REVIEW

Single nucleotide polymorphism in genome-wide association of human population: A tool for broad spectrum service

Mohd Fareed, Mohammad Afzal *

Human Genetics and Toxicology Lab., Section of Genetics, Department of Zoology, Faculty of Life Sciences, Aligarh Muslim University, Aligarh 202002, Uttar Pradesh, India

Abstract Genome-wide patterns of variation across individuals provide most powerful source of data for uncovering the history of migration, expansion, and adaptation of the human population. The arrival of new technologies that type more than millions of the single nucleotide polymorphisms (SNPs) in a single experiment has made SNP in genome-wide association (GWA) assay a prudent venture. SNPs represent the most widespread type of sequence variation in genomes, and known as valuable genetic markers for revealing the evolutionary history and common genetic polymorphisms that explain the heritable risk for common diseases. Characterizing the nature of gene variation in human populations and assembling an extensive catalog of SNPs in candidate genes in association with particular diseases are the major goals of human genetics. In this article we explore the recent discovery of SNP–GWA to revolutionize not only the process of genetic variation and disease detection but also the convention of preventative and curative medicine for future prospects.

Contents

* Corresponding author. Tel.: +91 0571 2700920x3442/+91 9897839601.
E-mail address: afzal1235@rediffmail.com (M. Afzal).
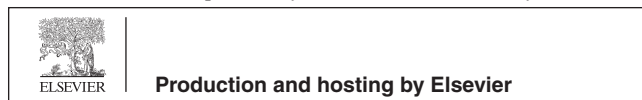
1. Introduction

Genome-wide data sets are increasingly being used to identify biological pathways and networks underlying complex diseases and in drug development process (Fig. 1). In particular, analyzing genomic data through sets defined by functional pathways offers the potential of greater power for discovery and natural connections to biological mechanisms. Much of genetic variation in the human genome is in the form of SNPs which is the result of point mutations that produce single base-pair differences (substitutions or deletions) among chromosome sequences. There are many laboratories and computational approaches to finding single nucleotide polymorphisms (SNPs) within a genome, but all involve some form of comparative analysis of the same DNA segment from different individuals or from different haplotypes. SNP identification can be based on expressed sequence tags (ESTs), which are generated by single-run sequencing of cDNAs obtained from different individuals and assembly of overlapping sequences for the same region permits novel SNP discoveries. The non-coding SNPs can be classified according to whether they are found in gene regulating segments of the genome. Many complex diseases may arise from quantitative, rather than qualitative, differences in gene products. Coding SNPs can be classified as to whether they alter the sequence of the protein encoded by the altered gene. Changes that alter protein sequences can be classified by their effects on protein structure [1].

GWA studies have become an important tool for discovering susceptibility genes for complex diseases. Information, including genotype frequencies, linkage disequilibrium (LD), and recombination rates, across populations help researchers to conduct GWA analysis using millions of SNP markers. The differences in association results among populations for phenotypes of interest are partially explained by HapMap information such as population specific common variants and linkage disequilibrium blocks [2,3]. Moreover, the phased haplotypes of HapMap samples are used as a reference for imputing untyped markers. One million SNPs can be increased to up to 2.5 million by imputing haplotypes from HapMap phased haplotypes based on the pattern of observed genotypes [4].

2. SNPs in human population and concerning issues

GWA studies build directly on recent efforts to map the patterns of inheritance for the most common form of genomic variation by the use of SNPs [5]. An estimated 10 million common SNPs, those with a minor-allele frequency of at least 5%, are transmitted across generations in blocks, allowing a few
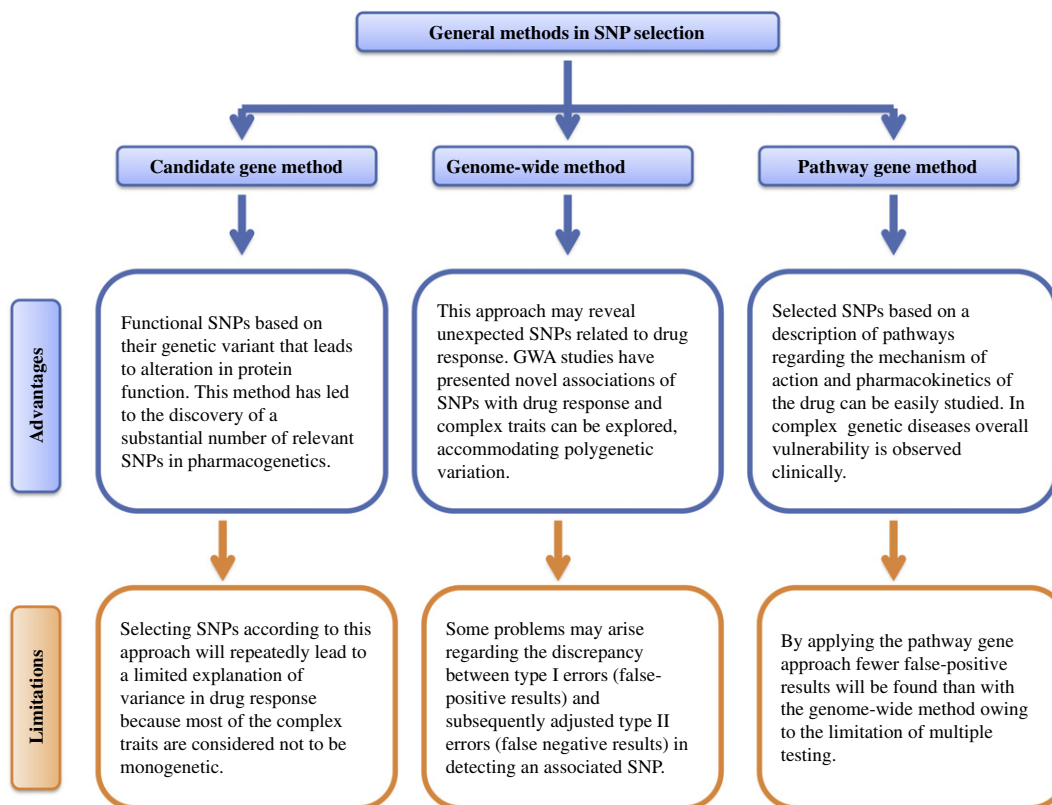

Figure 1   Selection of SNP made through three different pathways (approaches) exploring the mechanism of genome variation and drug development with some advantages and limitations [105–113].

particular, or tag, SNPs to capture the great majority of SNP variation within each block [6]. Rapid advances in technology and quality control now permit affordable, reliable genotyping of up to 1 million SNPs in a single scan of a person's DNA [7]. Genotyping hundreds of thousands of SNPs has led to a great accuracy with new models and approaches in inferring population structure, thus reconstructing population histories from given SNPs data [8].

A complication of genome wide association studies is the enormous number of tests of association required (at least one per SNP); thresholds of statistical significance are stringent, making it necessary to work with very large samples [9]. One frequently used approach to managing size is the tiered design, in which a subset of SNPs found to be significant in the GWA study (sometimes called the discovery set) is genotyped in a second tier (a replication set), yielding a smaller subset of significantly associated SNPs that are then tested in a third tier (a second replication set), and so on [10,11]. This process helps to identify false positive associations. Carrying forward a large number of SNPs identified through a GWA study into a test of replication also minimizes false negative results [12], while raising the bar for the establishment of true positive results. The pooling of results obtained in GWA studies (Fig. 2) under the auspices of large consortia is often required for the detection of variants with small effects on the risk of disease. Such pooled studies, like all genetic association studies, must be examined and controlled for differences in allele frequency between groups that can lead to spurious (false positive) associations [13].

3. SNPs reveal functional polymorphism

Regulatory polymorphisms at DNA level can potentially cause variations in gene expression. A SNP in a regulatory DNA binding site may alter the affinity with the regulatory protein, resulting in different gene expressions as shown in Fig. 2. SNPs in the osteopontin promoter have been shown to modify DNA binding affinity to transcription factors SP1/SP3 [14]. A GWA study revealed a G-to-A substitution in the 5_ untranslated region (5_-UTR) of the FOXE1 gene to associate with thyroid cancer susceptibility [15]. The T-to-C substitution located in the 5_-UTR of the GDF5 gene causes a different interaction with DEAF-1, a trans-acting factor for GDF5, leading to a reduced gene expression [16].

Opioid receptors (mainly μ-, δ- and κ-receptors) in the endogenous opioid system regulate neuronal activity via neurotransmission or neuromodulation. Variation in opioid receptor genes may change the expression level or activity of opioid

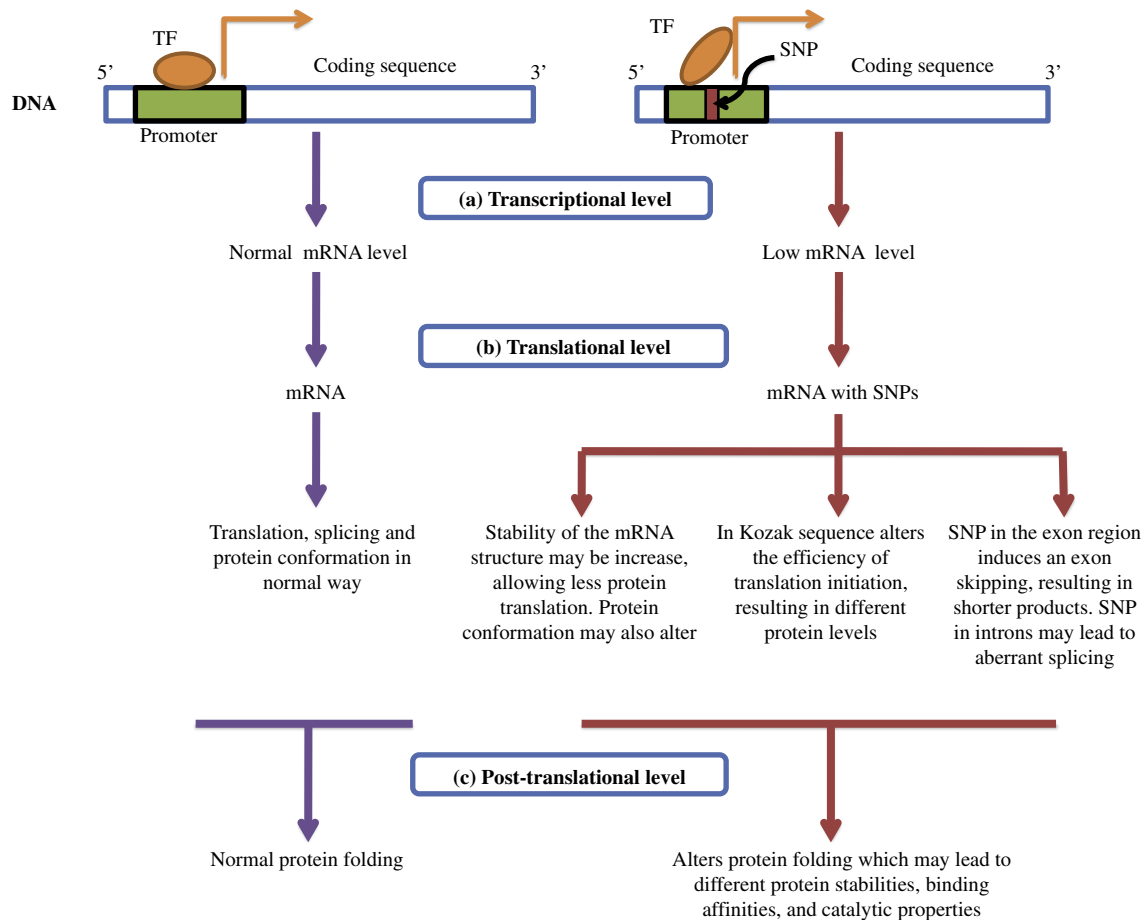**Figure 2** SNP in a regulatory DNA binding site may alter the affinity of the regulatory protein such as transcription factor (TF) resulting into different mRNA levels. The SNP in promoter region leads to decreased affinity as shown above, resulting in low gene expression. Several examples of the effects of SNPs on mRNA, translational levels and post-transcriptional levels [114–118] are there.

receptors, leading to increased risk for drug or alcohol dependence. The minor G-allele of SNP rs569356 may enhance transcription factor binding and increase δ-opioid receptor gene (*OPRD1*) expression [17].

At the protein and post-translational levels, variation in protein stability due to SNPs in coding sequences can cause different levels of enzyme activities. Thiopurine S-methyltransferase (TPMT) catalyzes the S-methylation of thiopurine drugs. Several human TPMT variant alleles that alter the encoded amino acid sequence of the enzyme generate less stable proteins [18]. Therefore, patients with those alleles have very low TPMT activity and suffer severe, life-threatening drug toxicity when treated with "standard" doses of thiopurine drugs [19]. The μ opioid receptors (MOPRs), belonging to the seven transmembrane receptor (7TMR) family [also called the G protein-coupled receptor (GPCR) family], mediate the pharmacological effects of morphine and other μ-preferring compounds. Activation of the MOPRs produces analgesia, reward, mood changes, sedation, respiratory depression, immunosuppression, decreased gastrointestinal motility, increased locomotor activity, tolerance and dependence. A common A118G SNP reduces MOPR N-glycosylation and protein stability [20].

Non-synonymous SNPs can affect post-translational modifications. Because protein phosphorylation is one of the key elements in signal transduction, an altered phosphorylation pattern can cause different responses to the environment. The human ether-a-go-go-related gene 1 (ERG1) protein channel polymorphism is associated with cardiac arrhythmias [21]. A SNP leading to a Lys897Thr substitution creates a phosphorylation site in ERG1, which in turn inhibits channel activity for the downstream signal transduction [22].

4. GWA study in human disease

GWA studies in which hundreds of thousands of single-nucleotide polymorphisms (SNPs) are tested for association with a disease in hundreds or thousands of persons have revolutionized the search for genetic influences on complex traits [23,24]. Such conditions, in contrast with single-gene disorders, are caused by many genetic and environmental factors working together, each having a relatively small effect and few if any being absolutely required for a disease to occur. Although complex conditions have been referred to as the geneticist's nightmare [25], in recent years GWA studies have identified SNPs implicating more than hundreds of robustly replicated loci for common diseases such as cancer, infectious diseases and many other traits shown in Tables 1–3 respectively.

GWA studies use dense maps of SNPs covering the human genome to look for allelic frequency differences between the cases (patients with a specific disease or individuals with a certain trait) and controls. Linkage disequilibrium (LD) mapping, recognizes that a mutation which is shared by affected individuals through common descent is surrounded by other alleles at nearby loci, thus representing the haplotype of the ancestral chromosome region where the mutation first occurred. Two polymorphic sites are said to be in LD when their specific alleles are correlated in a population. High LD means that the SNP alleles are almost always inherited together; information about the allele of one SNP in an individual is strongly

Table 1 GWA studies showing SNPs association with cancer.

Disease	Chromosomal location	Gene	Strongest SNPs-risk allele	*P*-value	Reference
Breast cancer	10q26.13	*FGFR2*	rs2981582-G	2×10^{-76}	[11,36-38]
	16q12.1	*TNCR9/LOC643714*	rs3803662-C	1×10^{-36}	
	5q11.2	*MAP3K1*	rs889312-A	7×10^{-20}	
	11p15.5	*LSP1*	rs3817198-T	3×10^{-9}	
	8q24.21	Intergenic	rs13281615-T	5×10^{-12}	
	2q34	*ERBB4*	rs13393577-C/T	8.8×10^{-14}	
	10q26.13	*FGFR2*	rs1219648-G	1×10^{-10}	
	2q35	Intergenic	rs13387042-A	1×10^{-13}	
Basal cell carcinoma	1q42.13	*RHOU*	rs801114-G	6×10^{-12}	[39-44]
	1p36.13	*PADI4, PADI6, RCC2, ARHGEF10L*	rs7538876-A	4×10^{-12}	
	12q12.13	*KRT5*	rs11170164-A	2×10^{-9}	
	9p21	*CDKN2A/B*	rs2151280-C	7×10^{-9}	
	8q24.21	Intergenic	rs16901979-A	1×10^{-12}	
	8q24.21	Intergenic	rs6983267-G	9×10^{-13}	
	17q12	*TCF2*	rs4430796-A	1×10^{-11}	
	11p15	*IGF2, IGF2A, INS, TH*	rs7127900-A	3×10^{-33}	
Colorectal cancer	8q24.21	*ORF, DQ515897*	rs10505477-A	3×10^{-11}	[45,46]
	18q21.1	*SMAD7*	rs4939827-T	1×10^{-12}	
Lung cancer	15q25.1	*CHRNA3, CHRNA5, CHRNB4, PSMA4, LOC123688*	rs8034191-C	5×10^{-20}	[47]
	6p22.1	*TRNAA-UGC*	rs4324798-A	2×10^{-8}	[48]
Melanoma	20q11.22	*CDC91L1*	rs910873-T	1×10^{-15}	[49,50]
	22q13.1	Intergenic	rs2284063-G	2×10^{-9}	
Neuroblastoma	6p22.3	*FLJ22536, FLJ44180*	rs6939340-G	9×10^{-15}	[51,52]
	2q35	*BARD1*	rs6435862-G	9×10^{-18}	
Ovarian cancer	9p22.2	Intergenic	rs3814113-T	5×10^{-19}	[53]
Thyroid cancer	9q22.23	*FOXE1*	rs965513-A	2×10^{-27}	[15]
	14q13.3	*NKX2-1*	rs944289-T	2×10^{-9}	

Table 2 SNPs significantly associated with infectious disease phenotypes in genome-wide studies.

Disease	SNP location	Strongest SNPs-risk allele	P-value	Reference
Creutzfeldt–Jakob disease	PRNP	rs1799990	2.0×10^{-27}	[54]
Dengue shock Syndrome	MICB	rs3132468	4.4×10^{-11}	[55]
Hepatitis B	HLA-DPA1	rs3077	2.3×10^{-38}	[56]
	HLA-DPB1	rs9277535	6.3×10^{-39}	
Hepatitis C	IL28B	rs8099917	6.1×10^{-9}	[57]
HIV-1 and AIDS	HLA-C	rs9264942	5.9×10^{-32}	[58–64]
	HLA-B, HCP5	rs2395029	4.5×10^{-35}	
	HLA-B	rs2523608	5.6×10^{-10}	
	HLA-C	rs9264942	2.8×10^{-35}	
	MICA	rs4418214	1.4×10^{-34}	
	HLA-B, HCP5	rs2395029	9.7×10^{-26}	
	PSORS1C3	rs3131018	4.2×10^{-16}	
	HLA-B	rs2523608	8.9×10^{-20}	
	Intergenic	rs2255221	3.5×10^{-14}	
	HLA-B	rs2523590	1.7×10^{-13}	
	Intergenic	rs9262632	1.0×10^{-8}	
	ZNRD1, RNF39	rs9261174	1.8×10^{-8}	
	PARD3B	rs11884476	3.4×10^{-9}	
	HLA-B, HCP5	rs2395029	6.8×10^{-10}	
	CXCR6	rs2234358	9.7×10^{-10}	
Leprosy	LACC1	rs3764147	3.7×10^{-54}	[65]
	NOD2	rs9302752	3.8×10^{-40}	
	RIPK2	rs42490	1.4×10^{-16}	
	CCDC122	rs3088362	1.4×10^{-31}	
	TNFSF15	rs6478108	3.4×10^{-21}	
Meningococcal disease	CFH	rs1065489	2.2×10^{-11}	[66]
Severe malaria	HBB	rs11036238	3.7×10^{-11}	[67]
Tuberculosis	18q11.2 GATA6, TAGE1, RBBP8, CABLES1	rs4334126	6.8×10^{-9}	[68]

predictive of the allele of the other SNP on that chromosome. The first example of LD between a DNA polymorphism and a disease mutation was provided by an association between an allele of an RFLP in the β-globin gene and the sickle-cell form of hemoglobin [26].

Next-generation sequencing projects are revolutionizing our understanding of genetic variation. The quality of data from the next-generation technology and the availability of analysis tools are both rapidly increasing. Just considering the pilot data from the 1000 Genome Projects, this new resource has provided the location, allele frequency, and local haplotype structure of approximately 15 million SNPs [27]. Successful GWA studies are the most visible and exciting outcome of HapMap which has also been invaluable in developing genotyping and analytic methods to realize advances in the prevention and treatment of common diseases. The tool of sequencing enables scientists to pinpoint functional variants from association studies and improve the knowledge available to researchers interested in evolutionary biology, and hence may lay the foundation for predicting disease susceptibility and drug response.

5. Role of SNPs in drug development

GWAS have uncovered many genetic loci that are associated with human diseases, but two fundamental limitations have hampered our ability to translate these results into clinically useful predictors of disease and drug targets [28]. The U.S. Food and Drug Administration's Adverse Event Reporting System (AERS) allows us to develop an initial understanding

of the context within which molecular level drug-target interactions can lead to distal effectors that results in adverse phenotypes at the organ and organismal levels [29]. Targeting drugs for complex diseases and predicting therapeutic efficacy and adverse risk for individuals with allelic variation (like SNPs) are the major goals of pharmacology. The allelic variants of the genes can lead to adverse effect of drugs and open new ways to drug development. Some of the genes related to severe diseases with pharmacogenomic targets and recommended drug therapy are shown in Table 4.

Drug transporters are now widely known as important determinants governing drug absorption, excretion, and, in many cases, extent of drug entry into target organs. Transporters of the solute carrier (SLC) and ATP-binding cassette (ABC) superfamilies are considered to be of major importance in drug therapy and outline how understanding the expression, function, and genetic variation in such drug transporters will result in better strategies for optimal drug design and tissue targeting as well as reduce the risk for drug–drug interactions and adverse drug responses. SNPs in genes encoding Organic cation transporters (OCTs) have been identified and characterized [30].

Aldehyde oxidase (AO) is a complex molybdoflavoprotein that belongs to the xanthine oxidase family. Human aldehyde oxidase (hAOX1) encoded by *AOX1* gene has an important role in the metabolism of drugs, based on its broad substrate specificity oxidizing aromatic aza–heterocycles, e.g., N1-methylnicotinamide and N-methylphthalazinium, or aldehydes, such as benzaldehyde, retinal and vanillin. SNP based functionally inactive hAOX1 allelic variants and also variants coding for enzymes with different catalytic activities are well

considered for the design of future drugs [31]. *VKORC1* encodes the vitamin K-epoxide reductase protein, the target enzyme of warfarin. *VKORC1* catalyzes the conversion of vitamin K-epoxide into vitamin K, which is the rate-limiting step

Table 3 SNP-GWA association studies in various traits.

Disease/trait	Chromosomal location	Gene	Strongest risk allele	P-value	References
Autoimmune diseases					
Rheumatoid arthritis	2p14	SPRED2	rs934734-A	3.2×10^{-7}	[69–71]
	5q11	ANKRD55, IL6ST	rs6859219-C	2.5×10^{-9}	
	3p14	PXK	rs13315591-T	3.7×10^{-7}	
	4p15	RBPJ	rs874040-G	1.9×10^{-7}	
	3 6q27	CCR6	rs309302-G	3.3×10^{-7}	
	31 7q32	IRF5	rs104886-T	2.8×10^{-6}	
	9q34	*TRAF1-C5*	rs3761847-G	1×10^{-14}	
	6q23.3	Near *TNFAIP3, OLIG3*	rs10499194-C	1×10^{-9}	
Systemic lupus erythematosus	4q24	*BANK1*	rs10516487-G	4×10^{-10}	[72]
Cardiovascular conditions					
Atrial fibrillation/atrial flutter	4q25	Intergenic	rs2200733-T	3×10^{-41}	[73]
		Near *PITX2*	rs10033464-T	7×10^{-11}	
Coronary disease	9p21.3	*CDKN2A/B*	rs1333049-C	1×10^{-13}	[74,75]
	9p21.3	*CDKN2A/B*	rs1333049-C	3×10^{-19}	
	6q25.1	*MTHFD1L*	rs6922269-A	3×10^{-8}	
	2q36.3	Pseudogene	rs2943634-C	2×10^{-7}	
Coronary heart disease (CHD)	9q33	DAB2IP	rs7025486-?	0.003	[76]
Myocardial	9p21.3	*CDKN2A/B*	rs10757278-G	1×10^{-20}	[77]
Diabetes					
Type 1 diabetes	19q13.33	*FUT2*	rs601338A > G	4.3×10^{-18}	[74,78–80]
	12q24.13	*C12orf30*	rs17696736-G	2×10^{-16}	
	12q13.2	*ERBB3*	rs2292239-A	2×10^{-20}	
	16p13.13	*KIAA0350*	rs12708716-A	3×10^{-18}	
	18p11.21	*PTPN2*	rs2542151-C	1×10^{-14}	
	18q22.2	*CD226*	rs763361-A	1×10^{-8}	
	12q13.2	*ERBB3*	rs11171739-C	1×10^{-11}	
	16q13.13	*TRAFD1, PTPN11, KIAA0350*	rs12708716-A	5×10^{-7}	
	16p13.13	*KIAA0350*	rs2903692-G	7×10^{-11}	
Type 2 diabetes	11p15	*KCNQ1*	rs151290-?	0.002	[25,81-84]
	8q24.11	*SLC30A8*	rs13266634-C	6×10^{-8}	
	16q12.2	*FTO*	rs8050136-A	1×10^{-12}	
	10q23.33	*HHEX*	rs5015480-C	6×10^{-10}	
	6p22.3	*CDKAL1*	rs10946398-C	4×10^{-11}	
	9p21.3	*CDKN2B*	rs10811661-T	8×10^{-15}	
	3q27.2	*IGFBP2*	rs4402960-T	9×10^{-16}	
	9p21.3	*CDKN2A/B*	rs10811661-T	8×10^{-15}	
	3q27.2	*IGF2BP2*	rs4402960-T	9×10^{-16}	
	6p22.3	*CDKAL1*	rs7754840-C	4×10^{-11}	
	6p22.3	*CDKAL1*	rs7754840-C	4×10^{-11}	
	9p21.3	*CDKN2A/B*	rs10811661-T	8×10^{-15}	
	3q27.2	*IGF2BP2*	rs4402960-T	9×10^{-16}	
	11p12	Intergenic	rs9300039-C	4×10^{-7}	
Gastrointestinal disorders					
Celiac disease	4q27	*KIA1109, TENR, IL2, IL21*	rs6822844-G	1×10^{-14}	[85]
Crohn disease	5p13.1	Intergenic	rs1373692-?	2×10^{-12}	[74,86-88]
	1q24.3	Intergenic	rs12035082-?	2×10^{-7}	
	18p11.	*PTPN2*	rs2542151-?	3×10^{-8}	
	3p21.31	Many genes	rs9858542-?	5×10^{-8}	
	5q33.1	*IRGM*	rs13361189-?	2×10^{-10}	
	21q22.2	Intergenic	rs2836754-?	5×10^{-7}	
	10q24.2	*NKX2–3*	rs10883365-?	4×10^{-18}	
	1q31.2	Intergenic	rs10801047-?	3×10^{-8}	
	2q37.1	*ATG16L1*	rs2241880-G	1×10^{-13}	
	3p21.31	*BSN, MST1*	rs9858542-A	4×10^{-8}	
	5q33.1	*IRGM*	rs1000113-T	3×10^{-7}	
	10q24.2	*NKX2–3*	rs10883365-G	6×10^{-8}	
	18p11.21	*PTPN2*	rs2542151-G	2×10^{-7}	

Table 3 (*continued*)

Gallstones	2p24.2	*ABCG8*	rs11887534-C	1×10^{-14}	[89]
	12q24.13	*SH2B3/LNK*	rs17696736-G	2×10^{-14}	
Inflammatory bowel disease	1p31	*IL23R*	rs11209026-A	4×10^{-11}	[90]
Lipid metabolism					
HDL-cholesterol	1q42.13	*GALNT2*	rs4846914-G	2×10^{-13}	[91,92]
	12q24.11	*MVK/MMAB*	rs2338104-G	3×10^{-8}	
LDL-cholesterol	1p13.3	*CELSR2, PSRC1*	rs599839-G	1×10^{-7}	[91,93]
	1p13.3	*CELSR2, PSRC1, SORT1*		3×10^{-29}	
	19p13.11	*CILP2, PBX4*	rs16996148-G	3×10^{-8}	
Triglycerides	7q11.23	*BCL7B, TBL2, MLXIPL*	rs17145738-T	7×10^{-22}	[91,92,94]
	19p13.11	*CILP2, PBX4*	rs16996148-G	4×10^{-9}	
	8q24.13	*TRIB1*	rs17321515-A	4×10^{-17}	
	1q42.13	*GALNT2*	rs4846914-G	7×10^{-15}	
	1p31.3	*ANGPTL3, DOCK7, ATG4C*	rs12130333-C	2×10^{-8}	
	2p23.3	*GCKR*	rs780094-T	6×10^{-32}	
	8q24.13	*TRIB1*	rs17321515-A	7×10^{-13}	
	19p13.3	*NCAN/CILP2*	rs16996148-G	3×10^{-9}	
	7q11.23	*MLXIPL*	rs17145738-C	2×10^{-12}	
	1p31.3	*ANGPTL3*	rs1748195-C	2×10^{-10}	
	7q11.23	*MLXIPL*	rs3812316-C	1×10^{-10}	
Neuropsychiatric conditions					
Amyotrophic lateral sclerosis	7q36.2	*DPP6*	rs10260404-C	5×10^{-8}	[95]
APOE*e4 with late-onset Alzheimer disease	11q14.1	*GAB2*	rs2373115-G	1×10^{-10}	[96]
Bipolar disorder	13q14.11	*DGKH*	rs1012053-A	2×10^{-8}	[74,97]
	16p12.1	*PALB2, NDUFAB1, DCTN5*	rs420259-A	6×10^{-8}	
Multiple sclerosis	16p13	*KIAA 0350*	rs6498169-?	?	[98,99]
	10p15.1	*IL2RA*	rs12722489-C	3×10^{-8}	
	5p13.2	*IL7RA*	rs6897932-C	3×10^{-7}	
Restless legs syndrome	6p21.2	*BTBD9*	rs3923809-A	1×10^{-17}	[100,101]
	2p14	*MEIS1*	rs2300478-G	3×10^{-28}	
	6p21.2	*BTBD9*	rs9296249-T	4×10^{-18}	
	15q23	*MAP2K5, LBXCOR1*	rs12593813-G	1×10^{-15}	
Schizophrenia	Xp22.33/Yp11.32	*CSF2RA*	rs4129148-C	4×10^{-7}	[102]
Sporadic amyotrophic lateral sclerosis	10q26.13	Intergenic	rs4363506-?	7×10^{-7}	[103]

N/R = not reported, ? = not described.

in vitamin K recycling. A common non-coding variant ($-1639G > A$, rs9923231) is significantly associated with warfarin sensitivity and its polymorphism alters a *VKORC1* transcription factor binding site, leading to lower protein expression [32].

Fig. 3 represents allelic variants associated with drug-induced liver injury (DILI). Pharmacogenomic study reveals the associations between DILI susceptibility and several polymorphisms in various genes, such as specific alleles in manganese superoxide dismutase glutathione peroxidase, and glutathione *S*-transferase that are significantly more frequent in DILI patients than in controls and shared across multiple drugs [33–36]. In the context of large genetic variation, GWA studies help us in identifying SNPs as pharmacodynamic models which can lead to develop a mechanistic understanding of drug action in the context of population as well as an individual's genomic status.

6. Conclusion

Natural selection exerts its influence by changing allele frequencies of polymorphic markers to eliminate a deleterious phenotype from a population or otherwise fix a beneficial one. The human complex traits as a matter of fact, show variation with SNPs having adverse effects on drug exposure which could be lethal. Many SNPs have been explored as a high-resolution marker for accelerating the pace of gene mapping related to diseases or traits. The GWA–SNPs have been studied in different human populations and their quantification for population structure within and between the populations has been attempted for association studies. Interindividual variability in drug response is influenced by variation in genes that control the absorption, distribution, metabolism and excretion of drugs. The Wellcome Trust Case Control Consortium (WTCCC) is engaged in exploring the utility, design and analysis of GWA studies. The International HapMap resource, documents patterns of genome-wide variation and linkage disequilibrium in population samples and greatly facilitates both the design and analysis of association studies. The availability of advance genotyping chips, containing sets of large number of SNPs provides good coverage of the human genome, to which means that GWA studies can be used for thousands of cases and controls for exploring the basis of complex traits.

Identification of the genetic basis for polymorphic expression of a gene is done through intronic or exomic SNPs which abolishes the need for different mechanisms for explaining the variability in drug metabolism. SNPs based variations in membrane receptors lead to multidrug resistance (MDR) and the drug–drug interactions. Even drug induced toxicity and many adverse effects can be explained by GWA studies. The aim of

Table 4 Some genes identified by Genome-wide studies which are pharmacogenomic targets with recommended drug therapy.

Therapeutic area	Gene/marker	Drug involved
Allergy	CYP2D6	Desloratadine, Pseudoephedrine
Analgesics	CYP2D6	Codeine, Tramadol, Acetaminophen
	CYP2C9	Celecoxib
Antiarrhythmics	CYP2D6	Quinidine
Antifungals	CYP2C19	Voriconazole
	CYP2D6	Terbinafine
Antiinfectives	NAT1; NAT2	Rifampin, Isoniazid, Pyrazinamide
Antivirals	CCR5	Maraviroc
	IL28B	Boceprevir
Cardiovascular	ApoE2	Pravastatin
	CYP2C19	Ticagrelor, Clopidogrel
	CYP2D6	Metoprolol, Carvedilol, Propranolol
Dermatology and dental	PML/RARα	Tretinoin
	DPD	Fluorouracil
	CYP2D6	Cevimeline
Gastroenterology	UCD NAGS; CPS; ASS; OTC; ASL; ARG	Sodium Phenylacetate, Sodium Benzoate
	CYP2C19	Rabeprazole, Esomeprazole, Pantoprazole, Dexlansoprazole,
Hematology	VKORC1	Warfarin
	CYP2C9	Warfarin
Metabolic and Endocrinology	LDL receptor	Atorvastatin
Musculoskeletal	CYP2C19	Carisoprodol
Neurology	CYP2D6	Dextromethorphan, Quinidine, Tetrabenazine, Galantamine
	CYP2C19	Clobazam
	HLA-B*1502	Carbamazepine
Oncology	PDGFR	Imatinib
	PML/RARα	Arsenic Trioxide
	CYP2D6	Gefitinib
	CD20 antigen	Tositumomab
	KRAS	Cetuximab, Panitumumab
	Her2/neu	Lapatinib, Trastuzumab
	CD30	Brentuximab Vedotin
	C-Kit	Imatinib
	DPD	Capecitabine
	TPMT	Mercaptopurine, Thioguanine, Cisplatin
	FIP1L1-PDGFRα	Imatinib
	ER &/PgR receptor	Exemestane
	ER receptor	Fulvestrant, Tamoxifen
	Ph Chromosome	Nilotinib, Dasatinib, Busulfan, Imatinib
	BRAF	Vemurafenib
	ALK	Crizotinib
	EGFR	Panitumumab, Cetuximab
Psychiatry	UCD NAGS; CPS; ASS; OTC; ASL; ARG	Valproic Acid
	CYP2C19	Diazepam, Citalopram, Fluvoxamine, Modafinil
	CYP2D6	Citalopram, Clomipramine, Desipramine, Aripiprazole, Iloperidone, Fluoxetine, Pimozide, Doxepin, Protriptyline, Chlordiazepoxide, Amitriptyline, Modafinil, Nefazodone, Nortriptyline, Thioridazine, Trimipramine, Venlafaxine, Clozapine, Risperidone, Fluoxetine, Olanzapine
Pulmonary	CFTR G551D	Ivacaftor
Reproductive and urologic	CYP2D6	Tolterodine
Rheumatology	TPMT	Azathioprine
	CYP2C9	Flurbiprofen

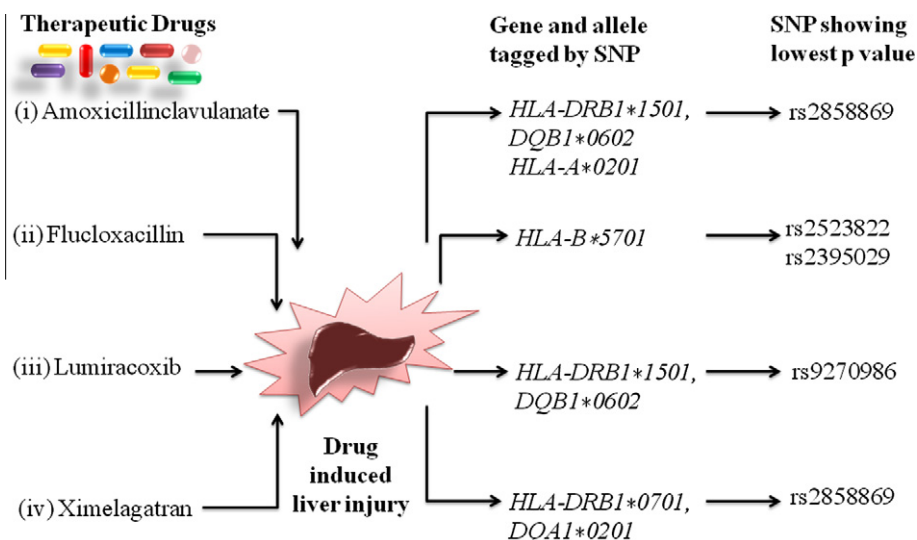Compiled from Food and Drug Administration, 2012 [104].

Figure 3 GWA approach to uncover SNPs for drug induced hepatotoxicity. The study has shown statistically significant associations with particular HLA class I or II alleles, suggesting that T cell responses contribute to the liver injury. The adverse reactions by the corresponding SNPs against HLA genes in response to drugs are shown on the right hand side of the above figure which is compiled from the data of some recent studies [119–122] based on GWA studies.

this review is to throw fresh light on human genetic variation using SNPs to predict drug response and its wide applications in medical, health and pharmacogenetic studies. The tool of SNPs can help understand the alteration in the amino-acid sequence of the encoded protein and make a common cause of pharmacologically relevant functional variation. The multitude of SNPs help in understanding gene pharmacokinetic (PK) or pharmacodynamic (PD) pathways. The association of a wide range of human diseases like cancer, infectious diseases (AIDS, leprosy, hepatitis, etc.) autoimmune, neuropsychiatric and many other diseases (Tables 1–4) with different SNPs can be made as relevant pharmacogenomic targets for drug therapy.

Disclosure statement

The authors declare that there is no conflict of interest.

References

[1] Chakravarti A. To a future of genetic medicine. Nature 2001;409(6822):822–3.

[2] Kato N, Takeuchi F, Tabara Y, et al. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. Nat Genet 2011;43:531–8.

[3] Soranzo N, Spector TD, Mangino M, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. Nat Genet 2009;41:1182–90.

[4] Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39:906–13.

[5] Frazer KA, Ballinger DG, Cox DR, et al. International Hap-Map Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–61.

[6] Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science 2002;296:2225–9.

[7] Spencer CC, Su Z, Donnelly P, Marchini J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. PLoS Genet 2009;5:e1000477.

[8] Sirén J, Marttinen P, Corander J. Reconstructing population histories from single nucleotide polymorphism data. Mol Biol Evol 2011;28:673–83.

[9] Hunter DJ, Kraft P. Drinking from the fire hose–statistical issues in genomewide association studies. N Engl J Med 2007;357:436–9.

[10] Hoover RN. The evolution of epidemiologic research: from cottage industry to "big" science. Epidemiology 2007;18:13–7.

[11] Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007;447:1087–93.

[12] Thomas G, Jacobs KB, Yeager M, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet 2008;40:310–5.

[13] Pearson TA, Manolio TA. How to interpret a genome-wide association study. JAMA 2008;299:1335–44.

[14] Giacopelli F, Marciano R, Pistorio A, Catarsi P, Canini S, Karsenty G, et al. Polymorphisms in the osteopontin promoter affect its transcriptional activity. Physiol Genomics 2004;20:87–96.

[15] Gudmundsson J, Sulem P, Gudbjartsson DF, et al. Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. Nat Genet 2009;41:460–4.

[16] Egli RJ, Southam L, Wilkins JM, Lorenzen I, Pombo-Suarez M, Gonzalez A, et al. Functional analysis of the osteoarthritis susceptibility-associated GDF5 regulatory polymorphism. Arthritis Rheum 2009;60:2055–64.

[17] Zhang H, Gelernter J, Gruen JR, Kranzler HR, Herman AI, Simen AA. Functional impact of a single-nucleotide polymorphism in the OPRD1 promoter region. J Hum Genet 2010;55:278–84.

[18] Salavaggione OE, Wang L, Wiepert M, Yee VC, Weinshilboum RM. Thiopurine S-methyltransferase pharmacogenetics: variant allele functional and comparative genomics. Pharmacogenet Genomics 2005;15:801–15.

[19] Lennard L, Van Loon JA, Weinshilboum RM. Pharmacogenetics of acute azathioprine toxicity: relationship to thiopurine

methyltransferase genetic polymorphism. Clin Pharmacol Ther 1989;46:149–54.

[20] Huang P, Chen C, Mague SD, Blendy JA, Liu–Chen LY. A common single nucleotide polymorphism A118G of the μ opioid receptor alters its N-glycosylation and protein stability. Biochem J 2012;441:379–86.

[21] Curran ME, Splawski I, Timothy KW, Vincent GM, Green ED, Keating MT. A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. Cell 1995;80:795–803.

[22] Gentile S, Martin N, Scappini E, Williams J, Erxleben C, Armstrong DL. The human ERG1 channel polymorphism, K897T, creates a phosphorylation site that inhibits channel activity. Proc Natl Acad Sci USA 2008;105:14704–8.

[23] Hardy J, Singleton A. Genomewide association studies and human disease. N Engl J Med 2009;360:1759–68.

[24] Manolio TA, Brooks LD, Collins FSA. A HapMap harvest of insights into the genetics of common disease. J Clin Invest 2008;118:1590–605.

[25] Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007;316:1341–5.

[26] Kan YW, Dozy AM. Polymorphism of DNA sequence adjacent to human beta–globin structural gene: relationship to sickle mutation. Proc Natl Acad Sci USA 1978;75:5631–5.

[27] Buchanan CC, Torstenson ES, Bush WS, Ritchie MD. A comparison of cataloged variation between International Hap-Map Consortium and 1000 Genomes Project data. J Am Med Inform Assoc 2012;19:289–94.

[28] Schadt EE. Molecular networks as sensors and drivers of common human diseases. Nature 2009;10461:218–23.

[29] Zhao S, Iyengar R. Systems pharmacology: network analysis to identify multiscale mechanisms of drug action. Annu Rev Pharmacol Toxicol 2012;52:505–21.

[30] Sakata T, Anzai N, Kimura T, Miura D, Fukutomi T, et al. Functional analysis of human organic cation transporter OCT3 SLC22A3 polymorphisms. J Pharmacol Sci 2010; 113:263–6.

[31] Hartmann T, Terao M, Garattini E, Teutloff C, Alfaro JF, Jones JP, et al. The impact of single nucleotide polymorphisms on human aldehyde oxidase. Drug Metab Dispos 2012;40:856–64.

[32] Johnson JA, Gong L, Whirl-Carrillo M, Gage BF, Scott SA, Stein CM, et al. Clinical pharmacogenetics implementation consortium guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. Clin Pharmacol Ther 2011;90:625–9.

[33] Lucena MI, Andrade RJ, Martinez C, Ulzurrun E, Garcia-Martin E, Borraz Y, et al. Glutathione S-transferase m1 and t1 null geno-types increase susceptibility to idio-syncratic drug-induced liver injury. Hepatology 2008;48:588–96.

[34] Lucena MI, García–Martín E, Andrade RJ, et al. Mitochondrial superoxide dismutase and glutathione peroxidase in idiosyncratic drug-induced liver injury. Hepatology 2010;52:303–12.

[35] Andrade RJ, Agundez JA, Lucena MI, Martinez C, Cueto R, Garcia-Martin E. Pharmacogenomics in drug induced liver injury. Curr Drug Metab 2009;10:956–70.

[36] Kim HC, Lee JY, Sung H, et al. A genome-wide association study identifies a breast cancer risk variant in ERBB4 at 2q34: results from the Seoul Breast Cancer Study. Breast Cancer Res 2012;14:R56.

[37] Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007;39:870–4.

[38] Stacey SN, Manolescu A, Sulem P, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 2007;39:865–9.

[39] Stacey SN, Gudbjartsson DF, Sulem P, et al. Common variants on 1p36 and 1q42 are associated with cutaneous basal cell carcinoma but not with melanoma or pigmentation traits. Nat Genet 2008;40:1313–8.

[40] Stacey SN, Sulem P, Masson G, et al. New common variants affecting susceptibility to basal cell carcinoma. Nat Genet 2009;41:909–14.

[41] Gudmundsson J, Sulem P, Manolescu A, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet 2007;39:631–7.

[42] Yeager M, Orr N, Hayes RB, Jacobs KB, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet 2007;39:645–9.

[43] Gudmundsson J, Sulem P, Steinthorsdottir V, et al. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet 2007;39:977–83.

[44] Eeles RA, Kote-Jarai Z, Al Olama AA, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet 2009;41:1116–21.

[45] Zanke BW, Greenwood CM, Rangrej J, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat Genet 2007;39:989–94.

[46] Broderick P, Carvajal–Carmona L, Pittman AM, et al. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. Nat Genet 2007;39:1315–7.

[47] Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 2008;452:633–7.

[48] Landi MT, Chatterjee N, Yu K, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet 2009;85:679–91.

[49] Brown KM, Macgregor S, Montgomery GW, et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. Nat Genet 2008;40:838–40.

[50] Bishop DT, Demenais F, Iles MM, et al. Genome-wide association study identifies three loci associated with melanoma risk. Nat Genet 2009;41:920–5.

[51] Maris JM, Mosse YP, Bradfield JP, et al. Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. N Engl J Med 2008;358:2585–93.

[52] Capasso M, Devoto M, Hou C, et al. Common variations in BARD1 influence susceptibility to high-risk neuroblastoma. Nat Genet 2009;41:718–23.

[53] Song H, Ramus SJ, Tyrer J, et al. A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. Nat Genet 2009;419:996–1000.

[54] Mead S, Poulter M, Uphill J, et al. Genetic risk factors for variant Creutzfeldt–Jakob disease: a genome-wide association study. Lancet Neurol 2009;8:57–66.

[55] Khor CC, Chau TN, Pang J, et al. Genome-wide association study identifies susceptibility loci for dengue shock syndrome at MICB and PLCE1. Nat Genet 2011;43:1139–41.

[56] Kamatani Y, Wattanapokayakit S, Ochi H, et al. A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. Nat Genet 2009;41:591–5.

[57] Rauch A, Kutalik Z, Descombes P, et al. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. Gastroenterology 2010;138:1338–45.

[58] Fellay J, Shianna KV, Ge D, et al. A whole-genome association study of major determinants for host control of HIV-1. Science 2007;317:944–7.

[59] Fellay J, Ge D, Shianna KV, Colombo S, et al. Common genetic variation and the control of HIV-1 in humans. PLoS Genet 2009;5:e1000791.

[60] Pelak K, Goldstein DB, Walley NM, et al. Host determinants of HIV-1 control in African Americans. J Infect Dis 2010;201:1141–9.

[61] Pereyra F, Jia X, McLaren PJ, et al. The International HIV Controllers Study. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science 2010;330:1551–7.

[62] Troyer JL, Nelson GW, Lautenberger JA, et al. Genome-wide association study implicates PARD3B-based AIDS restriction. J Infect Dis 2011;203:1491–502.

[63] Limou S, Le Clerc S, Coulonges C, et al. Genomewide association study of an AIDS-nonprogression cohort emphasizes the role played by HLA genes ANRS genomewide association study 02. J Infect Dis 2009;199:419–26.

[64] Limou S, Coulonges C, Herbeck JT, et al. Multiple-cohort genetic association study reveals CXCR6 as a new chemokine receptor involved in long-term nonprogression to AIDS. J Infect Dis 2010;202:908–15.

[65] Zhang FR, Huang W, Chen SM, et al. Genomewide association study of leprosy. N Engl J Med 2009;361:2609–18.

[66] Davila S, Wright VJ, Khor CC, et al. Genome-wide association study identifies variants in the CFH region associated with host susceptibility to meningococcal disease. Nat Genet 2010;42:772–6.

[67] Jallow M, Teo YY, Small KS, et al. Genome-wide and fine-resolution association analysis of malaria in West Africa. Nat. Genet 2009;41:657–65.

[68] Thye T, Vannberg FO, Wong SH, et al. Genome--wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2. Nat Genet 2010;42:739–41.

[69] Stahl EA, Raychaudhuri S, Remmers EF, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 2010;42:508–14.

[70] Plenge RM, Seielstad M, Padyukov L, et al. TRAF1-C5 as a risk locus for rheumatoid arthritis – a genomewide study. N Engl J Med 2007;357:1199–209.

[71] Plenge RM, Cotsapas C, Davies L, et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007;39:1477–82.

[72] Kozyrev SV, Abelson AK, Wojcik J, et al. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet 2008;40:211–6.

[73] Gudbjartsson DF, Arnar DO, Helgadottir A, et al. Variants conferring risk of atrial fibrillation on chromosome 4q25. Nature 2007;448:353–7.

[74] Burton PR, Clayton DG, Cardon LR, et al. Wellcome Trust Case Control Consortium (WTCCC). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–78.

[75] Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. N Engl J Med 2007;357:443–53.

[76] Harrison SC, Cooper JA, Li K, et al. Association of a sequence variant in DAB2IP with coronary heart disease. Eur Heart J 2012;33:881–8.

[77] Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007;316:1491–3.

[78] Smyth DJ, Cooper JD, Howson JM, Clarke P, Downes K, Mistry T, et al. FUT2 nonsecretor status links type 1 diabetes susceptibility and resistance to infection. Diabetes 2011;60:3081–4.

[79] Todd JA, Walker NM, Cooper JD, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39:857–64.

[80] Hakonarson H, Grant SF, Bradfield JP, et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 2007;448:591–4.

[81] van Vliet-Ostaptchouk JV, van Haeften TW, Landman GW, et al. Common variants in the type 2 diabetes KCNQ1 gene are associated with impairments in insulin secretion during hyperglycaemic glucose clamp. PLoS One 2012;7:e32148.

[82] Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007;445:881–5.

[83] Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007;316:1336–41.

[84] Saxena R, Voight BF, Lyssenko V, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007;316:1331–6.

[85] Van Heel DA, Franke L, Hunt KA, et al. A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat Genet 2007;39:827–9.

[86] Libioulle C, Louis E, Hansoul S, Sandor C, et al. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. PLoS Genet 2007;3:e58.

[87] Parkes M, Barrett JC, Prescott NJ, et al. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. Nat Genet 2007;39:830–2.

[88] Rioux JD, Xavier RJ, Taylor KD, et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. Nat Genet 2007;39:596–604.

[89] Buch S, Schafmayer C, Völzke H, et al. A genome-wide association scan identifies the hepatic cholesterol transporter ABCG8 as a susceptibility factor for human gallstone disease. Nat Genet 2007;39:995–9.

[90] Duerr RH, Taylor KD, Brant SR, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 2006;314:1461–3.

[91] Kathiresan S, Melander O, Guiducci C, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat Genet 2008;40:189–97.

[92] Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet 2008;40:161–9.

[93] Wallace C, Newhouse SJ, Braund P, et al. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. Am J Hum Genet 2008;82:139–49.

[94] Kooner JS, Chambers JC, Aguilar-Salinas CA, et al. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. Nat Genet 2008;40:149–51.

[95] van Es MA, van Vught PW, Blauw HM, et al. Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis. Nat Genet 2008;40:29–31.

[96] Reiman EM, Webster JA, Myers AJ, et al. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. Neuron 2007;54:713–20.

[97] Baum AE, Akula N, Cabanero M, et al. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. Mol Psychiatry 2008;13:197–207.

[98] Al Jumah M, Al Balwi M, Hussein M, Kojan S, Al Khathaami A, Al Fawaz M, et al. Association of SNPs rs6498169 and rs10984447 with multiple sclerosis in Saudi patients: a model of the usefulness of familial aggregates in identifying genetic linkage in a multifactorial disease. Mult Scler 2012. http://dx.doi.org/10.1177/1352458512440832.

[99] Hafler DA, Compston A, Sawcer S, et al. International Multiple Sclerosis Genetics Consortium. Risk alleles for multiple sclerosis

identified by a genomewide study. N Engl J Med 2007;357:851–62.

[100] Stefansson H, Rye DB, Hicks A, et al. A genetic risk factor for periodic limb movements in sleep. N Engl J Med 2007;357:639–47.

[101] Winkelmann J, Schormair B, Lichtner P, et al. Genome-wide association study of restless legs syndrome identifies common variants in three genomic regions. Nat Genet 2007;39:1000–6.

[102] Lencz T, Morgan TV, Athanasiou M, Dain B, Reed CR, Kane JM, et al. Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. Mol Psychiatry 2007;12:572–80.

[103] Schymick JC, Scholz SW, Fung HC, et al. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. Lancet Neurol 2007;6:322–8.

[104] Food and Drug Administration, 2012. Table of pharmacogenomic biomarkers in drug labels. Available at: http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm.

[105] Ando Y, Saka H, Ando M, Sawa T, Muro K, Ueoka H, et al. Polymorphisms of UDP-glucuronosyltransferase gene and irinotecan toxicity: a pharmacogenetic analysis. Cancer Res 2000;60:6921–6.

[106] Mallal S, Phillips E, Carosi G, et al. HLA-B*5701 screening for hypersensitivity to abacavir. N Engl J Med 2008;358:568–79.

[107] Halder I, Shriver MD. Measuring and using admixture to study the genetics of complex diseases. Hum Genomics 2003;1:52–62.

[108] Hu D, Ziv E. Confounding in genetic association studies and its solutions. Methods Mol Biol 2008;448:31–9.

[109] Liu C, Batliwalla F, Li W, et al. Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in rheumatoid arthritis. Mol Med 2008;14:575–81.

[110] Barrett JC, Hansoul S, Nicolae DL, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 2008;40:955–62.

[111] Yang Q, Cui J, Chazaro I, Cupples LA, Demissie S. Power and type I error rate of false discovery rate approaches in genome-wide association studies. BMC Genet 2005;6(Suppl 1):S134.

[112] Van der Helmvan Mil AH, Padyukov L, Toes RE, Klareskog L, Huizinga TW. Genome-wide single-nucleotide polymorphism studies in rheumatology: hype or hope? Arthritis Rheum 2008;58:2591–7.

[113] Johnson JA, Lima JJ. Drug receptor/effector polymorphisms and pharmacogenetics: current status and challenges. Pharmacogenetics 2003;13:525–34.

[114] Jacobson EM, Concepcion T, Oashi Y, Tomer A. Graves' disease-associated Kozak sequence single-nucleotide polymorphism enhances the efficiency of CD40 gene translation: a case for translational pathophysiology. Endocrinology 2005;146(6):2684–91.

[115] Wachtel C, Manley JL. Splicing of mRNA precursors: the role of RNAs and proteins in catalysis. Mol Biosyst 2009;5:311–6.

[116] Pagani F, Stuani C, Tzetis M, Kanavakis E, Efthymiadou A, Doudounakis S, et al. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon. Hum Mol Genet 2003;12(10):1111–20.

[117] Giacopelli F, Marciano R, Pistorio A, Catarsi P, Canini S, Karsenty G, et al. Polymorphisms in the osteopontin promoter affect its transcriptional activity. Physiol Genomics 2004;20(1):87–96.

[118] Wang Z, Moult J. SNPs, protein structure, and disease. Hum Mutat 2001;4:263–70.

[119] Daly AK, Donaldson PT, Bhatnagar P, Shen Y, Pe'er I, et al. HLA-B*5701 genotype is a major determinant of drug–induced liver injury due to flucloxacillin. Nat Genet 2009;41:816–9.

[120] Kindmark A, Jawaid A, Harbron CG, Barratt BJ, Bengtsson OF, et al. Genome-wide pharmacogenetic investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. Pharmacogenomics J 2008;8:186–95.

[121] Lucena MI, Molokhia M, Shen Y, Urban TJ, Aithal GP, et al. Susceptibility to amoxicillinclavulanate–induced liver injury is influenced by multipleHLA class I and class II alleles. Gastroenterology 2011;141:338–47.

[122] Singer JB, Lewitzky S, Leroy E, Yang F, Zhao X, et al. A genome-wide study identifies HLA alleles associated with lumiracoxib-related liver injury. Nat Genet 2010;42:711–4.