Genetic variation of the ABC transporter gene *ABCC1* (Multidrug resistance protein 1 – MRP1) in the Polish population

Marcin Słomka¹, Marta Sobalska-Kwapis¹, Małgorzata Korycka-Machała², Grzegorz Bartosz², Jarosław Dziadek² and Dominik Strapagiel*¹

**Abstract**

**Background:** Multidrug resistance-associated protein 1 (MRP1), encoded by the ABCC1 gene, is an ATP-binding cassette transporter mediating efflux of organic anions and xenobiotics; its overexpression leads to multidrug resistance. In this study, 30 exons (from 31 in total) of the ABCC1 gene as well as and their flanking intron sequences were screened for genetic variation, using the High Resolution Melting (HRM) method, for 190 healthy volunteers representing the Polish population. Polymorphism screening is an indispensable step in personalized patient therapy. An additional targeted SNP verification study for ten variants was performed to verify sensitivity of the scanning method.

**Results:** During scanning, 46 polymorphisms, including seven novel ones, were found: one in 3’UTR, 21 in exons (11 of them non-synonymous) and 24 in introns, including one deletion variant. These results revealed some ethnic differences in frequency of several polymorphisms when compared to literature data for other populations. Based on linkage disequilibrium analysis, 4 haplotype blocks were determined for 9 detected polymorphisms and 12 haplotypes were defined. To capture the common haplotypes, haplotype-tagging single nucleotide polymorphisms were identified.

**Conclusions:** Targeted genotyping results correlated well with scanning results; thus, HRM is a suitable method to study genetic variation in this model. HRM is an efficient and sensitive method for scanning and genotyping polymorphic variants. Ethnic differences were found for frequency of some variants in the Polish population compared to others. Thus, this study may be useful for pharmacogenetics of drugs affected by MRP1-mediated efflux.

**Keywords:** High resolution melting, ABCC1, MRP, Gene scanning, Genotyping

**Background**

The human multidrug resistance-associated protein 1 (MRP1) is a member of the ATP-binding cassette (ABC) transporter superfamily and is encoded by the *ABCC1* gene [1]. MRP1 was first described by Cole et al. [2] who cloned the overexpressed transporter from lung cancer cell line H69AR. The gene is mapped to chromosome 16p13.1. Its length is around 200,000 base pairs; the gene contains 31 exons encoding a protein of 1,531 amino acids, with molecular weight of 190 kDa [3, 4]. The protein has 3 hydrophobic transmembrane domains (TMDs), also called membrane-spanning domains (MSDs), named TMD0 (N-terminal), TMD1 (middle) and TMD2 (C-terminal). They contain 17 transmembrane helices (TM), 5 in TMD0 and 6 in both TMD1 and TMD2. The intracellular region of the protein contains 2 hydrophilic nucleotide binding domains (NBDs), NBD1 associated with TMD1 and NBD2 associated with TMD2 [5, 6]. They are both involved in binding and hydrolyzing ATR, which is indispensable for substrate transport. Each of NBDs contains highly conserved motifs: Walker A motif which binds the β-phosphate of ATP and Walker B motif that interacts with Mg²⁺ ions. There is also a third, 13 amino acid sequence motif, between Walker A and Walker B called the ABC signature [1, 7].

MRP1 is ubiquitously expressed in many human tissues and physiological barriers like blood–brain or blood-testis...
barriers. High levels of MRP1 expression were detected in lung, kidney, heart, testis, placenta, adrenal glands and skeletal muscles, while lower levels of expression were found in intestine, colon, brain, peripheral blood mononuclear cells and liver [1, 3, 8]. With the exception of brain cells, MRP1, in contrast to other ABC transporters, is expressed at the basolateral membrane of polarized epithelial cells where it plays a protective role against xenobiotics and their metabolites [3, 9].

Many drugs are good substrates for MRP1, so its overexpression leads to multidrug resistance (MDR), especially during cancer chemotherapy. This effect can be observed in many types of cancer cells including solid tumors (lung cancer, breast cancer, gastric and colon carcinomas, melanoma, prostate cancer, neuroblastoma) as well as in various types of leukemias [3, 9]. MRP1-overexpressing cells are capable of transporting a large group of substrates, including anticancer drugs like folate-based antimetabolites, anthracyclines, Vinca alkaloids, antiandrogens and even some inorganic anions such as arsenite and arsenate. MRP1 is also an active transporter of a broad range of endogenous compounds including glutathione (GSH), glucuronate and sulfate-conjugates of organic anions, such as inflammatory mediator leukotriene C₄ or estradiol-17-β-D-glucuronide (E₂17βG). Therefore, MRP1 plays a crucial role in efflux of organic anions, anionic drugs, xenobiotic conjugates and additionally in the maintenance of redox balance by participation in cellular response to oxidative stress [3, 7, 9–11].

A large number of single nucleotide polymorphisms (SNPs) in ABCC1 gene were detected so far in studies of various populations, especially Asian and Caucasian. Saito et al. [12] reported 779 genetic variations in eight ABC genes in Japanese population, 95 among them concerning ABCC1. Subsequently, Fukushima-Uesaka et al. [13], screening ABCC1 in the same population, identified 86 genetic variations including 31 novel ones. Similar screening of Chinese population revealed 32 SNPs [14]. Leschziner et al. [15] screened five ABC transporter genes in Caucasian individuals and detected 221 variants, including 61 for ABCC1, among them 22 novel for this gene. They identified also large blocks of high linkage disequilibrium and low haplotype diversity across all the genes in the examined population. Comparative study including polymorphisms from four different populations: Chinese, Malay, Indian and Caucasian showed that apart from population-specific SNPs, only 6 of 71 detected SNPs altered amino acid sequence, so ABCC1 is considered to be a highly conserved gene [16]. Majority of previously reported genetic variations were found in intron sequences. However, there is evidence in literature that some SNPs change protein functioning and effectiveness of cancer chemotherapy. Such cases were reported for neuroblastoma [17], breast cancer [18], ovarian cancer [19] and other serious diseases like: chronic obstructive pulmonary disease (COPD) [20, 21] cystic fibrosis [22] and major depression [23].

Previously mentioned studies have shown importance of ABCC1 polymorphisms and its genetic variation. These data can have great relevance for pharmacogenetics in personalized patient therapy. Based on the data described above for several populations, there exist ethnic differences in the frequency of ABCC1 polymorphic variants. Some of them are so strongly linked to each other that they are always inherited together, forming haplotypes. Haplotype is a combination of polymorphic variants inherited on the same chromosome and sometimes has stronger clinical relevance for drug response or adverse events than a single polymorphism [24]. Moreover, there is a large haplotype diversity across different populations [25]. Therefore, prior comprehensive analysis of the whole gene is required in the specific population before application of genetic variation data in pharmacotherapy.

Our previous studies performed during the project TESTOPLEK in which we have examined genetic variation of multidrug transporter genes in the Polish population indicated that ABCC1 polymorphisms were present at significant level in a sample of Polish individuals. However, various SNPs in the ABCC1 gene differed significantly in their frequency in tested DNA samples [26]. This prompted us to verify the whole coding sequence of ABCC1 gene. We conducted high throughput analysis and screened 30 exons (from 31 in total) of ABCC1 gene for polymorphism variants in the Polish population (which is exclusively Caucasians). This is the first study of genetic variation of ABCC1 gene in the Polish population.

Methods
Materials and genomic DNA samples

Human genomic DNA samples were derived from anonymous Polish unrelated volunteers who declared to be healthy. Samples were randomly selected from the “normal Polish population” genetic collection at the Biobank Lab, Department of Molecular Biophysics, University of Lodz. Genetic material for this collection was sampled in 2011 – 2012 within the EU-funded TESTOPLEK project. All subjects gave their written informed consent to participate in the study. This study was approved by the relevant regional ethical committee (Research Ethics Commission, University of Lodz - Decision no. 8/KBBN-UL/II/2014 and Statement of the Research Ethics Commission, University of Lodz from 17th of June, 2010) and all procedures were performed in accordance with the Declaration of Helsinki.

Saliva was collected into Oragene OG-500 DNA collection/storage receptacles (DNA Genotek, Kanata, Canada) and genomic DNA was subsequently isolated by the
Screening of ABCC1 by High Resolution Melting (HRM) Method

Investigation of ABCC1 genetic variation was conducted using High Resolution Melting method. The method is based on precise measurements of DNA melting profile. The dsDNA melting temperature depends on length and nucleotide composition of the PCR product. Even a single base variation generates a different melting curve, and many variants can be detected in this way. The list of primers used for HRM scanning of all the areas is presented in Table 1. The single reaction mixture (10 μl) was prepared using Janus® Automated Workstation (Perkin Elmer Inc., Waltham, USA) and composed of GoTaq® Colorless Master Mix (2×) (Promega, Madison, USA), LC Green Plus® dye (10x) (BioFire Defense Inc., Salt Lake City, USA), 0.5 μl of 10 μM primers mixture, 3 μl DNA, and filled up to the final volume with water. Reaction was performed on 384-micro well plates using CFX384® real-time PCR system (Bio-Rad Laboratories Inc., Hercules, USA) on duplicate samples. The reaction conditions were as follows: initial denaturation at 95 °C for 3 min, 50 amplification cycles of denaturation at 95 °C for 30 s and annealing at specific temperature depending on the primers used, for 30 s. The plate was read after each cycle. Directly afterwards, melting curve was determined, the plate being incubated at 90 °C for 60 s, 40 °C for 60 s and from 65 °C to 95 °C with an increment of 0.2 °C for 10 s with plate reading. The obtained data were analysed with the Bio-Rad Precision Melt Analysis Software, version 1.2 (Bio-Rad Laboratories Inc., Hercules, USA). Based on HRM melting curve analysis, confirmation of genetic variation for every melting cluster was obtained by direct sequencing method for several samples selected from each cluster (or all samples from clusters less than 3 samples). We used the data obtained from genotyping to compare them with HRM scanning results (comparison of minor allel frequencies - MAFs) and thereby to validate the accuracy of the screening method.

DNA sequencing

For sequencing, selected samples were prepared by PCR. Reaction mixture filled with water to a final volume of 50 μl per sample was composed using GoTaq® Green Master Mix (2x) (Promega, Madison, USA), 5 μl of 10 μM primers mix and 5 μl DNA. Reaction conditions were as follows: an initial denaturation at 95 °C for 3 min, 45 amplification cycles of denaturation at 95 °C for 30 s, annealing at specific temperature depending on the primers used for 30 s and extension at 72 °C for 45 s, followed by final extension at 72 °C for 5 min (list of primers for exon sequencing is shown in Table 1). Afterwards, PCR products were purified using NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany). DNA concentration and size of PCR product was determined by agarose electrophoresis of 2 μl of each sample. Based on the intensity of bands the samples were diluted from 5 to 50 times and applied for sequencing reaction using BigDye Terminator V3.1 (Applied Biosystems, USA) according to the manufacturer’s protocol. The PCR-Sequencing product was purified using BigDye X-Terminator kit following the manufacturer’s protocol (Applied Biosystems, USA). Further, 30 μl of each purified sample was applied to 96 wells of titration plates and analyzed in 3500 Genetic Analyzer (Applied Biosystems, USA).

Polymorphism detection

The ABCC1 genomic DNA sequence (NG_028268.1) was obtained from GenBank (http://www.ncbi.nlm.nih.gov) and used as a reference sequence during analysis of sequencing results by CodonCode Aligner software (CodonCode Corporation, Centerville, USA). Sequencing results of selected samples were compared with respective HRM clusters. Based on this, genotype for each melting cluster was established and genetic variation for each sample was verified.

Availability of supporting data

For each of the polymorphisms detected in this study in ABCC1 gene, the following parameters were assigned: dbSNP IDs (rs numbers) (http://www.ncbi.nlm.nih.gov/SNP/), nucleotide position within or relative to the coding sequence based on the NM_004996.3, and amino acid position in the protein for SNPs in the coding sequence.
| Exon | Number of scanned bp upstream from the exon (without primer) | Forward primer (5' → 3') | Reverse primer (5' → 3') | Number of scanned bp downstream from the exon (without primer) | Size (bp) including primers |
|------|------------------------------------------------|--------------------------|--------------------------|------------------------------------------------|--------------------------|
| 1    | 104*                                            | CACGCGTACGCGCGACAGG      | CCCGGTGCCTTTCTCCCTC      | 49                                             | 234                      |
| 2    | 31                                              | GTCCCTGCGGGGTGTGCTCT    | TCTGAAATGAGCTCGGTCAGT    | 201                                            |                          |
| 3    | 20                                              | GCTTTCTACACGCTCTTCTGCA  | TTACACACCCACCTACCTCC     | 52                                             | 203                      |
| 4    | 129                                             | AAGGCTGGGAGGAAATCTCA     | AAGGTAGCAGACAGCTGAGC     | 163                                            |                          |
| 5    | 27                                              | CAGCCGCGAATGTGACAGG      | CACACAGCAGACACGACT       | 19                                             | 212                      |
| 6    | 20                                              | TTTCCTCTTCTCCCAACAC     | AGCTGAGCATGTTCATCTGGT    | 59                                             | 182                      |
| 7    | 33                                              | CTCCTCTTCTGCACAGCT      | GAGGCGGACTAGGCTGACAG     | 33                                             | 175                      |
| 8    | 74                                              | TTTCCTGACAGGCTCTCATC    | GCCACAGAGGCTAGGCTAGC     | 104                                            | 184                      |
| 9    | 14                                              | GACCGCTTCGTGCTGGG       | CTGCCCACAGCGAAGTCCG     | 38                                             | 157                      |
| 10   | 27                                              | CACACCGCAGGCTCTCATC     | GCCACAGAGGCTAGGCTAGC     | 29                                             | 199                      |
| 11   | 21                                              | TTTCCTCTTCTCAGGAGAAGY   | TGGCCTATACAATCTCATAGGAGAG | 49                                           | 209                      |
| 12   | 41                                              | GTTCACTCTGAGGCTCTACT    | GCCAGACTCTCTGACACT       | 217                                            |                          |
| 13   | 62                                              | GTTCCTCTGAGGCTCTACT     | GCCGAGACTCTCTGACACT      | 75                                             | 206                      |
| 14   | 49                                              | TTACCGCTGGGTTTCTGTTG    | GCCAGACTCTCTGACACT       | 51                                             | 228                      |
| 15   | 41                                              | GTTCTGTGATGTGGAGGTGCG   | GGCTCTCTTACACCATACAC     | 22                                             | 179                      |
| 16   | 73                                              | TTTCCTACTGAGGTTAATGGGAGT | CTGAGACGGAGGAGACTTTACT  | 177                                            |                          |
| 17   | 12                                              | ATCCCGAGGAGGTGCTTTG     | GCCAGCTCTCGACACCCAGCAG   | 37                                             | 171                      |
| 18   | 20                                              | ATCTGGAGGAGGAGAGTGGAG   | CAGCCACCTGAGATGGTACAG    | 60                                             | 134                      |
| 19   | 26                                              | CTCACAGCATGTGCTCTCCAG   | TCTGTGCTGGCATAGGTACG     | 55                                             | 166                      |
| Table 1 | Pairs of primers used for HRM screening, sequencing and genotyping ABCC1 (Continued) |
|---------|-----------------------------------------------------------------------------------|
| 20      | GCCAAGCTAGGCAGTCTCAC GGCAAGTAGCTCATGCTGTG | 99 |
| 21      | TCTGAGATGGGCTCTCTACCA CAACTCAAGAGGCAATG | 56 |
| 22      | GGTGCGTGATGTGCTAAAG ATGGGGTCATCGTCCAGAG | 184 |
| 23      | GCCATCTGCATTGTGGAGT GTCCACTGTGCTCCAGTCT | 214 |
| 24      | ATAGGCTCTTGGAGCGGAC GGCCTTGGAGCTGCAAGTAAAG | 44 |
| 25      | GCCAAGCTAGGCAGTCTCAC GGCAAGTAGCTCATGCTGTG | 99 |
| 26      | CACCTTACCAGACGTGGGACTCTGCAAGTAAAG | 44 |
| 27      | GCCATCTGCATTGTGGAGT GTCCACTGTGCTCCAGTCT | 214 |
| 28      | ATAGGCTCTTGGAGCGGAC GGCCTTGGAGCTGCAAGTAAAG | 44 |
| 29      | GCCAAGCTAGGCAGTCTCAC GGCAAGTAGCTCATGCTGTG | 99 |
| 30      | CACCTTACCAGACGTGGGACTCTGCAAGTAAAG | 44 |
| 31      | GCCAAGCTAGGCAGTCTCAC GGCAAGTAGCTCATGCTGTG | 99 |
|         | Sequencing                                                                        |
| 1       | -                                                                                |
| 2       | TCTCGAACTCCCAGCTAAA TAAGCGGCAGAGCAAAGATT | 749 |
| 3       | GACGATGGTGATTAGGACACAA TCCCTGACATCGTCTGAC | 638 |
| 4       | GTGGTGAAACCCCAAGCTTTAA CCTTGGAAGCAAACACAGCAGA | 604 |
| 5       | CCCAGTAGAGCTGGCCGTTA ACCATGCCAAAGTGGAAAC | 606 |
| 6       | GGGTGGTGATGGGAGACTGAA AAGGACAAACAGTGCTGGGAT | 622 |
| 7       | CTGCACTGACCAAGACAGAATG CAAACTCCCACAAACAGT | 651 |
| 8       | TTTTCCACTTCTGGCTCTTG TAGGGAGCTAGTGCTTTT | 691 |
| 9       | AGACATTAGTGCAAACCCCA CAACTCCCACAAACAGT | 613 |
| 10      | ACGCATCGAAGCGTGAA AATCTCCAGAGCAGCTCTACA | 664 |
| 11      | AGGCGTGAGACCTGTTATTC CACCAGCTACAAGGTCTAA | 608 |
| 12      | GATGTTGAGGTAGGGCCGCTGA CTTGGGCACATGTGACCT | 363 |
| 13      | ATACTGCCAGGTTTTTTC CAAAAGGATAGGAATGCGG | 611 |
Table 1 Pairs of primers used for HRM screening, sequencing and genotyping ABCC1 (Continued)

|   | Forward primer (5’→3’) | Reverse primer (5’→3’) | Size (bp) including primers |
|---|------------------------|------------------------|-----------------------------|
| 14 | CTGGGGAAACCTTTGAAAGTT | CCAAGGGAAAGAATGCAAG    | 263                         |
| 15 | CCGAAGTGTGTCGTCGAGAT  | TACGCAAGCCAAGAATCTCC   | 632                         |
| 16 | TCTCCCTTCAGACCTCGGCT  | CCTGCTCTTCTAGGACAGCAC  | 676                         |
| 17 | ACTCTTTGGAAGACGAGGTGT | AGTGAACCTGGAACCCACAC   | 693                         |
| 18 | CTCAAAGGCACTCTCCCACCTC| GAAGGCCAGCCCTGAGACTTA  | 717                         |
| 19 | CATGTTCCACCTTCGACACT  | CCAGCTTAACTCCGTGCTTC   | 748                         |
| 20 | CTGCACAGTTGCAAGACACT  | GCAGAAGACCTGGAACAAAAG  | 620                         |
| 21 | CGTCTCTTTATGGCCATGGCT| GTTCAATCCCCCATGTCGCC   | 461                         |
| 22 | GCATGTGCTAAGCTGCCCTTA| TGTAATGAGGACACACTGGA   | 485                         |
| 23 | CTCAGATGATCCACCCACCT  | TGCTTCAAAGACACCACT     | 617                         |
| 24 | GAGGTAGCCTTTCGAGCC    | ATGATTCATCTGCATCCCGG   | 791                         |
| 25 | AAAGCAGCTCAGATTCGA    | AAAGCATCCCCCTGTACCT    | 674                         |
| 26 | AAAATGCAAGGAGGACACAG  | AAATCTGTGGGCTTACATT    | 618                         |
| 27 | CTGCTTGGATGGGAGAGTT   | GAAAGCATGAAATGGGCCAA   | 669                         |
| 28 | GAGAGGCTGACGCTGAGCT   | ACACTGAAATGGAGGGAAG    | 689                         |
| 29 | TCTGTGAGCTGCTGTTGTTGA| AAAGAGGAGGATCCACACTG   | 556                         |
| 30 | CCTGATTTGCTTGGGTTAG   | AAGACACAGAATAACACTT    | 640                         |
| 31 | GGGAAGCAGCTCAGATTCGA  | AAAGCAGGAAAACACTTGG    | 51                          |

Genotyping SNP

| Genotyping SNP | Forward primer (5’→3’) | Reverse primer (5’→3’) | Size (bp) including primers |
|---------------|------------------------|------------------------|-----------------------------|
| c.128G > C    | TCAGAAGACCGTGCTCTCG    | ACAGGGCCAGAGTAAAAA     | 49                          |
| c.616-7942A > G | AGCAACAGGGCAAAACACTC  | GACGGAGGCTTAAATGCCAG   | 76                          |
| c.816G > A    | CTTGTTCTCAGGAGCCG     | GCAGGATCCTTGGAGAGTA    | 51                          |
| c.825 T > C   | GCCGGAAGGTCCCAGATA    | ATTTGAGCTCTGAGGAGG     | 48                          |
| c.1062 T > C  | GATGTTGCTGAATTTCTGGA  | GTCTGGGGGCTTGTGTC      | 41                          |
| c.1218 + 8A > G | CATGAGATTCAAGAGCGCTG | GTGGAAGCTCGGCCCCACAT   | 84                          |
| c.1299G > T   | CGAGAAATCTCCACGGTG    | GTAGTGAGGCAAGTCTGATGA | 80                          |
| c.1684 T > C  | GGTCTGCTTGCGAGGAGCC   | TTCAGGGGAAACCGGAGGAT  | 125                         |
| c.1704C > T   | TGTTGCTGAGCTGAGGAG    | ATCAGGGGAAACCGGAGGAT  | 104                         |
| c.2012G > T   | TACATTCTTGCTATCCCCG   | CACTTTGTGCTACATCAGGCA | 67                          |
| c.2168G > A   | CAGGCTTGATTCAGAATA    | ATGATTTCTCCAGCTGACAT  | 67                          |
| SNP       | Forward Primer          | Reverse Primer          | Position |
|-----------|-------------------------|-------------------------|----------|
| c.2581G > A | CAGGAGCTGGCTGGCTCG      | TCTGTGCTGGCATAGGTACG    | 62       |
| c.2965G > A | TCCTTTTCATGTGTAACCATGTG | CAATAGTTGGAAGCCAGCG     | 46       |
| c.3140G > C | CTACTCCATGCGCCGTGTC    | AGGATGCTGTCAGCGAGGT     | 75       |
| c.3436G > A | GAGGTTCTACGCTGGCTTCC  | GAATAGACCGGGGAGCG      | 69       |
| c.4002G > A | GACAAGTCCGATGCCAG     | CGAAATAAGCCCGGATGC     | 112      |
| c.4009A > G | AGCTGGGAAGTCTCCCT    | TGATCTCTCCGCGAGG       | 65       |

104* - upstream from the START codon. 21** - downstream from the STOP codon. Bold variants means these ones which genetic variation was detected during genotyping method.
based on the reference sequence NP_004987.2. These data were obtained from GenBank (http://www.ncbi.nlm.nih.gov) and were used in this paper as the nomenclature of variants.

The obtained data as supporting materials are presented in details in Additional files 1, 2, 3 and 4.

**Prediction of impact of SNPs upon protein functionality**

Using PolyPhen-2 v2.2.2r398 tool (http://genetics.bwh.harvard.edu/pph2/), we performed automatic prediction of effects of missense SNPs and amino acid substitutions for protein functionality and structure. PolyPhen-2 uses damaging alleles datasets, like HumDiv or HumVar, to calculate naïve Bayes posterior probability and to classify a mutation deleterious effect as benign, possibly damaging or probably damaging. This ternary classification depends on thresholds of false positive rate (FPR) denoting the chance that mutation would be falsely qualified as deleterious. For HumDiv-trained model, the thresholds are 5 % and 10 % FPR while for HumVar-trained model they are 10 % and 20 % FPR, respectively [28].

**Linkage disequilibrium and haplotype blocks analysis**

The observed genotype distribution was determined for all the detected polymorphisms by performing the Hardy-Weinberg equilibrium (HWE) exact test assuming consistency for P-value higher than 0.001. Linkage disequilibrium (LD) and haplotype block analysis was performed by Haplovie 4.2 software (http://www.broad.mit.edu/mpg/haplovie/). The LD analysis was made using $D'$ and $r^2$ parameters for each variation pair. $D'$ coefficient was preferentially used to model recombination rates in examined population, and $r^2$ parameter to model association power [29, 30].

Studying haplotypes in the human genome requires definition of the location of haplotype blocks, which are regions with little evidence for historical recombination and within which only a few common haplotypes are observed [31]. To determine haplotype blocks in *ABCC1*...
for the Polish population, we used the method described previously by Gabriel et al. [31] based on confidence bounds on $D'$. For pairs of SNPs being in strong linkage, 95% upper confidence bound $D'$ was higher than 0.98 and the lower bound was above 0.7 while for pairs of SNPs with strong evidences of historical recombination upper bound on $D'$ is lower than 0.9. According to this model, polymorphisms with MAF [22] lower than 0.05 were excluded during the creation of blocks and 16 polymorphic variants with higher MAF were included. Haplotypes with frequency lower than 1% were grouped as "others".

Results

ABCC1 genetic variation detected in this study

We screened coding regions and exon/intron boundaries as the most important gene areas (Additional file 1). To assure optimal detection level, the maximal length of tested PCR product was never longer than 250 and 125 base pairs for scanning and genotyping, respectively. For scanning, most of exons had to be divided into separate scanned areas, which provided 51 fragments for analysis. We performed complete polymorphism screening in 30 exons in total. We were unable to scan exon 1 because of problems with amplification during PCR, caused by extremely GC-rich sequence region.

In the scanning study on 190 Polish volunteers, we found 46 different SNPs in the scanned areas, and seven of them being novel and not reported yet (Table 2). 21 SNPs were located in exons and 11 of them change amino acid sequence as non-synonymous variants, including one novel polymorphic variant detected in this study: c.596C > T (p.Ser199Leu) with MAF = 0.003. Among all detected non-synonymous variants, only one, c.2012G > T (p.Gly671Val), occurred as a homozygote with estimated MAF = 0.077. The rest of SNPs found homozygously in exons were synonymous variants which resided at: c.825 T > C (p.Val275=), c.1062 T > C (p.Asn354=), c.1684 T > C (p.Leu562=), c.1704C > T (p.Tyr568=) and c.4002G > A (p.Ser1334=) with MAF values equal to 0.309, 0.332, 0.130, 0.088 and 0.277, respectively. Apart from polymorphisms in coding regions, 23 SNPs were located in introns and a last one (c.3°C > G) in the 3’ UTR. Furthermore, one intron deletion variant c.2461-39_2461-38delAT was found during the scanning. Two polymorphic variants located in introns, c.2871 + 26C > T and c.3079 + 10G > A, occurred only as non-reference homozygotes in all subjects.

We used the data from genotyping study to validate our scanning results and compare accuracy of HRM methods (Additional file 2). For genotyping, we tested a subset of 17 SNPs, selected according to the literature or detected during scanning. Seven of them have not been found during scanning in examined samples and also did not show any genetic variation by genotyping method. Genotype validation by sequencing was performed for all of examined SNPs (Table 3). We calculated MAFs to enable us to assess the accuracy of scanning method compared to genotyping. Results obtained for eight of SNPs were quite similar and the differences were not statistically significant (Table 3). We found statistically significant difference in MAF values obtained in both our studies for only two loci – c.1062 T > C (p.Asn354+) and c.2012G > T (p.Gly671Val).

All the amino acids altered by non-synonymous variants were located in cytoplasmic domains of the protein. Variants c.596C > T (p.Ser199Leu) and c.814C > T (p.Pro272Ser) were located in the third intracellular loop (between TM5 and TM6), variant c.1299G > T (p.Arg433Ser) in the fourth intracellular loop, variant c.1396C > T (p.Arg1066Trp) in the seventh intracellular loop (between TM7 and TM8). Additional four variants located in the loop containing NBD1 alter amino acids sequence: c.1898G > A (p.Arg633Gln) and c.2012G > T (p.Gly671Val) are located 44 and 6 amino acids upstream of the Walker A motif, respectively, while c.2168G > A (p.Arg723Gln) and c.2876A > G (p.Lys959Arg) are located 37 amino acids downstream of this motif, respectively. Similarly, three variants changing amino acids in the loop containing NBD2 were detected: c.3886C > T (p.Arg1296Trp) and c.3901C > T (p.Arg1301Cys) are located 30 and 25 amino acids upstream of the Walker A motif, respectively, while c.4093G > A (p.Asp1365Asn) is located 30 amino acids downstream of the motif. Detected variant c.4002G > A (p.Ser1334=) occurred in the last position of Walker A motif in NBD2, however it does not change the amino acid sequence.

Analysis of all the non-synonymous variants detected in this study by the PolyPhen-2 tool (data in Additional file 3) showed for HumDiv-trained model that five of them: c.814C > T (p.Pro272Ser), c.1898G > A (p.Arg633Gln), c.2168G > A (p.Arg723Gln), c.2876A > G (p.Lys959Arg), the novel one c.4093G > A (p.Asp1365Asn) probably have benign influence on the functioning of the protein. For one polymorphism, c.3196C > T (p.Arg1066Trp) a possibly damaging effect on protein activity was expected. The novel variant c.596C > T (p.Ser199Leu) was estimated as a probably damaging substitution, likewise as four others: c.1299G > T (p.Arg433Ser), c.2012G > T (p.Gly671Val), c.3886C > T (p.Arg1296Trp) and c.3901C > T (p.Arg1301Cys). On the other hand, analysis for HumVar-trained model indicated that three polymorphisms: c.1299G > T (p.Arg433Ser), c.2012G > T (p.Gly671Val), c.3901C > T (p.Arg1301Cys), lead to probably damaging substitutions and two others, c.596C > T (p.Ser199Leu) and c.3886C > T (p.Arg1296Cys), are possibly damaging variants.
| Exon scanned by HRM | rsID | Variant position | Intron/amino acid residue | dbSNP ID | Observed genotypes$^*$ $^b$ (n) | HWE exact test P-value$^c$ | MAF$^d$ |
|---------------------|------|------------------|--------------------------|----------|--------------------------------|--------------------------|--------|
| 2                   | rs8187843 | c.225 + 26G > A | Intron                  | NM_004996.3: | 164 25 0 1 (A) | 0.066 |
| 4                   | rs587783373 | c.352-79G > A | Intron                  | NM_004996.3: | 185 1 0 1 (A) | 0.003 |
| 4                   | rs4148337 | c.352-66 T > C | Intron                  | NM_004996.3: | 15 80 91 0.727 (T) | 0.296 |
| 5                   | rs483352860 | c.596C > T | p.Ser199Leu             | NP_004987.2: | 186 1 0 1 (T) | 0.003 |
| 6                   | rs8187846 | c.677 + 17C > T | Intron                  | NM_004996.3: | 188 1 0 1 (T) | 0.003 |
| 7                   | rs483352864 | c.809 + 16C > T | Intron                  | NM_004996.3: | 188 1 0 1 (T) | 0.003 |
| 7                   | rs45609533 | c.809 + 31G > T | Intron                  | NM_004996.3: | 183 5 0 1 (T) | 0.003 |
| 7                   | rs903880  | c.809 + 54C > A | Intron                  | NM_004996.3: | 112 65 11 0.684 (A) | 0.231 |
| 7                   | rs246232  | c.809 + 64C > G | Intron                  | NM_004996.3: | 84 90 14 0.174 (G) | 0.314 |
| 8                   | rs54694313 | c.810-73C > T | Intron                  | NM_004996.3: | 187 1 0 1 (T) | 0.003 |
| 8                   | rs200194736 | c.814C > T | p.Pro272Ser             | NP_004987.2: | 187 1 0 1 (T) | 0.003 |
| 8                   | rs587783372 | c.816G > A | Intron                  | NM_004996.3: | 187 1 0 1 (T) | 0.003 |
| 9                   | rs35587   | c.1062 T > C | Intron                  | NM_004996.3: | 187 1 0 1 (T) | 0.003 |
| 9                   | rs35588   | c.1218 + 8A > G | Intron                  | NM_004996.3: | 62 91 16 0.245 (G) | 0.327 |
| 10                  | rs483352877 | c.1218 + 9C > T | Intron                  | NM_004996.3: | 188 1 0 1 (T) | 0.003 |
| 10                  | rs60782127 | c.1299G > T | p.Arg433Ser             | NP_004987.2: | 186 2 0 1 (T) | 0.005 |
| 12                  | rs17265551 | c.1677 + 56C > T | Intron                  | NM_004996.3: | 162 27 0 0.604 (T) | 0.072 |
| 13                  | rs35604   | c.1678-37G > A | Intron                  | NM_004996.3: | 2 45 142 0.745 (G) | 0.130 |
| 13                  | rs483352863 | c.1678-34G > A | Intron                  | NM_004996.3: | 188 1 0 1 (T) | 0.003 |
| 13                  | rs35605   | c.1684 T > C | p.Leu562=               | NP_004987.2: | 2 45 142 0.745 (T) | 0.130 |
| 14                  | rs8187858 | c.1704C > T | p.Tyr568=               | NP_004987.2: | 157 31 1 1 (T) | 0.088 |
| 14                  | rs11228109 | c.1898G > A | p.Arg633Gln             | NP_004987.2: | 187 1 0 1 (A) | 0.003 |
| 15                  | rs8187863 | c.2001C > T | p.Ser667=               | NP_004987.2: | 187 1 0 1 (T) | 0.003 |
| 16                  | rs45511401 | c.2012G > T | p.Gly671Val             | NP_004987.2: | 161 25 2 0.296 (T) | 0.077 |
| 17                  | rs4148356 | c.2168G > A | p.Arg723Gln             | NP_004987.2: | 181 9 0 1 (A) | 0.024 |
| 19                  | rs45607032 | c.2461-39_2461-38delAT  | Intron                  | NM_004996.3: | 179 9 0 1 (delAT) | 0.024 |
| 19                  | rs2074087 | c.2461-30C > G | Intron                  | NM_004996.3: | 0 44 144 0.083 (C) | 0.117 |
| 19                  | rs45492500 | c.2461-27G > A | Intron                  | NM_004996.3: | 172 14 2 0.056 (A) | 0.048 |
| 21                  | rs11075296 | c.2871 + 26C > T | Intron                  | NM_004996.3: | 0 0 189 1 (T) | - |
| 22                  | rs768191257 | c.2876A > G | p.Lys959Arg             | NP_004987.2: | 187 1 0 1 (G) | 0.003 |
| 22                  | rs3851716 | c.3079 + 10G > A | Intron                  | NM_004996.3: | 0 0 188 1 (T) | - |
| 22                  | rs34794353 | c.3079 + 24C > T | Intron                  | NM_004996.3: | 187 1 0 1 (T) | 0.003 |
| 22                  | rs3887893 | c.3079 + 62T > C | Intron                  | NM_004996.3: | 67 96 25 0.358 (C) | 0.388 |
| 23                  | rs19107838 | c.3171G > A | p.Leu1057=              | NP_004987.2: | 187 2 0 1 (A) | 0.005 |
| 23                  | rs199773531 | c.3196C > T | p.Arg1066Trp            | NP_004987.2: | 188 1 0 1 (T) | 0.003 |
| 25                  | rs41278168 | c.3591-5C > T | Intron                  | NM_004996.3: | 187 1 0 1 (T) | 0.003 |
| 27                  | rs200922662 | c.3886C > T | p.Arg1296Trp            | NP_004987.2: | 187 1 0 1 (T) | 0.003 |
| 27                  | rs20153167 | c.3901C > T | p.Arg1301Cys            | NP_004987.2: | 187 1 0 1 (T) | 0.003 |
Linkage disequilibrium analysis

Based on full genotype sets of 44 polymorphic variants confirmed by Hardy-Weinberg equilibrium exact test (Table 2), linkage disequilibrium analysis using \( r^2 \) and \( |D'| \) statistics was performed (Additional file 4). Accordingly to the \( r^2 \) parameter, perfect linkage (\( r^2 = 1 \)) was detected between the new variant c.809 + 16C > T and c.3196C > T, between c.1678-37G > A and c.1684 T > C, between c.2168G > A and c.2461-39_2461-38delAT. Almost perfect linkage was observed between c.1062 T > C and c.1218 + 8A > G (\( r^2 = 0.928 \)), between c.825 T > C (p.Val275=) and c.1062 T > C (p.Asn354=) (\( r^2 = 0.822 \)) and between c.825 T > C (p.Val275=) and c.1218 + 8A > G (\( r^2 = 0.894 \)). Apart from perfect linkage between variants c.1678-37G > A and c.1684 T > C, they were both strongly linked (\( r^2 = 0.705 \)) with c.2461-30C > G. Other linkages (\( r^2 = 0.624 \) and \( r^2 = 0.658 \), respectively) were found between c.225 + 26G > A and c.816G > A, and between c.809 + 31G > T, c.809 + 54C > A, c.809 + 64C > G.

For the \( |D'| \) values, perfect linkage (\( |D'| = 1 \)) was observed for the great majority of variants pairs (760 for 946). Strong linkage (\( 1 > |D'| \geq 0.8 \)) was detected for other 17 pairs. Small areas of low \( |D'| \) values, and hence weak or lacking of linkage, were defined between polymorphisms from group: c.225 + 26G > A, c.352-66 T > C, c.809 + 31G > T, c.809 + 54C > A, c.809 + 64C > G, c.816G > A, c.825 T > C, c.1062 T > C, c.1218 + 8A > G, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C, c.1677 + 56C > T, and those from group: c.1678-37G > A, c.1684 T > C.
c.1704C > T, c.2012G > T, c.2168G > A, c.2461-39_2461-38delAT, c.2461-30C > G, c.2461-27G > A, c.3079 + 62 T > C, c.4002G > A, c.4126-45G > A, c.4487 + 18G > A.

Haplotype block analysis

Based on the model of Gabriel et al. [31] a total of 12 common haplotypes were identified and their frequencies were defined in four blocks (Table 4). In block 1, variants c.809 + 54C > A and c.809 + 64C > G were included, and the most frequent haplotype in this block was that containing major alleles (frequency: 0.686), the next most frequent one was the haplotype with both minor alleles (frequency: 0.231). Block 2 spanned SNPs c.825 T > C (p.Val275=), c.1062 T > C (p.Asn354=) and c.1218 + 8A > G at a distance of 1 kb, with the most frequent haplotype containing all major alleles (frequency: 0.659), subsequently the haplotype with all minor alleles (frequency: 0.300), two other ones had low frequency (set alleles TCG: 0.021 and set alleles TCA: 0.011). Block 3 had 3 haplotypes composed of SNPs c.1678-37G > A and c.1684 T > C (p.Leu562=). The first haplotype was the set with major alleles (frequency: 0.870) and the second one the set with minor alleles (frequency: 0.130). Block 4 was created of SNPs c.4126-45G > A and c.4487 + 18G > A at a distance of around 2 kb. The most frequent haplotype was that of alleles combination AG (frequency: 0.447), the next ones with major alleles (frequency: 0.405) and allele combination GA (frequency: 0.148). The frequency of recombination was also calculated between blocks as a value of multilalllelic R’ coefficient, and equaled to 0.65 between block 1 and 2, 0.55 between block 2 and 3 and 0.27 between block 3 and 4. HaploView predicted 51 possible connections of haplotypes for recombination between blocks higher than 1 %. However, recombination between blocks higher than 10 % occurred only between the most frequent haplotype from block 1 and the most frequent one from block 2, between three most frequent haplotypes from block 2 and the most frequent one from block 3, and between the most frequent one from block 3 and three haplotypes from block 4. Apart from c.1684 T > C (p.Leu562=), the remaining SNPs in blocks were tagged in silico as haplotype tag SNPs (htSNPs).

Discussion

In this paper we highlighted the fact that the High Resolution Melting method is a powerful technique both for targeted genotyping of selected SNPs and for scanning of long parts of gene sequences. The obtained results for targeted genotyping and scanning of ABCCI gene regions did not show statistically significant differences in MAFs values for most SNPs. Marginal statistical significance (0.05 > P > 0.01) of difference between MAF values for scanning and targeted genotyping methods for two SNPs c.1062 T > C (p.Asn354=) and c.2012G > T (p.Gly671Val) is probably caused by sample size effect (there was no overlap between groups of individuals selected for both stages of the study). Seven SNPs from a subset of 17 individually genotyped ones were not detected during the scanning and did not show any genetic variation during genotyping either. Although these variants were previously indicated to may have clinical significance and were therefore included in our study, they were only identified during in vitro studies or in non Caucasian populations (Exome Aggregation Consortium (ExAC), Cambridge, MA (http://exac.broadinstitute.org), date accessed: Aug 2015). Thus, it was expected not to find them in our study for 380 genotyped samples, where technical or methodological problems with detection should not have caused any discrepancy.

During the scanning study we identified 46 polymorphic variants by screening 30 exons of the ABCCI gene with adjacent intronic fragments in 190 unrelated Polish individuals. A large fraction of variants represented rare or even completely novel (not previously published) ones. We were unable to compare them with the ones from the only other publication on ABCCI variant distribution in Caucasians because of unavailability of the raw data from that study [15].

Comparing our genetic variation results obtained using the scanning method with data for Asian populations, 18 variants were detected in the Polish population which were present also in the Japanese one (n = 153) [13], while 13 variants from the Polish population were also detected in the Chinese one (n = 27) [14]. Some of these variants exhibited significant differences in MAF values, confirmed by chi-square test for alleles (with Yates’ correction if any allele quantity was below 5) and showed as their P-values, assumed significant if <0.05. Higher MAF for the Polish population was detected for variant c.809 + 54C > A (MAF = 0.231, compared to 0.059 for Japanese (P < 0.001) and 0.093 for Chinese (P = 0.020) populations), synonymous SNP c.4002G > A (MAF = 0.277, compared to 0.196 for Japanese (P = 0.014) and only 0.111 for Chinese (P = 0.009) populations) and for variant c.4126-45G > A between Polish (MAF = 0.447) and Japanese (MAF = 0.304) populations (P < 0.001). On the other hand, lower MAF for Polish population was noticed for variants: c.809 + 64C > A (0.313 for Polish, compared 0.418 (P = 0.004) for Japanese and 0.481 (P = 0.014) for Chinese populations), c.1678-37G > A and synonymous SNP c.1684 T > C (for both SNPs, MAF = 0.130 for Polish, 0.245 (P < 0.001) for Japanese but not significant for Chinese 0.204 (P = 0.141), non-synonymous SNP c.2168G > A and deletion variant c.2461-39_2461-38delAT (for both variants, MAF = 0.024 for Polish, 0.065 (P = 0.007) for Japanese and not significant for Chinese 0.065 (P = 0.108)), c.2461-30C > G (for Polish MAF = 0.117, not significant for Chinese 0.167 (P = 0.299) but for Japanese 0.245.
| Variant position | Block 1 | Recombination between blocks | Block 2 | Recombination between blocks | Block 3 | Recombination between blocks | Block 4 | Recombination between blocks |
|------------------|---------|-------------------------------|---------|-------------------------------|---------|-------------------------------|---------|-------------------------------|
|                 | c.809 + 54C | c.809 + 64C > G | c.825 T>C | c.1062 T>C | c.1218 + 8A > G | c.1678-37G > A | c.1684 T>C | c.4126-45G > A | c.4487 + 18G > A |
| Variant frequency |
|------------------|---------|-------------------------------|---------|-------------------------------|---------|-------------------------------|---------|-------------------------------|
| Haplotypes       | C       | C                             | 0.686   | 0.65                          | T       | T                             | A       | 0.659                         | 0.45 | A                             | C       | 0.870                         | 0.27 | A                             | G       | 0.447                         |
|                  | A       | G                             | 0.231   | G                             | C       | C                             | G       | 0.300                         | G    | T                             | 0.130                           | G    | G                             | 0.405                           |
|                  | C       | G                             | 0.082   | G                             | T       | C                             | G       | 0.021                         | G    | A                             | 0.148                           |
| others           | 0.001   | T                             | C       | A                             | 0.011   | others                        | 0.009   |                                |      |                               |                                  |

*Position according to reference sequences for coding nucleotides NM_004996.3 and amino acid residues NP_004987.2. Bold variants signify haplotype tag SNPs (htSNPs). Major alleles in white boxes, minor alleles in shaded boxes; all the haplotypes with their frequencies in population; haplotypes below 1% grouped as “others”. Level of recombination between blocks as a value of multiallelic D*
is consistent with the conclusion of another study on Caucasians [15]. Although blocks for our results are small and the number of SNPs used for haplotype construction is low, the main core of blocks and their borders are retained across populations regardless of the blocks definitions used. Our blocks 1 and 2, usually connected as one in other studies, extend from intron 7 to intron 9. Our block 3 include SNPs from vicinity of exon 13 and block 4 contains SNPs from intron 28 to intron 30. It corresponds to block borders detected in other studies and suggest points of recombination which had taken place in intron 12 and in intron 27 or 28 [13–15].

Analysis using PolyPhen-2 which predicts the influence of detected non-synonymous SNPs on protein functionality suggested that six of them are suspect. However, each analysis is only a statistical calculation based on a full dataset and some experimental data do not support these results. Létourneau et al. [33] examined 10 missense SNPs of ABCC1 and almost all of them, classified by PolyPhen-2 software as deleterious, affected neither expression nor functionality of the protein. A similar observation was also reported previously for the SNP c.2012G > T (p.Gly671Val), detected also by us. Although this substitution affects a residue near the highly conserved motif Walker A, it reduced mRNA expression but did not change transport properties of the protein, although PolyPhen-2 analysis clearly showed its damaging impact [34]. Recently, other data confirmed these observations and none of the amino acid substitutions: p.Arg633Gln, p.Gly671Val, p.Arg723Gln, detected also in our study, was found to change functionality of MRP1 transporter. Two substitutions altered drug resistance and inhibitor sensitivity and we found one of them, c.4002G > A (p.Ser1334=) [35].

On the other hand, there are other papers showing contrasting data for above-mentioned variants, proving their clinical association or experimental impact on protein functionality. Thus, it was demonstrated that variant c.2168G > A (p.Arg723Gln) reduced resistance activity of MRP1-overexpressing cells to drugs like daunorubicin, doxorubicin, etoposide, vinblastine and vincristine in vitro [36]. Although correlation between carrying this SNP and chemotherapy response was not observed in lung cancer patients [37], this variant was linked to the response to chemotherapy in patients with ovarian cancer [19]. Two other polymorphic variants detected in this study c.825 T > C (p.Val275=) and c.2012G > T (p.Gly671Val), were correlated with febrile neutropenia as an effect of FEC-induced hematological toxicity in breast cancer patients [38]. The major allele variant c.825 T > C (p.Val275=), in combination with another one c.*543C > T (beyond our scanning), was also associated with anthracycline-induced cardiotoxicity after chemotherapy in children with acute lymphoblastic leukemia [39]. The
discrepancy observed in different studies on clinical significance of c.2012G > T (p.Gly671Val) and c.2168G > A (p.Arg723Gln) is of unclear origin. Cellular conditions in vitro in which mutants were expressed could be inadequate to reveal all their real functions, e.g. a change in mRNA structure could be more essential than the amino acid change caused by SNP [34]. Conseil and Cole [35] concluded from this discordance that comprehensive biochemical and pharmacological characterization of the respective MRP1 mutants should be performed to identify consequences of ABCC1 SNPs on phenotype because these effects might be quite selective. Additional nucleic acid-based methods should be also used to clarify the impact of individual SNPs. These examples show that only comprehensive examination of polymorphisms and studying their phenotypes can tell more about their real influence. Therefore, novel variants like c.596C > T (p.Ser199Leu), detected in this study and qualified in silico as damaging, need to be checked in an experimental investigation. Additionally, the fact that we found no novel mutation in important nucleotide binding motifs, nor any of the mutations identified previously in vitro as deleterious in NBD loops, proved that these highly conserved regions in the protein structure are unlikely to be affected by polymorphisms occurring in the healthy population.

Conclusions
We scanned 30 exons and short flanking intron sequences of ABCC1. HRM is an efficient and sensitive method for scanning and genotyping polymorphic variants. We compared results of long-range scanning to targeted genotyping of selected SNPs showing that both methods are highly sensitive tools to study genetic variation. During our study we found 45 polymorphic variants, some of them rare and eight of them novel not previously reported. This is the first study about ABCC1 genetic variation in the Polish population and we found some ethnic differences in frequency of genetic variants compared to other populations. These results need further complementation with a larger number of probands and new data including also introns and untranslated regions as sequences with polymorphic variants which may have clinical relevance as well.

Additional files

Additional file 1: Supplemental materials for scanning study (include HRM plots and sequencing traces). (DOCX 12174 kb)

Additional file 2: Supplemental materials for targeted genotyping study (include HRM plots and sequencing traces). (DOCX 5582 kb)

Additional file 3: Supplemental materials for structural and functional disruption - in silico analysis using PolyPhen-2 software. (DOCX 13 kb)

Abbreviations
ABC transporters: ATP-binding cassette transporters; E;17βG: Estradiol-17β-D-glucuronide; GSH: Glutathione; HRM: High resolution melting; hSNPs: Haplotype tag SNPs; HWE: Hardy Weinberg equilibrium; LD: Linkage disequilibrium; MAF: Minor allele frequency; MRP1: Multidrug resistance-associated protein 1; NBD: Nucleotide binding domain; TM: Transmembrane helix; TMD: Transmembrane domain.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
Conceived and designed the experiments: DS, MSI. Performed the experiments: MSI, MSO, MKM. Analysed data: MSI, MSO, DS. Contributed reagents/materials/analysis tools DS, GB. Contributed to the writing of manuscript: MSI, DS, MSO, JD, GB. All authors approved the final manuscript.

Acknowledgments
The authors would like to thank the reviewers for their comments that help to improve the manuscript. In addition we would like to thank Łukasz Pulaski for fruitful discussion and language editing. This work was supported by the Polish POIG grant 01.01.02-10-005/08 TESTOPLEK from the European Regional Development Fund.

Author details
1Biobank Lab, Department of Molecular Biophysics, Faculty of Biology and Environmental Protection, University of Łódź, Pilsarskiego 14/16, 90-231 Łódź, Poland. 2Institute of Medical Biology, Polish Academy of Sciences, Lodowa 106, 93-232 Łódź, Poland. 3Department of Molecular Biophysics, Faculty of Biology and Environmental Protection, University of Łódź, Pomorska 141/143, 90-236 Łódź, Poland.

Received: 14 April 2015 Accepted: 11 September 2015
Published online: 23 September 2015

References
1. Leslie EM, Deeleey RG, Cole SP. Multidrug resistance proteins: role of P-glycoprotein, MRP1, MRP2, and BCRP (ABC2) in tissue defense. Toxicol Appl Pharmacol. 2005;204(3):216–37.
2. Cole SP, Bhurdwaj G, Gerlich JH, Mackie JE, Grant CE, Almquist KC, et al. Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. Science. 1992;258(5088):1650–4.
3. Bakos E, Homolya L. Portrait of multifunctional transporter, the multidrug resistance-associated protein 1 (MRP1/ABCC1). Pfuicher Archs. 2007;453(5):621–41.
4. Chang XB. A molecular understanding of ATP-dependent solute transport by multidrug resistance-associated protein MRP1. Cancer Metastasis Rev. 2007;26(1):15–37.
5. Choudhuri S, Klaassen CD. Structure, function, expression, genomic organization, and single nucleotide polymorphisms of human ABCB1 (MDR1), ABCC (MRP), and ABCG2 (BCRP) efflux transporters. Int J Toxicol. 2006;25(4):231–59.
6. DeGorter MK, Conseil G, Deeleey RG, Campbell RL, Cole SP. Molecular modeling of the human multidrug resistance protein 1 (MRP1/ABCC1). Biochem Biophys Res Commun. 2008;365(1):29–34.
7. Leslie EM, Deeleey RG, Cole SP. Toxicological relevance of the multidrug resistance protein 1, MRP1 (ABCC1) and related transporters. Toxicology. 2001;167(1):3–23.
8. Flens MJ, Zaman GL, van der Vaak P, Izquierdo MA, Schoeijers AB, Scheffer GL, et al. Tissue distribution of the multidrug resistance protein. Ann J Pathol. 1996;148(4):1237–47.
9. Sharam FJ. ABC multidrug transporters: structure, function and role in chemoresistance. Pharmacogenomics. 2008;9(1):105–27.
10. Deeleey RG, Cole SP. Substrate recognition and transport by multidrug resistance protein 1 (ABCC1). FEBS Lett. 2006;580(4):1103–11.
11. Munoz M, Henderson M, Haber M, Norris M. Role of the MRP1/ABCC1 multidrug transporter protein in cancer. IUBMB Life. 2007;59(12):752–7.

12. Saito S, Ida A, Sekine A, Miura Y, Ogawa C, Kawauschi S, et al. Identification of 779 genetic variations in eight genes encoding members of the ATP-binding cassette, subfamily C (ABC/MRP/CFTR). J Hum Genet. 2004;49(4):147–71.

13. Fukushima-Linska A, Saito Y, Tohkin M, Maekawa K, Hasegawa R, Kawamoto M, et al. Genetic variations and haplotype structures of the ABC transporter gene ABC1C1 in a Japanese population. Drug Metab Pharmacokinet. 2007;22(1):48–60.

14. Wang H, Hao B, Zhou K, Chen X, Wu S, Zhou G, et al. Linkage disequilibrium and haplotype architecture for two ABC transporter genes (ABCC1 and ABCG2) in Chinese population: implications for pharmacogenomic association studies. Ann Hum Genet. 2004;68(8Pt 6):653–73.

15. Leschziner G, Zabaneh D, Pirmohamed M, Owen A, Rogers J, Coffey AJ, et al. Exon sequencing and high resolution haplotype analysis of ABC transporter genes implicated in drug resistance. Pharmacogenet Genomics. 2006;16(6):439–50.

16. Wang Z, Sew PH, Ambrose H, Ryan S, Chong SS, Lee EJ, et al. Nucleotide sequence analyses of the MRP1 gene in four populations suggest negative selection on its coding region. BMC Genomics. 2006;7:111.

17. Pajic M, Murray J, Marshall GM, Cole SP, Norris MD, Haber M, ABC1C1 G2012T single nucleotide polymorphism is associated with patient outcome in primary neuroblastoma and altered stability of the ABC1C1 gene transcript. Pharmacogenet Genomics. 2011;21(5):270–9.

18. Kunicka T, Vrana D, Pech V, Rauss K, et al. Non-coding polymorphisms in nucleotide binding domain 1 in ABC1C1 gene associate with transcript level and survival of patients with breast cancer. PLoS One. 2014;9(7):e107140.

19. Obata H, Yahata T, Quan J, Sekine M, Tanaka K. Association between single nucleotide polymorphisms of drug resistance-associated genes and response to chemotherapy in advanced ovarian cancer. Anticancer Res. 2002;22(3B):2227–32.

20. Siedlinckx M, Boezen HM, Boer JM, Smit HA, Postma DS. ABC1C1 polymorphisms contribute to level and decline of lung function in two population-based cohorts. Pharmacogenet Genomics. 2009;19(9):675–84.

21. Budulac SE, Postma DS, Hiemstra PS, Kunz LI, Siedlinckx M, Smit HA, et al. Multidrug resistance-associated protein-1 (MRP1) genetic variants, MRP1 protein levels and severity of COPD. Respir Res. 2010;11:60.

22. Maffini A, Ortombina M, Serret-Gaudelus I, Lebecque P, Leal T, Iansa P, et al. Impact of polymorphism of Multidrug Resistance-Associated Protein 1 (ABCC1) gene on the severity of cystic fibrosis. J Cyst Fibros. 2011;10(4):228–33.

23. Lee SH, Lee MS, Lee JH, Kim SW, Kang RH, Choi MJ, et al. MRPI polymorphisms associated with citralopram response in patients with major depression. J Clin Psychopharmacol. 2010;30(2):116–25.

24. Judson R, Stephens JC, Windemuth A. The predictive power of haplotypes in clinical response. Pharmacogenomics. 2002;3(1):15–26.

25. Wang Z, Wang B, Tang K, Lee EJ, Chong SS, Lee CG. A functional polymorphism within the MRP1 gene locus identified through its genomic signature of positive selection. Hum Mol Genet. 2005;14(14):2075–87.

26. Majewska M, Sobalska M, Siewierska A, Strapagiel D, Bartosz G. Single nucleotide polymorphisms of selected ABC genes in the Polish population. In 4th FEBS special meeting ATP-binding cassette (ABC) proteins: from multidrug resistance to genetic diseases; March 3–9, 2012; Innsbruck, Austria. FEBS. 2012. pp. 84.

27. Koszarska M, Kucinska N, Kiss K, Varady G, Szymanski DM, Antalffy G, et al. Screening the expression of ABCB6 in erythrocytes reveals an unexpectedly high frequency of 36 mutations in healthy individuals. PLoS One. 2014;9(10):e111590.

28. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248–9.

29. Shifman S, Kuypers J, Kokoris M, Yakir B, Danavi A. Linkage disequilibrium patterns of the human genome across populations. Hum Mol Genet. 2003;12(7):771–6.

30. Muller JC. Linkage disequilibrium for different scales and applications. Brief Bioinform. 2004;5(4):355–64.

31. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science. 2002;296(5576):2265–9.

32. Oselin K, Mrozikiewicz PM, Gaivkovitch E, Pahlka R, Roots I. Frequency of MRP1 genetic polymorphisms and their functional significance in Caucasians: detection of a novel mutation G618A in the human MRP1 gene. Eur J Clin Pharmacol. 2003;59(4):347–50.

33. Letoumeau IJ, Deelely RG, Cole SP. Functional characterization of non-synonymous single nucleotide polymorphisms in the gene encoding human multidrug resistance protein 1 (MRP1/ABCC1). Pharmacogenet Genomics. 2005;15(9):647–57.

34. Conrad S, Kauffmann HM, Ito K, Deelely RG, Cole SP, Schenk D. Identification of human multidrug resistance protein 1 (MRP1) mutations and characterization of a G671T substitution. J Hum Genet. 2001;46(11):656–63.

35. Consiell G, Cole SP. Two polymorphic variants of ABCC1 selectively alter drug resistance and inhibitor sensitivity of the multidrug and organic cation transporter multidrug resistance protein 1. Drug Metab Dispos. 2013;41(12):2187–96.

36. Yin JY, Huang Q, Yang Y, Zhang JT, Zhong MZ, Zhou HH, et al. Characterization and analyses of multidrug resistance-associated protein 1 (MRP1/ABCC1) polymorphisms in Chinese population. Pharmacogenet Genomics. 2009;19(3):206–16.

37. Yin JY, Han LF, Huang Q, Xu XJ, Zhou HH, Liu QZ. ABCC1 polymorphism Arg723Gln (2168G > A) is associated with lung cancer susceptibility in a Chinese population. Clin Exp Pharmacol Physiol. 2011;38(8):632–7.

38. Vulssteke C, Lambrechts D, Diederonni A, Harts S, Brouwers B, van Brussel T, et al. Genetic variability in the multidrug resistance associated protein-1 (ABCC1/MPR1) predicts hematological toxicity in breast cancer patients receiving (neo)-adjuvant chemotherapy with 5-fluorouracil, epirubicin and cyclophosphamide (FEC). Ann Oncol. 2013;24(6):1513–25.

39. Sensei AF, Erdelyi DJ, Ungvari I, Czagyili E, Hegyi MZ, Kizsel PS, et al. ABCC1 polymorphisms in anthracycline-induced cardio toxicity in childhood acute lymphoblastic leukaemia. Cell Biol Int. 2012;36(1):79–86.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit