Experimental validation of data mined single nucleotide polymorphisms from several databases and consecutive dbSNP builds

Hege Edvardsen[a], Grethe Irene Grenaker Alnæs[a], Anya Tsalenko[b], Tanya Mulcahy[c], Anton Yuryev[d], Marie Lindersson[e], Sigbjørn Lien[f], Stig Omholt[e], Ann-Christine Syvänen[e], Anne-Lise Børresen-Dale[a] and Vessela N. Kristensen[a]

Rapid development in the annotation of human genetic variation has increased the numbers of single nucleotide polymorphisms (SNPs) in candidate genes by several orders of magnitude. The selection of both useful target SNPs for disease–gene association studies and SNPs associated with the treatment response is therefore an increasingly challenging task. We describe a workflow for selecting SNPs based on their putative function and frequency in candidate genes extracted from PubMed resources. The annotation of each SNP and its frequency in a Caucasian population was assessed in several databases. Approximately 4000 SNPs were identified from an initial 233 candidate genes. In a case study, we performed actual genotyping of 1030 of these SNPs in 213 genes and obtained 710 successfully genotyped SNPs. Using the flow-chart outlined here, only 87 SNPs were monomorphic (approximately 12%). This study reports the frequency of SNPs in a Caucasian population, selected *in silico*, using a candidate gene approach and validated by actually genotyping 193 individuals. The selected genotypes represent a valuable set of verified candidate SNPs for pharmacogenetic studies in Caucasian populations.

Pharmacogenetics and Genomics **16:207–217 © 2006 Lippincott Williams & Wilkins.**

Pharmacogenetics and Genomics 2006, **16**:207–217

Keywords: cancer pharmacogenetics, data mining, dbSNP, SNP, SNP500, SNPper

[a]Department of Genetics, Institute for Cancer Research, the Norwegian Radium Hospital, Faculty Division of the University of Oslo, Oslo, Norway, [b]Agilent Technologies, Palo Alto, California, USA, [c]Beckman Coulter Inc., Fullerton, California, USA, [d]Ariadne Genomics Inc., Rockville Maryland, USA, [e]Department of Medical Sciences, Uppsala University, Uppsala, Sweden and [f]Center for Integrative Genetics, the Norwegian University of Life Sciences, Aas, Norway

Correspondence and requests for reprints to Hege Edvardsen, Department of Genetics, the Norwegian Radium Hospital, Montebello, 0310 Oslo, Norway
Fax: + 47 22934440; e-mail: hedvards@medisin.uio.no

Sponsorship: This project was supported by The Norwegian Cancer Association (project number D-03067 and D-99061) and The National Programme for Research in Functional Genomics in Norway (FUGE, project numbers 151924/150 and 15204/150), funded by The Research Council of Norway. Genotyping was partly performed at the Uppsala SNP technology platform, which is funded by the Swedish Wallenberg Consortium North (WCN).

Received 20 June 2005 **Accepted** 1 November 2005

Introduction

Genotype profiling of patients before administering therapy may help to identify those individuals who may not respond to a given treatment or who may suffer potentially severe side-effects. Although whole-genome single nucleotide polymorphism (SNP) analysis may potentially convey important information about the genetic structure of various populations, the candidate gene approach has not yet been exhaustively explored due to the still limited, but increasing, number of SNPs discovered in candidate genes [1]. At present, genome-wide scanning is unlikely to be realistic in clinical setting where a smaller number of SNPs must be analysed due to the costs involved. In the present study, the candidate gene approach was chosen, with emphasis being placed on the genes involved in the metabolism and signalling of reactive oxygen species (ROS), DNA repair and apoptosis to investigate an association of a known biochemical background with a specific treatment response.

When performing a large and expensive genotyping study, it is important to reduce the risk of evaluating too many false or monomorphic SNPs. The multiple SNP resources available on the Internet make searching for new SNPs easier but, at the same time, designing genotyping projects becomes more demanding. Although the time spent in the laboratory is greatly reduced by high-throughput genotyping systems, *in silico* mining and analysing data become the most time-consuming steps. We assessed differences in the quality and quantity of input from various databases and dbSNP builds (the Single Nucleotide Polymorphism database; a public-domain archive of SNPs) on a particular set of selected and genotyped SNPs. The quality and completeness of the public databases are variable; thus, by combining information from different databases, we aimed to obtain a set of SNPs with a high genotyping success rate. One of the approaches used when navigating a large number of SNPs is based on an assumption of functionality and

1744-6872 © 2006 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

frequency [2,3]. In the present study, we report the validation of a workflow for genotyping a large number of markers with an actual experiment including 193 women with breast cancer.

Materials and methods

The 193 women genotyped in this project had received radiation therapy or chemotherapy for their breast cancer, either as a sole treatment or as a complement to surgery. Each patient provided their informed consent and the study was approved by the ethical and protocol committee of The Norwegian Radium Hospital. The study aimed to extract and validate a panel of SNPs suitable for future association studies in this cohort.

A high-throughput solid-phase array method was used to perform the genotyping using the GenomeLab SNPstream genotyping system (Beckman Coulter, Fullerton, California, USA), where the amplified polymerase chain reaction (PCR) product (in 12-plex reactions) is used as a template in a second, modified, single-primer mini-sequencing reaction. Single-base extension is sufficient for the discrimination of the variant base and to achieve a detectable difference in allele-specific extension products [4].

Selecting genes and SNPs

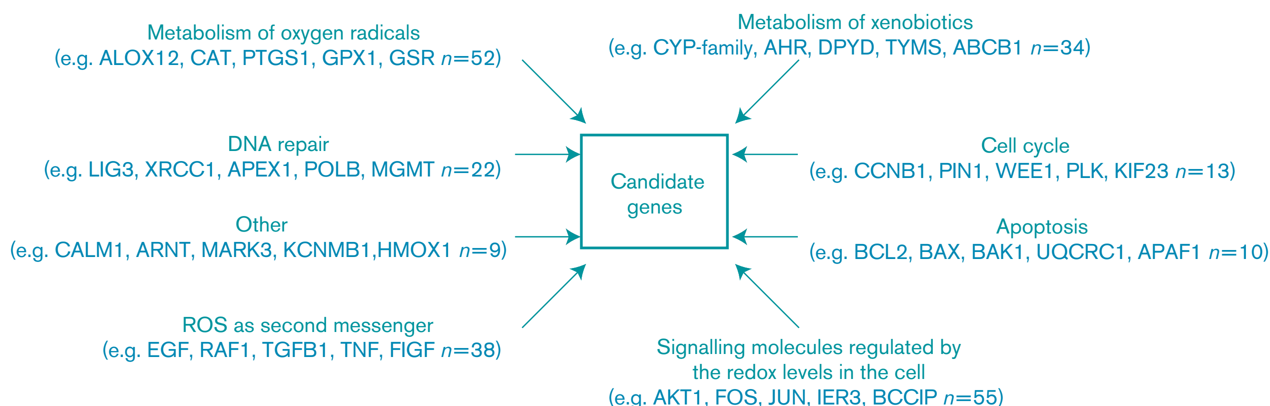
Candidate genes

Radiation therapy and chemotherapy both exert their anti-neoplastic effect by either directly attacking cellular macromolecules, including DNA, or indirectly, by generating ROS and their by-products. Therefore, the genes selected to create the genotype profile of patients treated with radiation therapy and chemotherapy are all involved in regulating the redox level in the cells, in signalling or in DNA damage repair caused by ROS [5–9]. Figure 1 shows the biochemical pathways investigated and also provides examples of the selected genes. A total of 233 genes were extracted from approximately 4000 MEDLINE entries. Some of these genes were selected as primary targets whereas others were selected based on pathway association with the primary genes. Some genes could be localized to more than one pathway, thus forming a network; their pathway classification was based on literature studies conducted during the SNP mining process. A complete list of all candidate genes screened for SNPs *in silico*, their chromosomal localization and pathway classification are presented in Table 1.

A combination of public and proprietary resources was used in the SNP search (Fig. 2). After establishing the official gene names in the HUGO nomenclature database (http://www.gene.ucl.ac.uk/nomenclature/), we extracted SNPs from within each of these genes from multiple resources. This resulted in approximately 4000 SNPs, which were subjected to further selection with the following inclusion criteria: (i) SNPs with the potential to influence function or expression levels (cSNPs and 5′-3′ SNPs) and/or (ii) with a known frequency in a Caucasian population. The exact positions of the SNPs in the annotated genes were reconfirmed, sequences (with 150–200 bases upstream and downstream of the SNP) were blasted and alu-repeats and non-specific areas were masked. The resulting file was submitted to an online primer design program (http://www.autoprimer.com/) which generated designs for 1030 SNPs.

Initially, SNP frequencies were extracted from the Chromosome Browser (software developed by Orchid Biosciences, Princeton, New Jersey, USA; currently the property of Beckman Coulter, Fullerton, California, USA), as well as the databases of Applied Biosystems (Foster

Fig. 1



An overview of the functional pathways from which candidate genes were extracted. Examples and the number of genes mined for single nucleotide polymorphisms (from a total of 233 genes) in each group are provided. The pathways are involved in the metabolism of reactive oxygen species (ROS), response to damage caused by ROS or influenced by the redox level in the cell.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Table 1 List of genes mined for single nucleotide polymorphisms showing the HUGO approved gene symbol, full gene name, chromosomal localization and internal pathway classification based on literature studies

HUGO approved gene name	Full gene name	Chromosomal localization	Internal classification
APAF1	Apoptotic protease activating factor isoform b	chr12:q23.1	Apoptosis
BAK1	BCL2-antagonist/killer 1	chr20:q11.21	Apoptosis
BAX	BCL2-associated X protein isoform gamma	chr19:q13.33	Apoptosis
BBC3	BCL2 binding component 3	chr19:q13.32	Apoptosis
BCL2	B-cell lymphoma protein 2 alpha isoform	chr18:q21.33	Apoptosis
MAP2K4	Mitogen-activated protein kinase kinase 4	chr17:p12	Apoptosis
MAP2K7	Mitogen-activated protein kinase kinase 7	chr19:p13.2	Apoptosis
MAP3K3	Mitogen-activated protein kinase kinase kinase 3	chr17:q23.3	Apoptosis
UQCRC1	Ubiquinol-cytochrome c reductase core protein I	chr3:p21.31	Apoptosis
UQCRC2	Ubiquinol-cytochrome c reductase core protein II	chr16:p12.1	Apoptosis
CCNB1	Cyclin B1	chr5:q13.2	Cell cycle
CCNG1	Cyclin G1	chr5:q34	Cell cycle
CDC25A	Cell division cycle 25A isoform a	chr3:p21.31	Cell cycle
CDC25B	Cell division cycle 25B isoform 5	chr20:p13	Cell cycle
CDC25C	Cell division cycle 25C protein isoform a	chr5:q31.2	Cell cycle
CENPE	Centromere protein E, 312 kDa	chr4:q24	Cell cycle
KIF23	Kinesin family member 23 isoform 2	chr15:q23	Cell cycle
KIF2C	Kinesin family member 2C	chr1:p34.1	Cell cycle
MDM2	Mouse double minute 2 homologue isoform MDM2	chr12:q15	Cell cycle
PIN1	Protein (peptidyl-prolyl cis/trans isomerase) NIMA-interacting 1	chr19:p13.2	Cell cycle
PLK1	Polo-like kinase 1 (Drosophila)	chr16:p12.1	Cell cycle
RRM2B	Ribonucleotide reductase M2 B (TP53 inducible)	chr8:q22.3	Cell cycle
WEE1	WEE1 homologue (S. pombe)	chr11:p15.4	Cell cycle
APEX1	APEX nuclease (multifunctional DNA repair enzyme) 1	chr14:q11.2	DNA repair
DDB1	Damage-specific DNA binding protein 1 (127 kDa)	chr11:q12.2	DNA repair
DDB2	Damage-specific DNA binding protein 2 (48 kDa)	chr11:p11.2	DNA repair
GADD45A	Growth arrest and DNA-damage-inducible, alpha	chr1:p31.2	DNA repair
GADD45B	Growth arrest and DNA-damage-inducible, beta	chr19:p13.3	DNA repair
GADD45G	Growth arrest and DNA-damage-inducible, gamma	chr9:q22.2	DNA repair
LIG3	Ligase III, DNA, ATP-dependent isoform beta	chr17:q12	DNA repair
LIG4	Ligase IV, DNA, ATP-dependent	chr13:q33.3	DNA repair
MGMT	O-6-methylguanine-DNA methyltransferase	chr10:q26.3	DNA repair
PCNA	Proliferating cell nuclear antigen	chr20:p12.3	DNA repair
POLB	Polymerase (DNA directed), beta	chr8:p11.21	DNA repair
POLD1	Polymerase (DNA directed), delta 1, catalytic subunit 125 kDa	chr19:q13.33	DNA repair
POLD2	Polymerase (DNA directed), delta 2, regulatory subunit 50 kDa	chr7:p13	DNA repair
POLE2	Polymerase (DNA directed), epsilon 2 (p59 subunit)	chr14:q21.3	DNA repair
POLE3	Polymerase (DNA directed), epsilon 3 (p17 subunit)	chr9:q32	DNA repair
RAD51	RAD51 homologue protein isoform 1	chr15:q15.1	DNA repair
TOP1	Topoisomerase (DNA) I	chr20:q12	DNA repair
TOP2A	DNA topoisomerase II, alpha isozyme	chr17:q21.2	DNA repair
TOP2B	DNA topoisomerase II, beta isozyme	chr3:p24.2	DNA repair
XPC	Xeroderma pigmentosum, complementation group C	chr3:p25.1	DNA repair
XRCC1	X-ray repair complementing defective repair in Chinese hamster cells 1	chr19:q13.31	DNA repair
XRCC4	X-ray repair complementing defective repair in Chinese hamster cells 4	chr5:q14.2	DNA repair
ALOX12	Arachidonate 12-lipoxygenase	chr17:p13.1	Metabolism of oxygen radicals
ALOX15	Arachidonate 15-lipoxygenase	chr17:p13.2	Metabolism of oxygen radicals
ALOX15B	Arachidonate 15-lipoxygenase, second type	chr17:p13.1	Metabolism of oxygen radicals
ALOXE3	Arachidonate lipoxygenase 3	chr17:p13.1	Metabolism of oxygen radicals
CAT	Catalase	chr11:p13	Metabolism of oxygen radicals
CCS	Copper chaperone for superoxide dismutase	chr11:q13.2	Metabolism of oxygen radicals
GPX1	Glutathione peroxidase 1 isoform 1	chr3:p21.31	Metabolism of oxygen radicals
GPX2	Glutathione peroxidase 2 (gastrointestinal)	chr14:q23.3	Metabolism of oxygen radicals
GPX3	Plasma glutathione peroxidase 3 precursor	chr5:q33.1	Metabolism of oxygen radicals
GPX4	Glutathione peroxidase 4 (phospholipids hydroperoxidase)	chr19:p13.3	Metabolism of oxygen radicals
GPX5	Glutathione peroxidase 5 precursor, isoform 1	chr6:p22.1	Metabolism of oxygen radicals
GSR	Glutathione reductase	chr8:p12	Metabolism of oxygen radicals
NDUFA1	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1, 7.5 kDa	chrX:q24	Metabolism of oxygen radicals
NDUFA10	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 10, 42 kDa	chr2:q37.3	Metabolism of oxygen radicals
NDUFA2	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2, 8 kDa	chr5:q31.3	Metabolism of oxygen radicals
NDUFA3	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3, 9 kDa	chr19:q13.42	Metabolism of oxygen radicals
NDUFA4	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4, 9 kDa	chr7:p21.3	Metabolism of oxygen radicals
NDUFA5	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5, 13 kDa	chr7:q31.32	Metabolism of oxygen radicals
NDUFA6	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 6, 14 kDa	chr22:q13.2	Metabolism of oxygen radicals
NDUFA7	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 7, 14.5 kDa	chr19:p13.2	Metabolism of oxygen radicals
NDUFA8	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 8, 19 kDa	chr9:q33.2	Metabolism of oxygen radicals
NDUFA9	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 9, 39 kDa	chr12:p13.32	Metabolism of oxygen radicals
NDUFB1	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 1, 7 kDa	chr14:q32.12	Metabolism of oxygen radicals
NDUFB10	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 10, 22 kDa	chr16:p13.3	Metabolism of oxygen radicals
NDUFB2	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 2, 8 kDa	chr7:q34	Metabolism of oxygen radicals
NDUFB3	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 3, 12 kDa	chr2:q33.1	Metabolism of oxygen radicals
NDUFB4	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 4, 15 kDa	chr3:q13.33	Metabolism of oxygen radicals
NDUFB5	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 5, 16 kDa	chr3:q26.33	Metabolism of oxygen radicals

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Table 1 (continued)

HUGO approved gene name	Full gene name	Chromosomal localization	Internal classification
NDUFB6	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 6, 17 kDa	chr9:p21.1	Metabolism of oxygen radicals
NDUFB7	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 7, 18 kDa	chr19:p13.12	Metabolism of oxygen radicals
NDUFB8	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 8, 19 kDa	chr10:q24.31	Metabolism of oxygen radicals
NDUFB9	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 9, 22 kDa	chr8:q24.13	Metabolism of oxygen radicals
NOX1	NADPH oxidase 1 isoform long	chrX:q22.1	Metabolism of oxygen radicals
NOX3	NADPH oxidase 3	chr6:q25.3	Metabolism of oxygen radicals
NOX4	NADPH oxidase 4	chr11:q14.3	Metabolism of oxygen radicals
NOX5	NADPH oxidase, EF hand calcium-binding domain 5	chr15:q23	Metabolism of oxygen radicals
PRDX2	Peroxiredoxin 2 isoform a	chr19:p13.13	Metabolism of oxygen radicals
PRDX2	Peroxiredoxin 2 isoform a	chr19:p13.13	Metabolism of oxygen radicals
PRDX4	Thioredoxin peroxidase	chrX:p22.11	Metabolism of oxygen radicals
PTGS1	Prostaglandin-endoperoxide synthase 1 isoform 1	chr9:q33.2	Metabolism of oxygen radicals
PTGS2	Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	chr1:q31.1	Metabolism of oxygen radicals
SOD1	Superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))	chr21:q22.11	Metabolism of oxygen radicals
SOD2	Superoxide dismutase 2, mitochondrial	chr6:q25.3	Metabolism of oxygen radicals
SOD3	Superoxide dismutase 3, extracellular	chr4:p15.2	Metabolism of oxygen radicals
TXN	Thioredoxin	chr9:q31.3	Metabolism of oxygen radicals
TXN2	Thioredoxin 2 precursor	chr22:q12.3	Metabolism of oxygen radicals
TXNDC	Thioredoxin domain containing	chr14:q22.1	Metabolism of oxygen radicals
TXNIP	Thioredoxin interacting protein	chr1:q21.1	Metabolism of oxygen radicals
TXNRD1	Thioredoxin reductase 1	chr12:q23.3	Metabolism of oxygen radicals
TXNRD2	Thioredoxin reductase 2 isoform 3	chr22:q11.21	Metabolism of oxygen radicals
TXNRD3	Thioredoxin reductase 3	chr3:p13-q13.33	Metabolism of oxygen radicals
XDH	Xanthine dehydrogenase	chr2:p23.1	Metabolism of oxygen radicals
ABCB1	ATP-binding cassette, sub-family B (MDR/TAP), member 1	chr7:q21.12	Metabolism of xenobiotics
ABCC1	ATP-binding cassette, sub-family C (CFTR/MRP), member 1	chr16:p13.11	Metabolism of xenobiotics
AHR	Aryl hydrocarbon receptor	chr7:p21.1	Metabolism of xenobiotics
AKR7A2	Aldo-keto reductase family 7, member A2 (aflatoxin aldehyde reductase)	chr1:p36.13	Metabolism of xenobiotics
CSK	c-src tyrosine kinase	chr15:q24.1	Metabolism of xenobiotics
CYP2C19	Cytochrome P450, family 2, subfamily C, polypeptide 19	chr10:q23.33	Metabolism of xenobiotics
CYP2C8	Cytochrome P450, family 2, subfamily C, polypeptide 8	chr10:q23.33	Metabolism of xenobiotics
CYP2C9	Cytochrome P450, family 2, subfamily C, polypeptide 9	chr10:q23.33	Metabolism of xenobiotics
CYP2E1	Cytochrome P450, family 2, subfamily E, polypeptide 1	chr10:q26.3	Metabolism of xenobiotics
CYP3A4	Cytochrome P450, family 3, subfamily A, polypeptide 4	chr7:q22.1	Metabolism of xenobiotics
DPYD	Dihydropyrimidine dehydrogenase	chr1:p21.3	Metabolism of xenobiotics
EPHX1	Epoxide hydrolase 1, microsomal (xenobiotic)	chr1:q42.12	Metabolism of xenobiotics
EPHX2	Epoxide hydrolase 2, cytoplasmic	chr8:p21.1	Metabolism of xenobiotics
GCLC	Glutamate-cysteine ligase, catalytic subunit	chr6:p12.1	Metabolism of xenobiotics
GCLM	Glutamate-cysteine ligase regulatory protein	chr1:p22.1	Metabolism of xenobiotics
GSTA1	Glutathione S-transferase A1	chr6:p12.2	Metabolism of xenobiotics
GSTA2	Glutathione S-transferase A2	chr6:p12.2	Metabolism of xenobiotics
GSTA3	Glutathione S-transferase A3	chr6:p12.2	Metabolism of xenobiotics
GSTA4	Glutathione S-transferase A4	chr6:p12.1	Metabolism of xenobiotics
GSTM1	Glutathione S-transferase M1 isoform 1	chr1:p13.3	Metabolism of xenobiotics
GSTM2	Glutathione S-transferase M2 (muscle)	chr1:p13.3	Metabolism of xenobiotics
GSTM3	Glutathione S-transferase M3 (brain)	chr1:p13.3	Metabolism of xenobiotics
GSTM4	Glutathione S-transferase M4 isoform 1	chr1:p13.3	Metabolism of xenobiotics
GSTM5	Glutathione S-transferase M5	chr1:p13.3	Metabolism of xenobiotics
GSTP1	Glutathione S-transferase pi	chr11:q13.2	Metabolism of xenobiotics
GSTT1	Glutathione S-transferase theta 1	chr22:q11.23	Metabolism of xenobiotics
GSTT2	Glutathione S-transferase theta 2	chr22:q11.23	Metabolism of xenobiotics
NAT1	N-acetyltransferase 1 (arylamine N-acetyltransferase)	chr8:p22	Metabolism of xenobiotics
NAT2	N-acetyltransferase 2 (arylamine N-acetyltransferase)	chr8:p22	Metabolism of xenobiotics
NQO1	NAD(P)H menadione oxidoreductase 1,	chr16:q22.1	Metabolism of xenobiotics
NQO2	NAD(P)H dehydrogenase, quinone 2	chr6:p25.2	Metabolism of xenobiotics
NR1I2	Pregnane X receptor isoform 1l	chr3:q13.33	Metabolism of xenobiotics
TPMT	Thiopurine S-methyltransferase	chr6:p22.3	Metabolism of xenobiotics
TYMS	Thymidylate synthetase	chr18:p11.32	Metabolism of xenobiotics
EGF	Epidermal growth factor (beta-urogastrone)	chr4:q25	ROS as second messengers
FGF2	Fibroblast growth factor 2 (basic)	chr4:q27	ROS as second messengers
FIGF	c-fos induced growth factor (vascular endothelial growth factor D)	chrX:p22.2	ROS as second messengers
ICAM1	Intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	chr19:p13.2	ROS as second messengers
IGF1	Insulin-like growth factor 1 (somatomedin C)	chr12:q23.2	ROS as second messengers
IGF1R	Insulin-like growth factor 1 receptor precursor	chr15:q26.3	ROS as second messengers
IGF2	Insulin-like growth factor 2 (somatomedin A)	chr11:p15.5	ROS as second messengers
IGF2R	Insulin-like growth factor 2 receptor	chr6:q25.3	ROS as second messengers
IL10	Interleukin 10 precursor	chr1:q31-q32	ROS as second messengers
IL1A	Interleukin 1, alpha proprotein	chr2:q13	ROS as second messengers
IL1B	Interleukin 1, beta proprotein	chr2:q13	ROS as second messengers
IL6	Interleukin 6 (interferon, beta 2)	chr7:p15.3	ROS as second messengers
IL8	Interleukin 8 precursor	chr4:q13.3	ROS as second messengers
LTA	Lymphotoxin alpha (TNF superfamily, member 1)	chr6:p21.33	ROS as second messengers

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Table 1 (continued)

HUGO approved gene name	Full gene name	Chromosomal localization	Internal classification
LTB	Lymphotoxin-beta isoform a	chr6:p21.33	ROS as second messengers
PDGFA	Platelet-derived growth factor alpha polypeptide	chr7:p22	ROS as second messengers
PDGFB	Platelet-derived growth factor beta isoform 2,	chr22:q13.1	ROS as second messengers
PDGFC	Platelet-derived growth factor C precursor	chr4:q32.1	ROS as second messengers
PDGFRL	Platelet-derived growth factor receptor-like	chr8:p22	ROS as second messengers
RAF1	v-raf-1 murine leukaemia viral oncogene homologue 1	chr3:p25.2	ROS as second messengers
TANK	TRAF interacting protein TANK isoform b	chr2:q24.2	ROS as second messengers
TGFB1	Transforming growth factor, beta 1 (Camurati–Engelmann disease)	chr19:q13.2	ROS as second messengers
TGFB2	Transforming growth factor, beta 2	chr1:q41	ROS as second messengers
TGFB3	Transforming growth factor, beta 3	chr14:q24.3	ROS as second messengers
TGFBR1	Transforming growth factor, beta receptor I (activin A receptor type II-like kinase, 53 kDa)	chr9:q22.33	ROS as second messengers
TGFBR2	Transforming growth factor, beta receptor II (70/80 kDa)	chr3:p24.1	ROS as second messengers
TGFBR3	Transforming growth factor, beta receptor III (betaglycan, 300 kDa)	chr1:p22.1	ROS as second messengers
TNF	Tumour necrosis factor (TNF superfamily, member 2)	chr6:p21.33	ROS as second messengers
TNFAIP2	Tumour necrosis factor, alpha-induced protein 2	chr14:q32.32	ROS as second messengers
TNFAIP3	Tumour necrosis factor, alpha-induced protein 3	chr6:q23.3	ROS as second messengers
TNFAIP6	Tumour necrosis factor, alpha-induced protein 6	chr2:q23.3	ROS as second messengers
TNFRSF6	Tumour necrosis factor receptor superfamily, member 6	chr10:q23.31	ROS as second messengers
TRAF1	TNF receptor-associated factor 1	chr9:q33.2	ROS as second messengers
TRAF2	TNF receptor-associated factor 2	chr9:q34.3	ROS as second messengers
TRAF3	TNF receptor-associated factor 3 isoform 1	chr14:q32.32	ROS as second messengers
TRAF4	TNF receptor-associated factor 4 isoform 1	chr17:q11.2	ROS as second messengers
TRAF5	TNF receptor-associated factor 5	chr1:q32.3	ROS as second messengers
TRAF6	TNF receptor-associated factor 6	chr11:p12	ROS as second messengers
AKT1	v-akt murine thymoma viral oncogene homologue 1	chr14:q32.33	Signalling molecules regulated by the redox levels in the cells
AKT2	v-akt murine thymoma viral oncogene homologue 2	chr19:q13.2	Signalling molecules regulated by the redox levels in the cells
AKT3	v-akt murine thymoma viral oncogene homologue 3 (protein kinase B, gamma)	chr1:q44	Signalling molecules regulated by the redox levels in the cells
BCCIP	BRCA2 and CDKN1A-interacting protein isoform	chr10:q26.2	Signalling molecules regulated by the redox levels in the cells
CCND1	Cyclin D1 (PRAD1: parathyroid adenomatosis 1)	chr11:q13.3	Signalling molecules regulated by the redox levels in the cells
CDK2	Cyclin-dependent kinase 2 isoform 1	chr12:q13.2	Signalling molecules regulated by the redox levels in the cells
CDKN1A	Cyclin-dependent kinase inhibitor 1A (p21, Cip1)	chr6:p21.31	Signalling molecules regulated by the redox levels in the cells
EGFR	Epidermal growth factor receptor isoform c	chr7:p11.2	Signalling molecules regulated by the redox levels in the cells
FOS	v-fos FBJ murine osteosarcoma viral oncogene homologue	chr14:q24.3	Signalling molecules regulated by the redox levels in the cells
HRAS	v-Ha-ras Harvey rat sarcoma viral oncogene homologue	chr11:p15.5	Signalling molecules regulated by the redox levels in the cells
IER3	Immediate early response 3 isoform short	chr6:p21.33	Signalling molecules regulated by the redox levels in the cells
IL10RA	Interleukin 10 receptor, alpha precursor	chr11:q23.3	Signalling molecules regulated by the redox levels in the cells
IL10RB	Interleukin 10 receptor, beta precursor	chr21:q22.11	Signalling molecules regulated by the redox levels in the cells
IL1R1	Interleukin 1 receptor, type I precursor	chr2:q11.2	Signalling molecules regulated by the redox levels in the cells
IL1R2	Interleukin 1 receptor, type II precursor	chr2:q11.2	Signalling molecules regulated by the redox levels in the cells
IL6R	Interleukin 6 receptor isoform 1 precursor	chr1:q21.3	Signalling molecules regulated by the redox levels in the cells
IL6ST	Interleukin 6 signal transducer isoform 1	chr5:q11.2	Signalling molecules regulated by the redox levels in the cells
IL8RA	Interleukin 8 receptor, alpha	chr2:q35	Signalling molecules regulated by the redox levels in the cells
IL8RB	Interleukin 8 receptor, beta	chr2:q35	Signalling molecules regulated by the redox levels in the cells
IL8RBP	Interleukin 8 receptor, beta pseudogene	chr2:q35	Signalling molecules regulated by the redox levels in the cells
JUN	v-jun sarcoma virus 17 oncogene homologue (avian)	chr1:p32.1	Signalling molecules regulated by the redox levels in the cells
MAP3K14	Mitogen-activated protein kinase kinase kinase 14	chr17:q21.31	Signalling molecules regulated by the redox levels in the cells
MAPK1	Mitogen-activated protein kinase 1	chr22:q11.21	Signalling molecules regulated by the redox levels in the cells
MAPK7	Mitogen-activated protein kinase 7 isoform 2	chr17:p11.2	Signalling molecules regulated by the redox levels in the cells

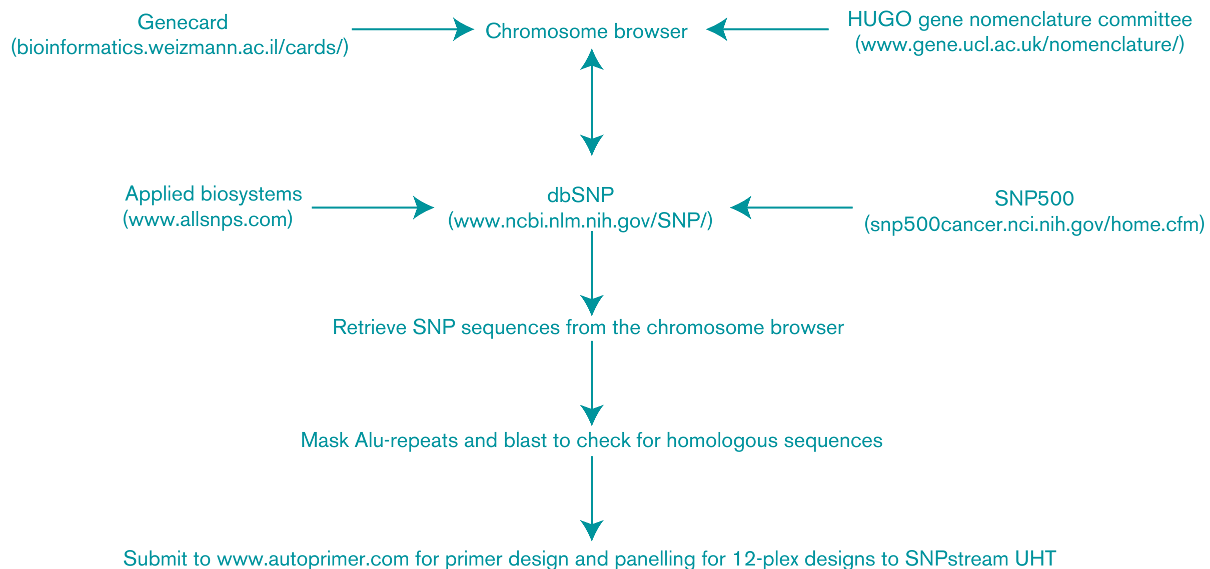
Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Table 1 (continued)

HUGO approved gene name	Full gene name	Chromosomal localization	Internal classification
MAPK8	Mitogen-activated protein kinase 8 isoform 2	chr10:q11.22	Signalling molecules regulated by the redox levels in the cells
MAPK9	Mitogen-activated protein kinase 9 isoform 1	chr5:q35.3	Signalling molecules regulated by the redox levels in the cells
MYB	v-myb myeloblastosis viral oncogene homologue (avian)	chr6:q23.3	Signalling molecules regulated by the redox levels in the cells
NFKB1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (p105)	chr4:q24	Signalling molecules regulated by the redox levels in the cells
NFKB2	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/ p100)	chr10:q24.32	Signalling molecules regulated by the redox levels in the cells
NFKBIA	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	chr14:q13.2	Signalling molecules regulated by the redox levels in the cells
NFKBIL2	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 2	chr8:q24.3	Signalling molecules regulated by the redox levels in the cells
PDGFRA	Platelet-derived growth factor receptor, alpha polypeptide	chr4:q12	Signalling molecules regulated by the redox levels in the cells
PDGFRB	Platelet-derived growth factor receptor, beta polypeptide	chr5:q32	Signalling molecules regulated by the redox levels in the cells
PDGFRL	Platelet-derived growth factor receptor-like	chr8:p22	Signalling molecules regulated by the redox levels in the cells
PIK3CA	Phosphoinositide-3-kinase, catalytic, alpha polypeptide	chr3:q26.32	Signalling molecules regulated by the redox levels in the cells
PLCG1	Phospholipase C gamma 1 isoform a	chr20:q12	Signalling molecules regulated by the redox levels in the cells
PLCG2	Phospholipase C, gamma 2 (phosphatidylinositol-specific)	chr16:q23.2	Signalling molecules regulated by the redox levels in the cells
PPM1A	Protein phosphatase 1A isoform 1	chr14:q23.1	Signalling molecules regulated by the redox levels in the cells
PPM1B	Protein phosphatase 1B isoform 1	chr2:p21	Signalling molecules regulated by the redox levels in the cells
PPM1G	Protein phosphatase 1G (formerly 2C), magnesium-dependent, gamma isoform	chr2:p23.3	Signalling molecules regulated by the redox levels in the cells
PPP1CA	Protein phosphatase 1, catalytic subunit, alpha isoform	chr11:q13.2	Signalling molecules regulated by the redox levels in the cells
PPP1R12B	Protein phosphatase 1, regulatory (inhibitor) subunit 12B	chr1:q32.1	Signalling molecules regulated by the redox levels in the cells
PPP1R14A	Protein phosphatase 1, regulatory (inhibitor) subunit 14A	chr19:q13.2	Signalling molecules regulated by the redox levels in the cells
PPP1R15A	Protein phosphatase 1, regulatory (inhibitor) subunit 15A	chr19:q13.33	Signalling molecules regulated by the redox levels in the cells
PPP1R1A	Protein phosphatase 1, regulatory (inhibitor) subunit 1A	chr12:q13.2	Signalling molecules regulated by the redox levels in the cells
PPP1R2	Protein phosphatase 1, regulatory (inhibitor) subunit 2	chr3:q29	Signalling molecules regulated by the redox levels in the cells
PPP1R3B	Protein phosphatase 1, regulatory (inhibitor) subunit 3B	chr8:p23.1	Signalling molecules regulated by the redox levels in the cells
PPP1R3C	Protein phosphatase 1, regulatory (inhibitor) subunit 3C	chr10:q23.32	Signalling molecules regulated by the redox levels in the cells
PPP1R7	Protein phosphatase 1, regulatory subunit 7	chr2:q37.3	Signalling molecules regulated by the redox levels in the cells
PPP1R9A	Protein phosphatase 1, regulatory (inhibitor) subunit 9A	chr7:q21.11-q31.33	Signalling molecules regulated by the redox levels in the cells
PPP2R4	Protein phosphatase 2A, regulatory subunit B' (PR 53)	chr9:q34.11	Signalling molecules regulated by the redox levels in the cells
PPP3CA	Protein phosphatase 3 (formerly 2B), catalytic subunit, alpha isoform (calcineurin A alpha)	chr4:q24	Signalling molecules regulated by the redox levels in the cells
PRKCA	Protein kinase C, alpha	chr17:q24.2	Signalling molecules regulated by the redox levels in the cells
PRKCABP	Protein kinase C, alpha binding protein	chr22:q13.1	Signalling molecules regulated by the redox levels in the cells
TP53	Tumour protein p53 (Li–Fraumeni syndrome)	chr17:p13.1	Signalling molecules regulated by the redox levels in the cells
ARNT	Aryl hydrocarbon receptor nuclear translocator	chr1:q21.2	Other
CALM1	Calmodulin 1 (phosphorylase kinase, delta)	chr14:q32.11	Other
CALM2	Calmodulin 2 (phosphorylase kinase, delta)	chr2:p21	Other
CALM3	Calmodulin 3 (phosphorylase kinase, delta)	chr19:q13.32	Other
HMOX1	Haeme oxygenase (decycling) 1	chr22:q12.3	Other
HMOX2	Haeme oxygenase (decycling) 2	chr16:p13.3	Other
KCNMB1	Potassium large conductance calcium-activated channel, subfamily M, beta member 1	chr5:q35.1	Other
MARK3	MAP/microtubule affinity-regulating kinase 3	chr14:q32.32	Other
MRPL49	Mitochondrial ribosomal protein L49	chr11:q13.1	Other

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Fig. 2



A flow-chart of the up-front single nucleotide polymorphism-mining and confirmation process.

City, California, USA) (www.allsnps.com, Celera), dbSNP (www.ncbi.nlm.nih.gov/SNP) and the SNP500Cancer project (http://snp500cancer.nci.nih.gov/home_1.cfm). At a later stage, the SNPper database (http://snpper.chip.org/bio/) and the HapMap project (http://www.hapmap.org/index.html.en) were also mined for frequency information. A summary of the basic characteristics of these resources for SNP mining is provided in Table 2.

Results and discussion

Of the 1030 SNPs analysed, 305 assays failed. A full list of the remaining 725 genotyped SNPs is provided in Supplementary Table 1. Also listed are the frequencies of the variant alleles in the Norwegian population and in different databases and database builds, as well as the call rates and the Hardy–Weinberg equillibrium. The PCR primers and mini-sequencing primers used for the 725 assays are listed in Supplementary Table 2. A failure rate as high as approximately 30%, could reflect the fact that this was a pilot experiment with no optimization and that users had no prior experience with the GenomeLab SNPstream genotyping system. The relatively high degree of multiplexing with the expected drop-outs is also likely to have had an influence. In addition, some of the primers designed by autoprimer.com were 'flagged' by that program as having the potential to give rise to PCR primer cross-hybridization or PCR-primer versus amplicon cross-hybridization. A higher percentage of SNPs generated by these susceptible primers were found in the 305 failed assays (33%) compared to the successful ones (23%). Of the remaining 725 assays, 15 had a lower than 75% call rate and were therefore rejected, leaving 710 successfully genotyped SNPs, of which 36 SNPs were not in Hardy–Weinberg equilibrium and 87 were monomorphic; an additional 40 SNPs occurred at a frequency of less than 1% in the 193 patients studied. The 710 SNPs were located in 206 different genes with 1–20 SNPs per gene and a median of nine genes per chromosome. Figure 3 shows the chromosomal distribution of the 206 genes. The successfully genotyped SNPs are listed in Supplementary Table 3, together with their chromosomal localization, gene affiliation and functional class from the different databases and builds.

The majority of SNPs were recruited from the dbSNP and the Applied Biosystems (Celera) databases. A comparatively large number of SNPs (relative to the actual size of the database) was derived from the SNP500 database, which proved to be of a high quality in terms of both annotation and frequency information. Figure 4(a) shows the number of SNPs with frequency information in the different databases prior to genotyping (Fig. 4a, left) and in current builds (Fig. 4a, right). The majority of the successfully genotyped SNPs in dbSNP build 114 were intronic (Fig. 4b, left) ($n = 349$). This was a result of prioritising SNPs with a known and reliable frequency. Notwithstanding this, 77 SNPs were in exons, another 137 in regulatory areas and two in splice sites. The functional localization of SNPs in the dbSNP build 124 is shown in Fig. 4(b, right). An update of database information for the SNPs studied, following genotyping in the originally mined databases, gives rise to broad material for crosschecking the quality and information

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Table 2 Summary of conditions, availability and options for different single nucleotide polymorphism resources

Resource	Availability	Batch search options	Used in this project	Comments
SNPer (http://snpper.chip.org/bio/)	Free for academic users	Yes	Yes	The SNPper database is based on information from dbSNP and the UCSC Human Genome Browser. Gives minimum–maximum frequency
DbSNP (http://www.ncbi.nlm.nih.gov/SNP/)	Public	Yes	Yes	A public repository for SNPs, microsatellite repeats and small insertions/deletions accepting submissions from all sources, public and private. Gives average allele frequencies
ABI/Celera (http://myscience.appliedbiosystems.com/common/search.jsp?assayType=genotyping)	Proprietary	Yes	Yes	Gives frequency information from Celera in different ethnical groups, including Caucasians. Links to dbSNP/JSNP
Chromosome Browser	Proprietary	Yes	Yes	A software application allowing the user to search the human genome for SNPs and SNP characteristics developed by Orchid Biosciences. Classification based on dbSNP, JSNP and HGBase
SNP500 (http://snp500cancer.nci.nih.gov/)	Public	No	Yes	Information from SNP500 is deposited in dbSNP. The frequency data are retrieved from a panel of DNA's from 102 lymphoblastoid cell lines from four ethnically diverse groups, including Caucasians
Ensembl (http://www.ensembl.org/index.html)	Public	Yes	No	Co-operation between Sanger Institute and EBI. Gives frequency data from different sources, included Perlegen, SNP500 and Sequenom
The UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway)	Public	Yes	No	Based on the NCBI Genome Build
The SNP consortium (http://snp.cshl.org/)	Public	Yes	No	Participant in the HapMap project; a non-profit foundation organized for the purpose of providing public genomic data. Provides frequencies in three populations, including Caucasians
The HapMap project (http://www.hapmap.org/index.html.en)	Public	No, but bulk download possible	Yes	Catalogue of genetic variation in populations of African, Asian and European ancestry giving information on frequency, linkage disequilibrium properties, genotypes and SNP assay

All the databases are searchable by reference cluster id (rs#) and are available online. In December 2002, Orchid sold part of its Life Sciences business to Beckman Coulter. At that time, Orchid stopped supporting Chromosome Browser, which was an activity undertaken as part of its Life Sciences business. Beckman Coulter are currently working on developing and updating Chromosome Browser for a release next year (personal communication: Dr Zhiming Jiang, Product Manager Strategic Marketing, Beckman Coulter).

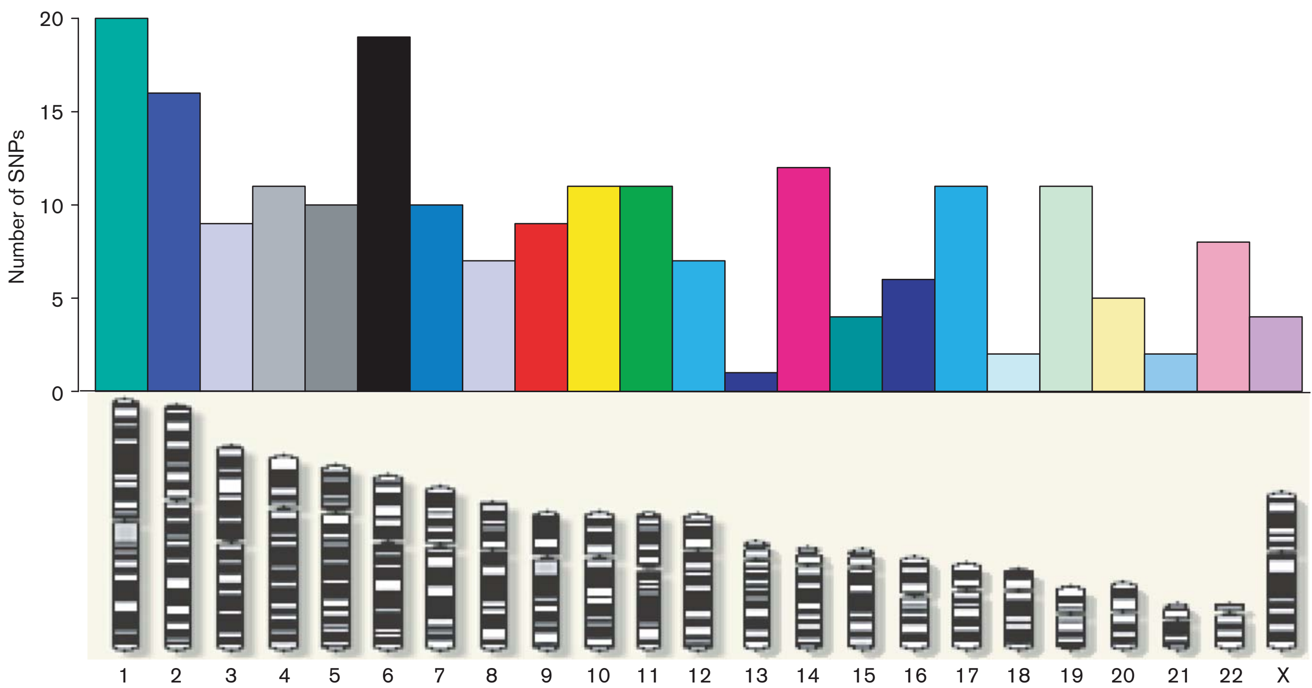
level of the different freezes. Comparing the number of SNPs with available frequency information in the different databases for the different builds for our 725 SNPs, revealed an increase in their information content (from 30% to 70%). The largest increase was in dbSNP (70%); but, as the population-specific frequency information is not given in this database, data should be verified in other databases. Frequency information was also extracted from the SNPper database and the HapMap project. Table 3 shows the number of SNPs with pair-wise frequency information in the databases (current builds) for our SNP set. The number of SNPs with unknown function decreased markedly, falling from 139 SNPs labelled as 'unclassified' SNPs or with 'sequences not matched' in dbSNP build 114 (Fig. 4b, left) to 57 SNPs with 'no data available' or 'no link established' in dbSNP build 124 (Fig. 4b, right).

It has recently been suggested that SNPs in the dbSNP database reported independently by at least two groups, and thereby listed as validated, are more likely to be confirmed as true variants [10,11]. dbSNP build 124 currently contains more than 10 million SNPs, approximately half of which have been validated. In our set of 1030 SNPs, we compared the genotyping success rate of validated and non-validated SNPs in dbSNP build 114 and 124. We found that validated SNPs were more often successfully genotyped than non-validated SNPs ($P = 0.003$ for both builds) (Table 4). We also compared the success rate of those SNPs with allele frequency information in at least two databases ('two hits') with those which had allele frequency information in either only one database ('one hit') or in none of the databases ('no hits') (Table 4), prior to genotyping. SNPs with one or two 'hits' were more likely to be successfully genotyped than SNPs with no frequency information ($P < 0.001$). These findings indicate that the presence of frequency information for a SNP in a Caucasian population can be taken as strong evidence that the same SNP will be found in the Norwegian population. This is in agreement with the reported bias towards populations of European descent as a source of the SNP information found in dbSNP [10].

Between June 2003 and March 2005, 17 SNPs (in a total of six genes including *CALM3* and *NQO1*) changed gene position in all the databases investigated; another 23 SNPs (in 17 genes including *PLCG2* and *TNFRSF6*) changed gene association in at least one of the databases. A search was conducted for these SNPs using the UCSC

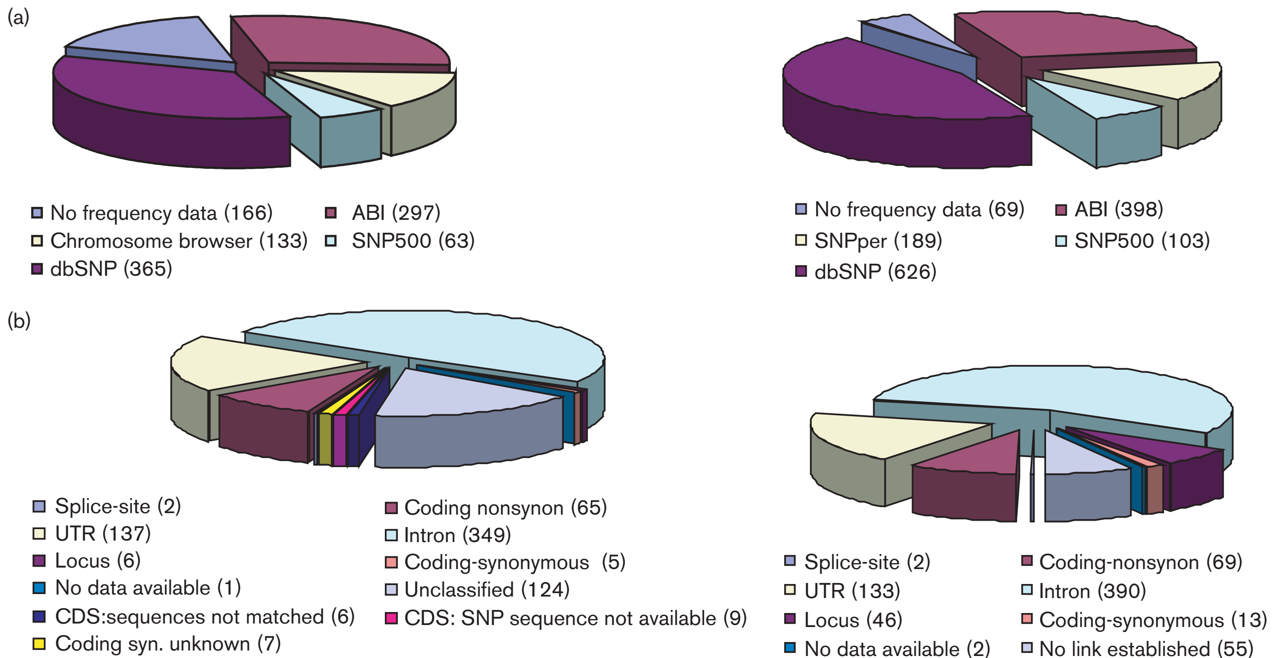
Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Fig. 3



Chromosomal distribution of the 206 genes with successfully genotyped single nucleotide polymorphisms (SNPs) (The chromosome picture is from http://www.ensembl.org/Homo_sapiens/, slightly modified).

Fig. 4



(a) Number of single nucleotide polymorphisms (SNPs) with frequency information according to database localization, prior (left panel) and after (right panel) genotyping. (b) Distribution of successfully genotyped SNPs according to functional classification in dbSNP build 114 (left panel) and dbSNP build 124 (right panel).

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Table 3 Overlap between the different databases mined for single nucleotide polymorphisms (SNPs) in this study with respect to the number of annotated SNPs with validated frequencies in a Caucasian population, in the case of SNPs selected based on the candidate gene approach

	dbSNP (build 124)	ABI (Human Genome build 35)	SNP500 (Human Genome build 35)	SNPper (dbSNP build 123, Goldenpath: hg17)	HapMap (CEPH) (Human Genome build 34)
dbSNP (build 124)	623	367	100	188	190
ABI (Human Genome build 35)	367	397	41	115	241
SNP500 (Human Genome build 35)	100	41	103	11	43
SNPper (dbSNP build 123, Goldenpath: hg17)	188	115	11	188	111
HapMap (CEPH) (Human Genome build 34)	190	241	43	111	334

Table 4 Evaluation of successfully and unsuccessfully genotyped single nucleotide polymorphisms (SNPs) with respect to validation status in dbSNP builds 114 and 124 (left panel), and available frequency information (right panel) with calculated *P*-values

	dbSNP				All databases, June 2003 (SNPs with frequency information)		
	Build 114		Build 124				
	Validated	Non-validated	Validated	Non-validated	2 hits	1 hit	0 hits
Good calls ($n=674$)	400 (59.3)	274 (40.7)	606 (89.9)	68 (10.1)	237 (35.2)	280 (41.5)	157 (23.3)
Bad calls ($n=356$)	178 (50.0)	178 (50.0)	296 (83.1)	60 (16.9)	84 (23.6)	126 (35.4)	146 (41.0)
P-value	$P=0.003$		$P=0.003$		$P<0.0001$		

Frequency information found in two or more of the databases mined is referred to as 'two hits', one database as 'one hit' and no frequency information available in any of the databases mined is referred to as 'no hits'. 'Good calls' are defined as SNPs successfully genotyped and in Hardy–Weinberg equilibrium. 'Bad calls' are defined as SNPs not successfully genotyped or not in Hardy–Weinberg equilibrium.

Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway) to verify their position; a list of the SNPs together with their gene association in the UCSC Genome Browser is provided in Supplementary Table 4.

The HapMap database lists only 442 of the 1030 SNPs genotyped in this project. Of these 442 SNPs, 108 were among the 305 SNPs that failed to be genotyped. The remaining 334 SNPs are listed in Supplementary Table 1, together with frequency and minor allele information for the the Centre d'Etude du Polymorphisme Humain (CEPH) family population. Of these 334 SNPs, a total of 23 were either not in Hardy–Weinberg equilibrium or had a failure rate higher than 25% in our genotyping project. Comparing information from the HapMap database with the results of our genotyping, we found 18 SNPs where the allele described as minor in the HapMap project was a major allele in the Norwegian population. Another 10 SNPs, which were monomorphic in the CEPH family, were found with a frequency of between 0.003 and 0.011 in the Norwegian population. Altogether, the difference in frequency between the two populations was between 0.5% and 42%.

This study indicates that reliance on only one database may prove to be expensive with respect to reagents and time spent on assays involving erroneously annotated SNPs/genes. However, constant improvements are being made in the Human Genome Reference DNA Sequence. This will result in increasingly accurate information and fewer discrepancies between the different databases. The availability of this more accurate information, combined with the data obtained from large genotyping projects (such as the large-scale genotyping of approximately 1.5 million SNPs in three different ethnic populations [12]) is likely to make the selection of SNPs for pharmacogenetic studies an easier task in the future. The results of our study indicate that the success rate of genotyping projects for a set number of selected SNPs will increase: (i) with the number of databases providing frequency information for the given SNPs or (ii) if the SNPs are validated in dbSNP. In the present study, we report a set of validated SNPs in a Caucasian population that constitutes a set of candidate SNPs for pharmacogenetics; these are presented along with their gene localization, chromosome localization, functional classification and frequency in a cohort of Norwegian women with breast cancer. These data will be useful for further pharmacogenetic studies in Caucasian populations.

Acknowledgement

The authors would like to thank Johanna C. Mylet, MS, from Beckman Coulter, for the assistance she provided during the genotyping process.

References

1 Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet* 2003; **33 (Suppl)**:228–237.
2 Wjst M. Target SNP selection in complex disease association studies. *BMC Bioinformatics* 2004; **5**:92.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

3 Gemignani F, Landi S, Vivant F, Zienolddiny S, Brennan P, Canzian F. A catalogue of polymorphisms related to xenobiotic metabolism and cancer susceptibility. *Pharmacogenetics* 2002; **12**:459–463.
4 Bell PA, Chaturvedi S, Gelfand CA, Huang CY, Kochersperger M, Kopla R, *et al.* SNPstream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery. *Biotechniques* 2002; **32 (Suppl 70–2, 74, 76–7)**:70–77.
5 Weston CR, Davis RJ. The JNK signal transduction pathway. *Curr Opin Genet Dev* 2002; **12**:14–21.
6 Davis W Jr, Ronai Z, Tew KD. Cellular thiols and reactive oxygen species in drug-induced apoptosis. *J Pharmacol Exp Ther* 2001; **296**:1–6.
7 Norbury CJ, Hickson ID. Cellular responses to DNA damage. *Annu Rev Pharmacol Toxicol* 2001; **41**:367–401.
8 Kamata H, Hirata H. Redox regulation of cellular signalling. *Cell Signal* 1999; **11**:1–14.
9 Bravard A, Luccioni C, Moustacchi E, Rigaud O. Contribution of antioxidant enzymes to the adaptive response to ionizing radiation of human lymphoblasts. *Int J Radiat Biol* 1999; **75**:639–645.
10 Carlson CS, Eberle MA, Rieder MJ, Smith JD, Kruglyak L, Nickerson DA. Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. *Nat Genet* 2003; **33**: 518–521.
11 Reich DE, Gabriel SB, Altshuler D. Quality and completeness of SNP databases. *Nat Genet* 2003; **33**:457–458.
12 Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, *et al.* Whole-genome patterns of common DNA variation in three human populations. *Science* 2005; **307**:1072–1079.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.