1263.
 68. Rudan I, Rudan D, Campbell H, Carothers A, Wright A, Smolej-Narancic N, Janicijevic B, Jin L, Chakraborty R, Deka R, Rudan P: **Inbreeding and risk of late onset complex disease.** *J Med Genet* 2003, **40**:925-932.
 69. Valdar W, Solberg LC, Gauguier D, Cookson WO, Rawlins JN, Mott R, Flint J: **Genetic and environmental effects on complex traits in mice.** *Genetics* 2006, **174**:959-984.
 70. Wright AF, Carothers AD, Pirastu M: **Population choice in mapping genes for complex diseases.** *Nat Genet* 1999, **23**:397-404.
 71. Terwilliger JD, Zollner S, Laan M, Pääbo S: **Mapping genes through the use of linkage disequilibrium generated by genetic drift: 'drift mapping' in small populations with no demographic expansion.** *Hum Hered* 1998, **48**:138-154.
 72. Fraumene C, Petretto E, Angius A, Pirastu M: **Striking differentiation of sub-populations within a genetically homogeneous isolate (Ogliastra) in Sardinia as revealed by mtDNA analysis.** *Hum Genet* 2003, **114**:1-10.
 73. Bourgain C, Genin E: **Complex trait mapping in isolated populations: Are specific statistical methods required?** *Eur J Hum Genet* 2005, **13**:698-706.
 74. Balaci L, Spada MC, Olla N, Sole G, Loddo L, Anedda F, Naitza S, Zuncheddu MA, Maschio A, Altea D, Uda M, Pilia S, Sanna S, Masala M, Crisponi L, Fattori M, Devoto M, Doratiotto S, Rattu S, Mereu S, Giua E, Cadeddu NG, Atzeni R, Pelosi U, Corrias A, Perra R, Torrazza PL, Pirina P, Ginesu F, Marcias S, et al: **IRAK-M is involved in the pathogenesis of early-onset persistent asthma.** *Am J Hum Genet* 2007, **80**:1103-1114.
 75. Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, Gudbjartsson DF, Walters GB, Ingvarsson T, Jonsdottir T, Saemundsdottir J, Center JR, Nguyen TV, Bagger Y, Gulcher JR, Eisman JA, Christiansen C, Sigurdsson G, Kong A, Thorsteinsdottir U, Stefansson K: **Multiple genetic loci for bone mineral density and fractures.** *N Engl J Med* 2008, **358**:2355-2365.
 76. Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, Jonsson GF, Jakobsdottir M, Bergthorsson JT, Gudmundsson J, Aben KK, Strobbe LJ, Swinkels DW, van Engelenburg KC, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Saez B, Lambea J, Godino J, Polo E, Tres A, Picelli S, Rantalä J, Margolin S, Jonsson T, Sigurdsson H, Jonsdottir T, Hrafnkelsson J, et al: **Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer.** *Nat Genet* 2008, **40**:703-706.
 77. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, Albers-Akkers MT, Swinkels DW, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Godino J, Garcia-Prats MD, Polo E, Tres A, Mouy M, Saemundsdottir J, Backman VM, Gudmundsson L, Kristjansson K, Bergthorsson JT, Kostic J, et al: **Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer.** *Nat Genet* 2007, **39**:865-869.
 78. Helgadóttir A, Thorleifsson G, Magnusson KP, Grétarsdottir S, Steinthorsdottir V, Manolescu A, Jones GT, Rinkel GJ, Blankensteijn JD, Ronkainen A, Jääskeläinen JE, Kyo Y, Lenk GM, Sakalihan N, Kostulas K, Gottsäter A, Flex A, Stefansson H, Hansen T, Andersen G, Weinsheimer S, Borch-Johnsen K, Jorgensen T, Shah SH, Quyyumi AA, Granger CB, Reilly MP, Austin H, Levey AI, Vaccarino

- V, *et al.*: **The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm.** *Nat Genet* 2008, **40**:217-224.
79. Engert JC, Lemire M, Faith J, Brisson D, Fujiwara TM, Roslin NM, Brewer CG, Montpetit A, Darmond-Zwaig C, Renaud Y, Doré C, Bailey SD, Verner A, Tremblay G, St-Pierre J, Bétard C, Platko J, Rioux JD, Morgan K, Hudson TJ, Gaudet D: **Identification of a chromosome 8p locus for early-onset coronary heart disease in a French Canadian population.** *Eur J Hum Genet* 2008, **16**:105-114.
80. Raelson JV, Little RD, Ruether A, Fournier H, Paquin B, Van Eerdewegh P, Bradley WE, Croteau P, Nguyen-Huu Q, Segal J, Debrus S, Allard R, Rosenstiel P, Franke A, Jacobs G, Nikolaus S, Vidal JM, Szego P, Laplante N, Clark HF, Paulussen RJ, Hooper JW, Keith TP, Belouchi A, Schreiber S: **Genome-wide association study for Crohn's disease in the Quebec Founder Population identifies multiple validated disease loci.** *Proc Natl Acad Sci USA* 2007, **104**:14747-14752.
81. Chen WM, Erdos MR, Jackson AU, Saxena R, Sanna S, Silver KD, Timpson NJ, Hansen T, Orrù M, Grazia Piras M, Bonnycastle LL, Willer CJ, Lyssenko V, Shen H, Kuusisto J, Ebrahim S, Sestu N, Duren WL, Spada MC, Stringham HM, Scott LJ, Olla N, Swift AJ, Najjar S, Mitchell BD, Lawlor DA, Smith GD, Ben-Shlomo Y, Andersen G, Borch-Johnsen K, *et al.*: **Variations in the G6PC2/ABCB11 genomic region are associated with fasting glucose levels.** *J Clin Invest* 2008, **118**:2620-2628.
82. Thorleifsson G, Magnusson KP, Sulem P, Walters GB, Gudbjartsson DF, Stefansson H, Jonsson T, Jonasdottir A, Jonasdottir A, Stefansdottir G, Masson G, Hardarson GA, Petursson H, Arnarsson A, Motallebipour M, Wallerman O, Wadelius C, Gulcher JR, Thorsteinsdottir U, Kong A, Jonasson F, Stefansson K: **Common sequence variants in the LOXL1 gene confer susceptibility to exfoliation glaucoma.** *Science* 2007, **317**:1397-1400.
83. Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen WM, Bonnycastle LL, Shen H, Timpson N, Lettre G, Usala G, Chines PS, Stringham HM, Scott LJ, Dei M, Lai S, Albai G, Crisponi L, Naitza S, Doheny KF, Pugh EW, Ben-Shlomo Y, Ebrahim S, Lawlor DA, Bergman RN, Watanabe RM, Uda M, Tuomilehto J, Coresh J, Hirschhorn JN, Shuldiner AR, *et al.*: **Common variants in the GDF5-UQCC region are associated with variation in human height.** *Nat Genet* 2008, **40**:198-203.
84. Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, Manolescu A, Thorleifsson G, Stefansson H, Ingason A, Stacey SN, Bergthorsson JT, Thorlacius S, Gudmundsson J, Jonsson T, Jakobsdottir M, Saemundsdottir J, Olafsdottir O, Gudmundsson LJ, Bjornsdottir G, Kristjansson K, Skuladottir H, Isaksson HJ, Gudbjartsson T, Jones GT, Mueller T, Gottsäter A, Flex A, Aben KK, de Vegt F, *et al.*: **A variant associated with nicotine dependence, lung cancer and peripheral arterial disease.** *Nature* 2008, **452**:638-642.
85. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Blondal T, Stacey SN, Helgason A, Gunnarsdottir S, Olafsdottir A, Kristinsson KT, Birgisdottir B, Ghosh S, Thorlacius S, Magnusdottir D, Stefansdottir G, Kristjansson K, Bagger Y, Wilensky RL, Reilly MP, Morris AD, Kimber CH, *et al.*: **Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes.** *Nat Genet* 2007, **39**:977-983.
86. Palo OM, Antila M, Silander K, Hennah W, Kilpinen H, Soronen P, Tuulio-Henriksson A, Kiesepää T, Partonen T, Lönnqvist J, Peltonen L, Paunio T: **Association of distinct allelic haplotypes of DISC1 with psychotic and bipolar spectrum disorders and with underlying cognitive impairments.** *Hum Mol Genet* 2007, **16**:2517-2528.
87. Chavarría-Siles I, Walss-Bass C, Quezada P, Dassori A, Contreras S, Medina R, Ramírez M, Armas R, Salazar R, Leach RJ, Raventos H, Escamilla MA: **TGFB-induced factor (TGIF): a candidate gene for psychosis on chromosome 18p.** *Mol Psychiatry* 2007, **12**:1033-1041.
88. Maniatis N, Collins A, Xu CF, McCarthy LC, Hewett DR, Tapper W, Ennis S, Ke X, Morton NE: **The first linkage disequilibrium (LD) maps: delineation of hot and cold blocks by diplotype analysis.** *Proc Natl Acad Sci USA* 2002, **99**:2228-2233.