



(51) International Patent Classification:

C40B 40/06 (2006.01) *C07H 21/00* (2006.01)
A61K 31/65 (2006.01) *C12Q 1/68* (2006.01)
A61P 35/00 (2006.01) *C40B 30/04* (2006.01)

(21) International Application Number:

PCT/CA2012/000529

(22) International Filing Date:

31 May 2012 (31.05.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/491,408 31 May 2011 (31.05.2011) US
61/540,436 28 September 2011 (28.09.2011) US

(71) Applicant (for all designated States except US): **THE UNIVERSITY OF BRITISH COLUMBIA** [CA/CA]; University - Industry Liaison Office, #103 - 6190 Agromony Road, Vancouver, British Columbia V6T 1Z3 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **HAYDEN, Michael** [CA/CA]; 4484 West 7th Avenue, Vancouver, British Columbia V6R 1W9 (CA). **CARLETON, Bruce** [CA/CA]; 349 Monteray Avenue, North Vancouver, British Columbia V7N 3E7 (CA). **ROSS, Colin** [CA/CA]; 5451 Heathdale Court, Burnaby, British Columbia V5B 2H7 (CA). **VISSCHER, Hendrik** [NL/NL]; Dorstige Harthof 21, 3512 NX Utrecht (NL).

(74) Agent: **MACINS, Andris, D.**; Smart & Biggar, 2300 - 1055 West Georgia Street, Box 1115 Royal Centre, Vancouver, British Columbia V6E 3P3 (CA).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: POLYMORPHISM PANELS PREDICTIVE OF ANTHRACYCLINE-INDUCED CARDIOTOXICITY (ACT)

(57) Abstract: Provided are methods, nucleic acids, and arrays for assessing the susceptibility of a subject to the development of cardiotoxicity in response to receiving one or more anthracycline compounds, the method including determining the presence or absence of one or more polymorphisms, wherein the presence or absence of one or more such polymorphisms is indicative of susceptibility to the development of cardiotoxicity.



POLYMORPHISM PANELS PREDICTIVE OF ANTHRACYCLINE-INDUCED CARDIOTOXICITY (ACT)

FIELD OF THE INVENTION

This invention relates to the field of genetic markers for adverse drug reactions. More specifically, methods useful for identifying individuals that may be at risk for an adverse drug reaction.

BACKGROUND

Anthracycline-induced cardiotoxicity (ACT) is one of the most important adverse drug reactions in childhood cancer therapy with potential life-long consequences causing substantial morbidity and mortality as well as limiting anthracycline use (Mertens, A.C. *et al.* J Clin Oncol (2001) 19:3163-3172; van Dalen, E.C. *et al.* Eur J Cancer (2006) 42:3191-3198; and Lipshultz, S.E. Heart (2008) 94:525-533). Nevertheless, anthracyclines are widely used – nearly 60% of childhood cancer patients receive anthracyclines, and help to improve cancer survival rates (van Dalen, E.C. *et al.* Eur J Cancer (2006) 42:3191-3198). ACT is a serious adverse drug reaction (ADR) of cancer therapy, and is, in part, mediated by genetic variation. Recently, several genetic variants predictive of ACT risk in children were identified and replicated. ADRs are a significant cause of illness, hospitalization and death for both children and adults in the Western world (LAZAROU *et al.* JAMA 1998; PIRMOHAMED *et al.* BMJ 2004). Estimates suggest that 15% of hospitalized children experience an ADR. Those that do survive the ADR may be left disabled (MITCHELL *et al.*, 1988 Pediatrics 82:24-9; MARTINEZ-MIR *et al.*, 1999. Br J Clin Pharmacol 47 :681-8).

Cardiotoxicity can occur early – during or within one year after therapy – or one or many years after treatment (late cardiotoxicity) (Lipshultz, S.E. Heart (2008) 94:525-533). It manifests as asymptomatic subclinical left ventricular dysfunction – usually diagnosed using echocardiography – in up to 57% of patients, and which can be progressive (Kremer, L.C. *et al.* Ann Oncol (2002) 13:819-829; Lipshultz, S.E. *et al.* J Clin Oncol (2005) 23:2629-2636; and van der Pal, H.J. *et al.* Arch Intern Med (2010) 170:1247-1255), or as severe clinical heart failure requiring treatment in as high as 16% (van Dalen, E.C. *et al.* Eur J Cancer (2006) 42:3191-3198; Lefrak, E.A. *et al.* Cancer (1973) 32:302-314; Von Hoff, D.D. *et al.* Ann Intern Med (1979) 91:710-717; and Kremer, L.C. *et al.* Ann Oncol (2002) 13:503-512).

Identification of patients at risk for ACT is important to allow for risk stratification and to inform treatment and monitoring options. Many clinical risk factors for ACT have been identified, most notably higher cumulative anthracycline doses and concomitant cardiac irradiation, with some of these known for

decades (Tukenova, M. *et al.* J Clin Oncol (2010) 28:1308-1315). More recently, several studies have started to unravel the genetic susceptibility to ACT – including in children – though only few variants have been replicated in independent cohorts (Blanco, J.G. *et al.* Cancer (2008) 112:2789-2795; Rajic, V. *et al.* Leuk Lymphoma (2009) 50:1693-1698; Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011; and Visscher H, Ross CJ, Rassekh SR *et al.* Validation of SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. Submitted to Cancer 2011).

Many approved drugs used in children are untested in pediatric populations. While it is known that children metabolize drugs differently than adults, in many cases pediatric dosage forms are not available. This is of particular concern with chemotherapy drugs, which may frequently be supplied as a single-dose package, and in combination with other agents, excipients and the like. Pediatric populations also represent a more varied population, and this increased variability may be due to developmental differences in the normal expression of drug metabolism genes.

Genetic factors are involved in variability in drug response – ranging from 20-95% in some studies. Age, sex, body weight, health, medical history and the like may be accounted for, but patient genotype is largely an unknown factor (EVANS *et al* 2003. NEJM 348:538-549; WEINSHILBOUM 2003. NEJM 348:529-537).

Anthracyclines are used as cytotoxic agents in chemotherapeutic protocols in both children and adults, for a variety of neoplasms. Examples of anthracyclines and anthracycline analogues include daunorubicin, doxorubicin, idarubicin and epirubicin. For example, anthracyclines may be used in the treatment of solid and hematologic cancers, such as breast cancer, acute myeloid leukemia, acute lymphoblastic leukemia, multiple myeloma, Hodgkin's disease or non-Hodgkin's lymphoma.

Cardiotoxicity is a serious problem in patient populations receiving anthracyclines, particularly pediatric patients (LIPSHULTZ 2006. Seminars in Oncology 33:S8-S14). Anthracycline-induced cardiotoxicity may result in cardiomyopathy and congestive heart failure and may be irreversible. Anthracycline-induced cardiotoxicity may be characterized by reduced ventricular wall thickness and mass, indicative of decreased cardiac muscle and depressed ventricular contractility. As mentioned, an increased dose, cumulative dose, nature of the particular anthracycline, administration route, age, sex and prior radiation treatment may affect onset and severity of cardiotoxicity. Administration of dexrazoxane may be beneficial in preventing or reducing cardiac injury during chemotherapy. However, the administration of enalapril, or antioxidants such as vitamin E, coenzyme Q10, carnitine, or glutathione, for example may

also be beneficial in preventing or reducing cardiac injury during chemotherapy. Other agents that may be administered to reduce anthracycline cardiotoxicity are described (WOUTERS et al 2005. Br. J Hematol 131:561-578).

Dose limits have been empirically set in the clinic, above which the cardiotoxicity is deemed to be unacceptable. Subclinical and clinical cardiotoxicity may occur below these doses (JOHNSON 2006. Seminars in Oncology 33:S33-70) and affect current and subsequent therapeutic regimens. Liposomal anthracycline compositions may demonstrate reduced cardiotoxicity (EWER et al 2004. Seminars in Oncology 31:161-181).

Proteomic methods have been developed for early detection of drug-induced cardiotoxicity (PETRICON *et al.* 2004. Toxicol Pathol 32:122-30).

Some polymorphisms have been associated with anthracycline-induced cardiotoxicity (WOJNOWSKI et al 2005. Circulation 112:3754-3762; Blanco, J.G. *et al.* Cancer (2008) 112:2789-2795; Rajic, V. *et al.* Leuk Lymphoma (2009) 50:1693-1698; Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011); and Rossi, D. et al. Leukemia (2009) 23(6):1118-26).

Genotype has been shown to alter response to therapeutic interventions. Genentech's HERCEPTIN® was not effective in its overall Phase III trial but was shown to be effective in a genetic subset of subjects with human epidermal growth factor receptor 2 (HER2)-positive metastatic breast cancer. Similarly, Novartis' GLEEVEC® is only indicated for the subset of chronic myeloid leukemia subjects who carry a reciprocal translocation between chromosomes 9 and 22.

SUMMARY

This invention is based in part on the identification that the particular nucleotide (allele) or genotype at the site of a given SNP may be associated with an increased likelihood of cardiotoxicity ('risk genotype') or a decreased likelihood of cardiotoxicity ('decreased risk genotype').

In accordance with a first embodiment, methods are provided for screening a subject having a neoplastic disease for cardiotoxicity risk, the method including: determining the identity for one or more of the following single nucleotide polymorphisms (SNPs): rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406;

rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto, for the subject, wherein the subject may be a candidate for anthracycline administration.

In accordance with a further embodiment, methods are provided for diagnosing a predisposition for cardiotoxicity risk in a human subject from anthracycline administration, the method including: a) determining an identity for one or more of the following single nucleotide polymorphisms (SNPs) in a biological sample from the subject: rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto from the sample; and b) making a cardiotoxicity risk determination based on the prevalence of risk alleles in the subject sample.

In accordance with a further embodiment, methods are provided for diagnosing a predisposition for cardiotoxicity risk in a human subject from anthracycline administration, the method including: a) obtaining a biological sample from the subject; b) determining an identity for one or more of the following single nucleotide polymorphisms (SNPs): rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto from the sample; and c) making a cardiotoxicity risk determination based on the prevalence of risk alleles in the subject sample.

In accordance with a first embodiment, methods are provided for screening a subject having a neoplastic disease for cardiotoxicity risk, the method including: determining the identity for one or more of the following single nucleotide polymorphisms (SNPs): rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto, for the subject, wherein the subject may be a candidate for anthracycline administration.

In accordance with a further embodiment, methods are provided for diagnosing a predisposition for cardiotoxicity risk in a human subject from anthracycline administration, the method including: a) determining an identity for one or more of the following single nucleotide polymorphisms (SNPs) in a biological sample from the subject: rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto from the sample; and b) making a cardiotoxicity risk determination based on the prevalence of risk alleles in the subject sample.

In accordance with a further embodiment, methods are provided for diagnosing a predisposition for cardiotoxicity risk in a human subject from anthracycline administration, the method including: a) obtaining a biological sample from the subject; b) determining an identity for one or more of the following single nucleotide polymorphisms (SNPs): rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; and rs17583889; one or more polymorphic sites in linkage disequilibrium thereto from the sample; and c) making a cardiotoxicity risk determination based on the prevalence of risk alleles in the subject sample.

In accordance with a first embodiment, methods are provided for screening a subject having a neoplastic disease for cardiotoxicity risk, the method including: determining the identity for one or more of the following single nucleotide polymorphisms (SNPs): rs7853758; rs885004; rs2305364; rs4982753; rs4149178; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto, for the subject, wherein the subject may be a candidate for anthracycline administration.

In accordance with a further embodiment, methods are provided for diagnosing a predisposition for cardiotoxicity risk in a human subject from anthracycline administration, the method including: a) determining an identity for one or more of the following single nucleotide polymorphisms (SNPs) in a biological sample from the subject: rs7853758; rs885004; rs2305364; rs4982753; rs4149178; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto from the sample; and b) making a cardiotoxicity risk determination based on the prevalence of risk alleles in the subject sample.

In accordance with a further embodiment, methods are provided for diagnosing a predisposition for cardiotoxicity risk in a human subject from anthracycline administration, the method including: a) obtaining a biological sample from the subject; b) determining an identity for one or more of the following single nucleotide polymorphisms (SNPs): rs7853758; rs885004; rs2305364; rs4982753; rs4149178; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto from the sample; and c) making a cardiotoxicity risk determination based on the prevalence of risk alleles in the subject sample.

In accordance with a first embodiment, methods are provided for screening a subject having a neoplastic disease for cardiotoxicity risk, the method including: determining the identity for one or more of the

following single nucleotide polymorphisms (SNPs): rs7853758; rs885004; rs2305364; rs4982753; and rs4149178; or one or more polymorphic sites in linkage disequilibrium thereto, for the subject, wherein the subject may be a candidate for anthracycline administration.

In accordance with a further embodiment, methods are provided for diagnosing a predisposition for cardiotoxicity risk in a human subject from anthracycline administration, the method including: a) determining an identity for one or more of the following single nucleotide polymorphisms (SNPs) in a biological sample from the subject: rs7853758; rs885004; rs2305364; rs4982753; and rs4149178; or one or more polymorphic sites in linkage disequilibrium thereto from the sample; and b) making a cardiotoxicity risk determination based on the prevalence of risk alleles in the subject sample.

In accordance with a further embodiment, methods are provided for diagnosing a predisposition for cardiotoxicity risk in a human subject from anthracycline administration, the method including: a) obtaining a biological sample from the subject; b) determining an identity for one or more of the following single nucleotide polymorphisms (SNPs): rs7853758; rs885004; rs2305364; rs4982753; and rs4149178; or one or more polymorphic sites in linkage disequilibrium thereto from the sample; and c) making a cardiotoxicity risk determination based on the prevalence of risk alleles in the subject sample.

The method may further include determining the identity of rs17863783 or one or more polymorphic sites in linkage disequilibrium thereto.

In accordance with a further embodiment, methods are provided for determining cardiotoxicity risk from anthracycline administration, the method including: determining the identity of the single nucleotide polymorphism (SNP) at each of the following polymorphic sites: rs7853758; rs885004; rs17863783; rs10426377; and rs2305364; or a polymorphic sites in linkage disequilibrium thereto, for a subject receiving or about to receive one or more anthracyclines.

In accordance with a further embodiment, methods are provided for determining cardiotoxicity risk from anthracycline administration, the method comprising: determining the identity of the single nucleotide polymorphism (SNP) at each of the following polymorphic sites: rs7853758; rs4148808; rs17863783; rs10426377; and rs2305364; or a polymorphic sites in linkage disequilibrium thereto, for a subject receiving or about to receive one or more anthracyclines.

In accordance with a further embodiment, methods are provided for determining cardiotoxicity risk from anthracycline administration, the method comprising: determining the identity of the single nucleotide polymorphism (SNP) at each of the following polymorphic sites: rs7853758; rs17863783; rs10426377; and rs2305364; or a polymorphic sites in linkage disequilibrium thereto, for a subject receiving or about to receive one or more anthracyclines.

The method may further include determining the identity of one or both of the following two SNPs: rs4982753; and rs4149178; or one or more polymorphic sites in linkage disequilibrium thereto.

The method may further include determining the cumulative anthracycline dose given to the subject and/or whether the subject received radiation therapy involving the heart region. The method may further include determining the sex of the subject or the age of the subject. The subject may preferably be female if the SNP is rs4148808 or rs1149222. The subject may preferably be male if the SNP is rs10426377. The subject may preferably be <5.3 yrs old if the SNP is rs17583889.

In accordance with a further embodiment, uses are provided for the manufacture of a medicament comprising an anthracycline compound having a cardiotoxicity risk for the treatment of a subject, where the subject is a candidate for anthracycline administration, and wherein the subject treated has a reduced cardiotoxicity risk genotype at one or more of the following polymorphic sites: rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628.

In accordance with a further embodiment, uses are provided for an anthracycline compound having a cardiotoxicity risk for the treatment of a subject, wherein the subject treated has a reduced cardiotoxicity risk genotype at one or more of the following polymorphic sites: rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628 for the subject, where the subject is a candidate for anthracycline administration.

The subject may also have a reduced cardiotoxicity risk genotype at rs17863783 or one or more polymorphic sites in linkage disequilibrium thereto.

In accordance with a further embodiment, there is provided an anthracycline for use in treating a neoplastic disease in a subject in need thereof, the method comprising: (a) selecting a subject having a

reduced risk of developing cardiotoxicity, wherein cardiotoxicity is based on the identity of a single nucleotide polymorphism (SNP) at one or more of the following polymorphic sites: rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto; and (b) administering said subject one or more anthracyclines.

The subject may also be selected based on a reduced cardiotoxicity risk genotype at rs17863783 or one or more polymorphic sites in linkage disequilibrium thereto.

In accordance with a further aspect of the invention, there is provided a method of selecting a therapeutic regimen for a subject, the therapeutic regimen including one or more opioids, the method including: determining the identity of a single nucleotide polymorphism (SNP) at one or more of the following polymorphic sites: rs7853758; rs17863783; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628, for the subject to assess the risk of toxicity.

The anthracycline may be selected from one or more of the following: anthracycline antibiotics such as daunorubicin (daunomycin, rubidomycin), doxorubicin, idarubicin, epirubicin, mitoxantrone, carminomycin, esorubicin, quelamycin, aclarubicin, esorubicin, zorubicin, pirarubicin, amrubicin, iododoxorubicin, detorubicin, marcellomycin, rodorubicin, and valrubicin. Alternatively, the anthracycline may be selected from one or more of the following: doxorubicin, and daunorubicin. The method may further include administering the anthracycline in accordance with the subject's risk of developing cardiotoxicity. The method may further include administering a cardioprotective agent. The method may further include administering an anthracycline that has a reduced toxicity risk associated therewith. The method may further include monitoring the subject for signs of cardiotoxicity.

The cardiotoxicity risk allele may be selected from one or more of: rs7853758G; rs885004G; rs17863783A (reverse); rs17863783T (forward); rs10426377C; rs2305364A; rs4982753G (reverse); rs4982753C (forward); rs4149178A; rs4148808A; rs1149222G; rs17583889A; rs4877847A; rs11625724T; rs12882406G; rs12896494T; and rs10426628A. The reduced cardiotoxicity risk allele may be selected from one or more of: rs7853758A; rs885004A; rs17863783C (reverse); rs17863783G (forward); rs10426377A; rs2305364G; rs4982753A (reverse); rs4982753T (forward); rs4149178G; rs4148808G; rs1149222T; rs17583889C; rs4877847C; rs11625724A; rs12882406C; rs12896494C; and rs10426628G. The identity of a single nucleotide polymorphism may be determined by one or more of

the following techniques: restriction fragment length analysis; sequencing; micro-sequencing assay; hybridization; invader assay; gene chip hybridization assays; oligonucleotide ligation assay; ligation rolling circle amplification; 5' nuclease assay; polymerase proofreading methods; allele specific PCR; matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectroscopy; ligase chain reaction assay; enzyme-amplified electronic transduction; single base pair extension assay; and reading sequence data.

In accordance with a further embodiment, there are provided two or more oligonucleotides or peptide nucleic acids of about 10 to about 400 nucleotides that hybridize specifically to a sequence contained in a human target sequence consisting of a subject's toxicity associated gene sequence, a complementary sequence of the target sequence or RNA equivalent of the target sequence and wherein the oligonucleotides or peptide nucleic acids are operable in determining the identity of two or more polymorphism(s) in the toxicity associated gene sequence selected from of the following polymorphic sites: rs7853758; rs885004; rs17863783; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628.

In accordance with a further embodiment, there are provided two or more oligonucleotides or peptide nucleic acids selected from the group:

- (a) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:1 having an A at position 301 but not to a nucleic acid molecule including SEQ ID NO:1 having a G at position 301;
- (b) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:1 having a G at position 301 but not to a nucleic acid molecule including SEQ ID NO:1 having an A at position 301;
- (c) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:2 having a G at position 201 but not to a nucleic acid molecule including SEQ ID NO:2 having an A at position 201;
- (d) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:2 having an A at position 201 but not to a nucleic acid molecule including SEQ ID NO:2 having an G at position 201;
- (e) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:3 having an A at position 301 but not to a nucleic acid molecule including SEQ ID NO:3 having a C at position 301;

- (f) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:3 having a C at position 301 but not to a nucleic acid molecule including SEQ ID NO:3 having an A at position 301;
- (g) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:4 having an A at position 501 but not to a nucleic acid molecule including SEQ ID NO:4 having a C at position 501;
- (h) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:4 having a C at position 501 but not to a nucleic acid molecule including SEQ ID NO:4 having an A at position 501;
- (i) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:5 having an A at position 251 but not to a nucleic acid molecule including SEQ ID NO:5 having a G at position 251;
- (j) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:5 having a G at position 251 but not to a nucleic acid molecule including SEQ ID NO:5 having an A at position 251;
- (k) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:6 having an A at position 501 but not to a nucleic acid molecule including SEQ ID NO:6 having a G at position 501;
- (l) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:6 having a G at position 501 but not to a nucleic acid molecule including SEQ ID NO:6 having an A at position 501;
- (m) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:7 having an A at position 401 but not to a nucleic acid molecule including SEQ ID NO:7 having a G at position 401;
- (n) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:7 having a G at position 401 but not to a nucleic acid molecule including SEQ ID NO:7 having an A at position 401;
- (o) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:8 having an A at position 301 but not to a nucleic acid molecule including SEQ ID NO:8 having a G at position 301;
- (p) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:8 having a G at position 301 but not to a nucleic acid molecule including SEQ ID NO:8 having an A at position 301;

- (q) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:9 having an A at position 201 but not to a nucleic acid molecule including SEQ ID NO:9 having a C at position 201;
- (r) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:9 having a C at position 201 but not to a nucleic acid molecule including SEQ ID NO:9 having an A at position 201;
- (s) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:10 having an A at position 301 but not to a nucleic acid molecule including SEQ ID NO:10 having a C at position 301;
- (t) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule including SEQ ID NO:10 having a C at position 301 but not to a nucleic acid molecule including SEQ ID NO:10 having an A at position 301;
- (u) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:11 having an A at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:11 having a C at position 301;
- (v) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:11 having a C at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:11 having an A at position 301;
- (w) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:12 having an A at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:12 having a T at position 301;
- (x) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:12 having a T at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:12 having an A at position 301;
- (y) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:13 having a C at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:13 having a G at position 501;
- (z) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:13 having a G at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:13 having a C at position 501;
- (aa) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:14 having a C at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:14 having a T at position 301;

- (bb) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:14 having a T at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:14 having a C at position 301;
- (cc) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having an A at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having a C, G, or T at position 501;
- (dd) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having a C at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having an A, G, or T at position 501;
- (ee) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having an G at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having a C, A, or T at position 501;
- (ff) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having a T at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having an A, G, or C at position 501;
- (gg) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having an G at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having an A at position 501; and
- (hh) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having an A at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having a G at position 501.

In accordance with a further embodiment, there is provided an array of oligonucleotides or peptide nucleic acids attached to a solid support, the array including two or more of the oligonucleotides or peptide nucleic acids described herein.

In accordance with a further embodiment, there is provided a composition including an addressable collection of two or more oligonucleotides or peptide nucleic acids, the two or more oligonucleotides or peptide nucleic acids consisting essentially of two or more nucleic acid molecules set out in SEQ ID NO:1-15 or compliments, fragments, variants, or analogs thereof.

The oligonucleotides or peptide nucleic acids described herein may further include one or more of the following: a detectable label; a quencher; a mobility modifier; a contiguous non-target sequence situated 5' or 3' to the target sequence or 5' and 3' to the target sequence.

The oligonucleotides or peptide nucleic acids may further include one or more of the following: a detectable label; a quencher; a mobility modifier; a contiguous non-target sequence situated 5' or 3' to the target sequence or 5' and 3' to the target sequence.

The oligonucleotides or peptide nucleic acids may alternatively be of about 10 to about 400 nucleotides, about 15 to about 300 nucleotides. The oligonucleotides or peptide nucleic acids may alternatively be of about 20 to about 200 nucleotides, about 25 to about 100 nucleotides. The oligonucleotides or peptide nucleic acids may alternatively be of about 20 to about 80 nucleotides, about 25 to about 50 nucleotides. The genotype may be determined using a nucleic acid sample from the subject. Genotype may be determined using one or more of the following techniques: restriction fragment length analysis; sequencing; micro-sequencing assay; hybridization; invader assay; gene chip hybridization assays; oligonucleotide ligation assay; ligation rolling circle amplification; 5' nuclease assay; polymerase proofreading methods; allele specific PCR; matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectroscopy; ligase chain reaction assay; enzyme-amplified electronic transduction; single base pair extension assay; and reading sequence data. A determination of whether a site is in linkage disequilibrium (LD) with another site may be determined based on an absolute r^2 value or D' value. When evaluating loci for LD those sites within a given population having a high degree of linkage disequilibrium (for example an absolute value for D' of ≥ 0.5 or $r^2 \geq 0.5$) are potentially useful in predicting the identity of an allele of interest (for example associated with the condition of interest). A high degree of linkage disequilibrium may be represented by an absolute value for D' of ≥ 0.6 or $r^2 \geq 0.6$. Alternatively, a higher degree of linkage disequilibrium may be represented by an absolute value for D' of ≥ 0.7 or $r^2 \geq 0.7$ or by an absolute value for D' of ≥ 0.8 or $r^2 \geq 0.8$. Additionally, a high degree of linkage disequilibrium may be represented by an absolute value for D' of ≥ 0.85 or $r^2 \geq 0.85$ or by an absolute value for D' of ≥ 0.9 or $r^2 \geq 0.9$. Two or more oligonucleotides or peptide nucleic acids may include 3 or more; 4 or more; 5 or more; 6 or more; 7 or more; 8 or more; 9 or more; 10 or more; 11 or more; 12 or more; 13 or more; 14 or more; 15 or more; 16 or more; 17 or more; 18 or more; 19 or more; or 20 or more.

Sequence variations may be assigned to a gene if mapped within 2 kb or more of an mRNA sequence feature. In particular, such a sequence may extend many kilobases (kb) from a gene and into neighbouring genes, where the LD within a region is strong.

DETAILED DESCRIPTION

1. Definitions

In the description that follows, a number of terms are used extensively, the following definitions are provided to facilitate understanding of the various embodiments of the invention.

An “anthracycline compound” or “anthracycline” or “anthracycline derivatives” or “anthracycline analogues” as used herein is typically an anthraquinone core attached to a carbohydrate moiety and derivative thereof (see for example, FAN et al. J. Org. Chem. (2007) 72:2917-2928; Goodman and Gilman’s The Pharmacological Basis of Therapeutics 8th edition editors Alfred Goodman Gilman, Theodore Rall, Alan Nies, Palmer Taylor. Pergamon Press. 1990 pg 1241-1244). For example, include anthracycline antibiotics such as daunorubicin (daunomycin, rubidomycin), doxorubicin, idarubicin, epirubicin, mitoxantrone, carminomycin, esorubicin, quelamycin, aclarubicin, esorubicin, zorubicin, pirarubicin, amrubicin, iododoxorubicin, detorubicin, marcellomycin, rodorubicin, and valrubicin. Alternatively, the anthracycline may be selected from daunorubicin and doxorubicin.

As used herein “anthracycline-induced cardiotoxicity” or “ACT” is defined based on CTCAEv3 (Common Terminology Criteria for Adverse Events – see Cancer Therapy Evaluation Program - Common Terminology Criteria for Adverse Events- Version 3 in edition 2003) as early- or late-onset left ventricular dysfunction measured by echocardiogram (shortening fraction, SF) and/or symptoms requiring intervention. We used a more stringent threshold of $SF \leq 26\%$ at any time during or after anthracycline therapy to better differentiate between cardiotoxicity cases and controls. To exclude transient acute cardiotoxicity, echocardiograms obtained <21 days after a dose of anthracyclines were excluded. Control patients were required to have normal echocardiograms ($SF \geq 30\%$) during and after therapy, with a follow-up of >5 years after completion of anthracycline therapy. Doxorubicin equivalents were used to calculate cumulative anthracycline doses (Altman A.J. editor, Children's Oncology Group. Supportive care of children with cancer: current therapy and guidelines from the Children's Oncology Group. Baltimore: Johns Hopkins University Press; 2004. 412 p.p.).

“Genetic material” includes any nucleic acid and can be a deoxyribonucleotide or ribonucleotide polymer in either single or double-stranded form.

A nucleotide represented by the symbol M may be either an A or C, a nucleotide represented by the symbol W may be either an T/U or A, a nucleotide represented by the symbol Y may be either an C or T/U, a nucleotide represented by the symbol S may be either an G or C, while a nucleotide represented by

the symbol R may be either an G or A, and a nucleotide represented by the symbol K may be either an G or T/U. Similarly, a nucleotide represented by the symbol V may be either A or G or C, while a nucleotide represented by the symbol D may be either A or G or T, while a nucleotide represented by the symbol B may be either G or C or T, and a nucleotide represented by the symbol H may be either A or C or T. A nucleotide represented by the symbol N may be an A or G or T or C.

A “polymorphic site” or “polymorphism site” or “polymorphism” or “single nucleotide polymorphism site” (SNP site) or single nucleotide polymorphism” (SNP) as used herein is the locus or position within a given sequence at which divergence occurs. A “polymorphism” is the occurrence of two or more forms of a gene or position within a gene (allele), in a population, in such frequencies that the presence of the rarest of the forms cannot be explained by mutation alone. The implication is that polymorphic alleles confer some selective advantage on the host. Polymorphic sites have at least two alleles, each occurring at frequency of greater than 1%, and may be greater than 10% or 20% of a selected population.

Polymorphic sites may be at known positions within a nucleic acid sequence or may be determined to exist. Polymorphisms may occur in both the coding regions and the noncoding regions (for example, promoters, introns or untranslated regions) of genes. Polymorphisms may occur at a single nucleotide site (SNPs) or may involve an insertion or deletion as described herein.

A “risk genotype” as used herein refers to an allelic variant (genotype) at one or more of the following polymorphic sites rs7853758; rs885004; rs17863783; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or a polymorphic site in linkage disequilibrium thereto, for the subject as described herein, as being indicative of an increased likelihood of cardiotoxicity following administration of an anthracycline. The risk genotype may be determined for either the haploid genotype or diploid genotype, provided that at least one copy of a risk allele is present. Risk genotype may be an indication of an increased risk of cardiotoxicity. Subjects having one copy (heterozygotes) or two copies (homozygotes) of the risk allele are considered to have the “risk genotype” even though the degree to which the subjects is at risk cardiotoxicity may increase, depending on whether the subject is a homozygote rather than a heterozygote. Such “risk alleles” or “risk genotypes” may be selected from the following: rs7853758G; rs885004G; rs17863783A (reverse); rs17863783T (forward); rs10426377C; rs2305364A; rs4982753G (reverse); rs4982753C (forward); rs4149178A; rs4148808A; rs1149222G; rs17583889A; rs4877847A; rs11625724T; rs12882406G; rs12896494T; and rs10426628A; or a polymorphic site in linkage disequilibrium thereto.

A “decreased risk genotype” as used herein refers to an allelic variant (genotype) at one or more of the following polymorphic sites: rs7853758; rs885004; rs17863783; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or a polymorphic site in linkage disequilibrium thereto, for the subject as described herein, as being indicative of a decreased likelihood of cardiotoxicity following administration of an anthracycline. “Decreased risk alleles” or “decreased risk genotypes” or “reduced risk genotypes” may be selected from the following: rs7853758A; rs885004A; rs17863783C (reverse); rs17863783G (forward); rs10426377A; rs2305364G; rs4982753A (reverse); rs4982753T (forward); rs4149178G; rs4148808G; rs1149222A (reverse); rs1149222T (forward); rs17583889C; rs4149178G; rs4148808G; rs1149222T; rs17583889C; rs4877847C; rs11625724A; rs12882406C; rs12896494C; and rs10426628G or a polymorphic site in linkage disequilibrium thereto (decreased risk alleles on the forward strand).

A “clade” is a group of haplotypes that are closely related phylogenetically. For example, if haplotypes are displayed on a phylogenetic (evolutionary) tree a clade includes all haplotypes contained within the same branch.

The pattern of a set of markers along a chromosome is referred to as a “Haplotype”. Accordingly, groups of alleles on the same small chromosomal segment tend to be transmitted together. Haplotypes along a given segment of a chromosome are generally transmitted to progeny together unless there has been a recombination event. Absence of a recombination event, haplotypes can be treated as alleles at a single highly polymorphic locus for mapping.

As used herein “haplotype” is a set of alleles of closely linked loci on a chromosome that tend to be inherited together. Such allele sets occur in patterns, which are called haplotypes. Accordingly, a specific SNP or other polymorphism allele at one SNP site is often associated with a specific SNP or other polymorphism allele at a nearby second SNP site or other polymorphism site. When this occurs, the two SNPs or other polymorphisms are said to be in Linkage Disequilibrium (LD) because the two SNPs or other polymorphisms are not just randomly associated (i.e. in linkage equilibrium).

In general, the detection of nucleic acids in a sample depends on the technique of specific nucleic acid hybridization in which the oligonucleotide is annealed under conditions of “high stringency” to nucleic acids in the sample, and the successfully annealed oligonucleotides are subsequently detected (see for example Spiegelman, S., Scientific American, Vol. 210, p. 48 (1964)). Hybridization under high stringency conditions primarily depends on the method used for hybridization, the oligonucleotide length,

base composition and position of mismatches (if any). High-stringency hybridization is relied upon for the success of numerous techniques routinely performed by molecular biologists, such as high-stringency PCR, DNA sequencing, single strand conformational polymorphism analysis, and in situ hybridization. In contrast to Northern and Southern hybridizations, these aforementioned techniques are usually performed with relatively short probes (e.g., usually about 16 nucleotides or longer for PCR or sequencing and about 40 nucleotides or longer for in situ hybridization). The high stringency conditions used in these techniques are well known to those skilled in the art of molecular biology, and examples of them can be found, for example, in Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y., 1998.

“Oligonucleotides” as used herein are variable length nucleic acids, which may be useful as probes, primers and in the manufacture of microarrays (arrays) for the detection and/or amplification of specific nucleic acids. Such DNA or RNA strands may be synthesized by the sequential addition (5'-3' or 3'-5') of activated monomers to a growing chain, which may be linked to an insoluble support. Numerous methods are known in the art for synthesizing oligonucleotides for subsequent individual use or as a part of the insoluble support, for example in arrays (BERNFELD MR. and ROTTMAN FM. J. Biol. Chem. (1967) 242(18):4134-43; SULSTON J. *et al.* PNAS (1968) 60(2):409-415; GILLAM S. *et al.* Nucleic Acid Res.(1975) 2(5):613-624; BONORA GM. *et al.* Nucleic Acid Res.(1990) 18(11):3155-9; LASHKARI DA. *et al.* Proc Nat Acad Sci (1995) 92(17):7912-5; MCGALL G. *et al.* PNAS (1996) 93(24):13555-60; ALBERT TJ. *et al.* Nucleic Acid Res.(2003) 31(7):e35; GAO X. *et al.* Biopolymers (2004) 73(5):579-96; and MOORCROFT MJ. *et al.* Nucleic Acid Res.(2005) 33(8):e75). In general, oligonucleotides are synthesized through the stepwise addition of activated and protected monomers under a variety of conditions depending on the method being used. Subsequently, specific protecting groups may be removed to allow for further elongation and subsequently and once synthesis is complete all the protecting groups may be removed and the oligonucleotides removed from their solid supports for purification of the complete chains if so desired.

“Peptide nucleic acids” (PNA) as used herein refer to modified nucleic acids in which the sugar phosphate skeleton of a nucleic acid has been converted to an N-(2-aminoethyl)-glycine skeleton. Although the sugar-phosphate skeletons of DNA/RNA are subjected to a negative charge under neutral conditions resulting in electrostatic repulsion between complementary chains, the backbone structure of PNA does not inherently have a charge. Therefore, there is no electrostatic repulsion. Consequently, PNA has a higher ability to form double strands as compared with conventional nucleic acids, and has a high ability to recognize base sequences. Furthermore, PNAs are generally more robust than nucleic acids. PNAs

may also be used in arrays and in other hybridization or other reactions as described above and herein for oligonucleotides.

An “addressable collection” as used herein is a combination of nucleic acid molecules or peptide nucleic acids capable of being detected by, for example, the use of hybridization techniques or by any other means of detection known to those of ordinary skill in the art. A DNA microarray would be considered an example of an “addressable collection”.

In general the term “linkage”, as used in population genetics, refers to the co-inheritance of two or more nonallelic genes or sequences due to the close proximity of the loci on the same chromosome, whereby after meiosis they remain associated more often than the 50% expected for unlinked genes. However, during meiosis, a physical crossing between individual chromatids may result in recombination.

"Recombination" generally occurs between large segments of DNA, whereby contiguous stretches of DNA and genes are likely to be moved together in the recombination event (crossover). Conversely, regions of the DNA that are far apart on a given chromosome are more likely to become separated during the process of crossing-over than regions of the DNA that are close together. Polymorphic molecular markers, like SNPs, are often useful in tracking meiotic recombination events as positional markers on chromosomes.

Furthermore, the preferential occurrence of a disease gene in association with specific alleles of linked markers, such as SNPs or other polymorphisms, is called “Linkage Disequilibrium” (LD). This sort of disequilibrium generally implies that most of the disease chromosomes carry the same mutation and the markers being tested are relatively close to the disease gene(s).

For example, in SNP-based association analysis and LD mapping, SNPs can be useful in association studies for identifying polymorphisms, associated with a pathological condition, such as sepsis. Unlike linkage studies, association studies may be conducted within the general population and are not limited to studies performed on related individuals in affected families. In a SNP association study the frequency of a given allele (i.e. SNP allele) is determined in numerous subjects having the condition of interest and in an appropriate control group. Significant associations between particular SNPs or SNP haplotypes and phenotypic characteristics may then be determined by numerous statistical methods known in the art.

Association analysis can either be direct or LD based. In direct association analysis, potentially causative SNPs may be tested as candidates for the pathogenic sequence. In LD based SNP association analysis,

SNPs may be chosen at random over a large genomic region or even genome wide, to be tested for SNPs in LD with a pathogenic sequence or pathogenic SNP. Alternatively, candidate sequences associated with a condition of interest may be targeted for SNP identification and association analysis. Such candidate sequences usually are implicated in the pathogenesis of the condition of interest. In identifying SNPs associated with cardiotoxicity, candidate sequences may be selected from those already implicated in the pathway of the condition or disease of interest. Once identified, SNPs found in or associated with such sequences, may then be tested for statistical association with an individual's prognosis or susceptibility to the condition or to the side effect of a medication.

For an LD based association analysis, high density SNP maps are useful in positioning random SNPs relative to an unknown pathogenic locus. Furthermore, SNPs tend to occur with great frequency and are often spaced uniformly throughout the genome. Accordingly, SNPs as compared with other types of polymorphisms are more likely to be found in close proximity to a genetic locus of interest. SNPs are also mutationally more stable than variable number tandem repeats (VNTRs) and short tandem repeats (STRs).

In population genetics linkage disequilibrium refers to the “preferential association of a particular allele, for example, a mutant allele for a disease with a specific allele at a nearby locus more frequently than expected by chance” and implies that alleles at separate loci are inherited as a single unit (Gelehrter, T.D., Collins, F.S. (1990). Principles of Medical Genetics. Baltimore: Williams & Wilkins). Accordingly, the alleles at these loci and the haplotypes constructed from their various combinations serve as useful markers of phenotypic variation due to their ability to mark clinically relevant variability at a particular position (see Akey, J. *et al.* Eur J Hum Genet (2001) 9:291-300; and Zhang, K. *et al.* (2002). Am J Hum Genet. 71:1386-1394). This viewpoint is further substantiated by Khoury *et al.* ((1993). Fundamentals of Genetic Epidemiology. New York: Oxford University Press at p. 160) who state, “[w]henver the marker allele is closely linked to the true susceptibility allele and is in [linkage] disequilibrium with it, one can consider that the marker allele can serve as a proxy for the underlying susceptibility allele.”

As used herein “linkage disequilibrium” (LD) is the occurrence in a population of certain combinations of linked alleles in greater proportion than expected from the allele frequencies at the loci. For example, the preferential occurrence of a disease gene in association with specific alleles of linked markers, such as SNPs, or between specific alleles of linked markers, are considered to be in LD. This sort of disequilibrium generally implies that most of the disease chromosomes carry the same mutation and that the markers being tested are relatively close to the disease gene(s). Accordingly, if the genotype of a first

locus is in LD with a second locus (or third locus etc.), the determination of the allele at only one locus would necessarily provide the identity of the allele at the other locus. When evaluating loci for LD those sites within a given population having a high degree of linkage disequilibrium (i.e. an absolute value for $r^2 \geq 0.5$) are potentially useful in predicting the identity of an allele of interest (i.e. associated with the condition of interest). A high degree of linkage disequilibrium may be represented by an absolute value for $r^2 \geq 0.6$. Alternatively, a high degree of linkage disequilibrium may be represented by an absolute value for $r^2 \geq 0.7$ or by an absolute value for $r^2 \geq 0.8$. Additionally, a high degree of linkage disequilibrium may be represented by an absolute value for $r^2 \geq 0.85$ or by an absolute value for $r^2 \geq 0.9$ or by an absolute value for $r^2 \geq 0.95$. Accordingly, two SNPs that have a high degree of LD may be equally useful in determining the identity of the allele of interest or disease allele. Therefore, we may assume that knowing the identity of the allele at one SNP may be representative of the allele identity at another SNP in LD. Accordingly, the determination of the genotype of a single locus can provide the identity of the genotype of any locus in LD therewith and the higher the degree of linkage disequilibrium the more likely that two SNPs may be used interchangeably. For example, SLC28A3 rs885004 showed a high degree of LD (i.e. $r^2=0.83$) with rs7853758.

LD may be useful for genotype-phenotype association studies. For example, if a specific allele at one SNP site (e.g. "A") is the cause of a specific clinical outcome (e.g. call this clinical outcome "B") in a genetic association study then, by mathematical inference, any SNP (e.g. "C") which is in significant LD with the first SNP, will show some degree of association with the clinical outcome. That is, if A is associated (~) with B, i.e. $A \sim B$ and $C \sim A$ then it follows that $C \sim B$. Of course, the SNP that will be most closely associated with the specific clinical outcome, B, is the causal SNP – the genetic variation that is mechanistically responsible for the clinical outcome. Thus, the degree of association between any SNP, C, and clinical outcome will depend on LD between A and C.

Until the mechanism underlying the genetic contribution to a specific clinical outcome is fully understood, LD helps identify potential candidate causal SNPs and also helps identify a range of SNPs that may be clinically useful for prognosis of clinical outcome or of treatment effect. If one SNP within a gene is found to be associated with a specific clinical outcome, then other SNPs in LD will also have some degree of association and therefore some degree of prognostic usefulness.

Polymorphisms in linkage disequilibrium may be identified, for example, using the Haploview program (BARRETT JC. et al. Bioinformatics (2005) 21(2):263-5 (<http://www.broad.mit.edu/mpg/haploview/>)) and the LD function in the Genetics Package in R (R Core Development Group, 2005 - R Development

Core Team (www.R-project.org). Linkage Disequilibrium between markers may be defined using r^2 whereby all SNPs available on Hapmap.org (phase II) (cohort H), all SNPs genotyped internally using the Illumina Goldengate assay (cohort I) and SNPs may be sequenced using the Sequenom Iplex Platform (cohort S) for genes of interest. A minimum r^2 of 0.5 may be used as the cutoff to identify LD SNPs.

Although this study had moderate to high statistical power to replicate the previous associations, some variants were not replicated. While this might suggest that these associations were false-positive, other explanations exist. For example, the current studies were powered to find similar effect sizes, but often effects may be smaller in replication studies (Chanock, SJ. *et al.* Nature (2007) 447:655-660). Despite trying to keep the replication cohorts similar to the original cohorts, small differences might exist, for example in ethnicity, (supportive) treatment or follow-up, which could potentially lead to non-replication due to different effects of the variants in specific populations or subgroups (Chanock, SJ. *et al.* Nature (2007) 447:655-660).

In the present analyses correction were made for the effects of several important clinical risk factors. Not unexpectedly, cumulative doses showed higher statistical significance in cases compared to controls in both the Dutch-EKZ and CPNDS cohort. Age at start of treatment was higher in CPNDS cases, whereas younger age was usually considered a risk factor (Kremer, LC. *et al.* Ann Oncol (2002) 13:503-512). This is likely in part due to the requirement in the present studies to have controls with at least a 5 year follow-up, selecting for relatively younger controls in the CPNDS cohort.

Numerous sites have been identified as polymorphic sites associated cardiotoxicity following anthracycline administration (see TABLE 1).

TABLE 1. Single Nucleotide Polymorphisms that Showed an Association with Anthracycline-Induced Cardiotoxicity (ACT)

Gene Symbol	SNP ID	Chromosome	SNP Alleles (* reverse strand)	ACT Predictive Variant
SLC28A3	rs7853758	9	A/G	G
SLC28A3	rs885004	9	A/G	G
UGT1A6	rs17863783	2	A/C* or G/T (forward strand)	A* or T
SULT2B1	rs10426377	19	A/C	C
SLC28A1	rs2305364	6	A/G	A
SLC22A17	rs4982753	14	A/G* or C/T (forward strand)	G* or C
SLC22A7	rs4149178	6	G/A	A
ABCB4	rs4148808	7	A/G	A

ABCB4	rs1149222	7	C/A* or G/T (forward strand)	C* or G
HNMT	rs17583889	2	A/C	A
SLC28A3	rs4877847	9	A/C	A
SLC22A17	rs11625724	14	A/T	T
SLC22A17	rs12882406	14	C/G	G
SLC22A17	rs12896494	14	C/T	T
SULT2B1	rs10426628	19	A/C/G/T (where A and G are most common)	A

TABLE 2. below shows the flanking sequences for the SNPs shown in TABLE 1 providing their rs designations and corresponding SEQ ID NO designations. Each polymorphism is identified by its position within the flanking sequence and is in bold.

TABLE 2. Sequence for Cardiotoxicity-Associated Polymorphisms Listed in TABLE 1, with SEQ ID NO designations

Gene Symbol	SNP ID	SNP Alleles (* reverse strand)	GENOMIC SEQUENCE	SEQ ID NO:
SLC28A3	rs7853758 (at position 301)	A/G	GAAGGAAGCTGGATTCTTGGGGAAGGGGCCAGGAGAGACTGACT GTCGTTGTGGGGAAGGAGGGGGATAGGAGACAGAGAAAAGGTGG GTGGGAAGTTGGGGGATGCATGAGAAGCTTCTACGGTGTGGAAG AGTCTACTGAGGTTAGGGTGGGCTGTTACAAACCTATTTTATT TTTAAACAAAGATAGGCAGAAACAAAACAGAGGGCAGGGGCGTG ATGTGATTATACCTCAAACTCAGCTGTGGGTAGTCAAACATGT TTCCAAACCAGGACAGGGCTGAATTCATAAAAGACA R CAGGGCCAGGAAGGCAATCAGATTCACAGCGATGTTGGCCACCA GGGAGATGGAGGAGGATGCTCCCTGTGTTGCAGCTTCTAGAAGA TTCCCTGAATCACTTTATCAAGAAATAGCAATTCCAGAATTACC AAGGAGTTGTCAGGGGATGGACACCATTGGTGCAGAAGTAGCAT AATCAGAGCTTAGATAAGAGACAAACTCTAAGGCCCTGCCTTG CTCAAATACTCACCTGAATCCCATGACATCATGAAGTGGGCAT ATGAGTATTTAACCATGTTTGTCTTGTTTTAAAG	1

SLC28A3	rs885004 (at position 201)	A/G	CTTTTATTCTAATTTGTCACAGGACTTTGACTTTTTTCCCCTGT CTTCTATGGTGAGCTTTTTTTTGCTCAAAGAAGGAATGAAGTTTG GGGGGATTTGTAATACCTAGATCAATACACTGAAGGCTGTCAGA AAATATTCTTGTCTGCTGCTAAAAAGACATGGGCTTCACCACT AGTTAGCAGAGTAACCTAAGTTGA R ATAAATTGTGAGCAACCTCAGAGCCTTATTCTGAGTAGCTGTAG AGATCAGGTAAGGCATGGTAGGTGCTGACACATAGTGCCCTCCTA Ctggatggcagacacacctgacagcaataacttcagcatgtcct gagaatgaccctatggtctaagcgtgtgtgtttggagttccaaa ctaaggaacctgggagcggccattccaaagattcattcccttacc tataaggaatatctgaaccttggccattccgtggaaagcagg ccatatgggggatccaggccctttgttttgggttaaatgaaggt tgggaggtggaggttgcctgggggaggttactaagtgagaatgct atataaactgcctgccttttacaacagtagcggttctcctctc tagcctgctgccactggactgccctgcatgtgagtcacctcaaa taaaccctatgtctcatttgcctggctccaggtcttttctttggc cgctcaaacatggtgccccccctactgaagtcattaggggtctg tcatgacCTTATGCAGTAAACATTCTTCTAGCTTTTAAAA TAATAGTGATTTTACCTATAGATTTTCAAATTACTAAAAAGTGA CTGCATTTTGGGAAGCTACTTTAGTTCTGTG	2
UGT1A6	rs17863783 (at position 301)	A/C* or G/T (forward strand)	GACAGAGTACAGGAATAACATGATTGTTATTGGCCTGTACTTCA TCAACTGCCAGAGCCTCCTGCAGGACAGGGACACCCTGAACCTC TTTAAGGAGAGCAAGTTTGATGCTCTTTTCACAGACCCAGCCTT ACCCTGTGGGGTGATCCTGGCTGAGTATTTGGGCCTACCATCTG TGTACCTCTTCAGGGGTTTTCCGTGTTCCCTGGAGCATACTTC AGCAGAAGCCCAGACCCTGTGTCCTACATTCCCAGGTGCTACAC AAAGTTTTTCAGACCACATGACTTTTTTCCCAACGAGT K GCCAACTTCCTTGTTAATTTGTTGGAGCCCTATCTATTTTATTG TCTGTTTTCAAAGTATGAAGAACTCGCATCAGCTGTCCCTCAAGA GAGATGTGGATATAATCACCTTATATCAGAAGGTCTCTGTTTGG CTGTTAAGATATGACTTTGTGCTTGAATATCCTAGGCCGGTCAT GCCCCAATGGTCTTCATTGGAGGTATCAACTGTAAGAAGAGGA AAGACTTGTCTCAGGTTGGTGGGTTTATTTCTTTTGGACTGCCT TGTTTCTTCCAGGCTCTGTCTCCCTCACTCATTTG (TAACAAGGAA GTTGGC M ACTCGTTGGG AAAAAG)	3

SULT2B1	rs10426377 (at position 501)	A/C	actccagcctgggcaacaaaagcaaaactccgtctcaaaaaaaaa aaCCCCATatattatactacatattataataatttaatagtat tacaatatactttaattataatttaatatatttatatgtttaacata taataacataatgcaatataatataatacaTAAAATGTCTGATG GGATAAATTTACATTGGGGGGAGAGGGCTGGAGATAGAGTTAGA GGTGGGTGTGGGGCTGCAATTATAAATAAGGAGGTCAAGGGGGC CCCCAGtgagtaaaaacccgggaggtgatgggtgggagccacaga ggtttctagaggaagagcatttcaggcaagagggaaaagcaggt gcaaaggccctgaggtgggtgtatctgaggtgcaagagggaggt cgggtgtggccggagccaagtgagcaagtggggaaaggagtgag atgaggtccgggtggggaggcaacagggggccaaatgtgcaggg ccgcgtgggcccgggtg M ggacttgagctctgaccgagtgacgaggcagcgccggcagggatc gcagcagaggaggagcccagcgggcttaggcttcaccggcgcc ctctggcggtcagatggggaccgactgtcgagagagcagaagcc gggaagcccagagggcgccgctggccgggggtccaggggagggga cgggtggctggactaggggtgatgggtcacggaggtattttgaaggt gagaccgggaggatttctgtgacagactggatgtgggtgtgagag aaggggatgagtcagggtgactccaaggtttcggcggaagcaa ctgtcaggggtggggcaactgggaaggggaggaggaggggagca ggcgggggaggaagaCACGGGCTCTGCGGCGGCCATCCCACGT CCAGCAGAGGCTGCCGGTGGGCACAGAAATGCACCCATCGGGAT CTGAGGGGAGAGACTTGGGCTGACACGTCCATTTGCAATTCCTT GGCACATATATTCTGT	4
SLC28A1	rs2305364 (at position 251)	A/G	AGGGAGCTATTGCAGCGGGACTAGTAAGGGCTGGCTTCCTGGCA GAGGTGGCCTGTGAGGGGGCCCTGAAGGAGGAGCGCCTGAGAGGA TCATTTGAGCCTAAGCCTTCCTGGAAACTGCTCTTTCTGGGCCC AGAGTCTCTGCAGCACATCAGGGACTAAGGTCTCAGCAGGAGCT GTCCCCATGACCCCTACCCCACTTCTTGTCTCAGCTACTCAG TAATTCTATAATTTGGAGGTTTCATCTCTAA R AAAAAACATACAGAAATATCTTACTTTTGCAAATTTTACAAAA TATAGCAAAAACATCTCATGAGGTGAACATTGCTGGGCTCCTCC CGGGCCCCAGGCTTAAGGTTCACTAACTTCCCAGTACGGCGACC TCTGCCCCACAGGACTCCTCCCCCTCCACCTGGGCCCCCACCATGC CCAGACTCCCCTCTAGCTGGAGTCGCTGGGCGGTGGCTGCTAAA GGCCTCTGGGACCCTGAGTCAAGCTGACCT	5

SLC22A17	rs4982753 (at position 501)	A/G* or C/T (forward strand)	TTCTctgaagccttcatcttcatgataaaaacctaggtctccaaa accccttatcttaacccaaacattccttttctactgataattact ctttcaaccaattgccaatcagaatatgtttaaatctaactacg gcctggaagcccctggccctgcctttgagttgtcccggtcttc cagatccaaccattgtaaatcctgcacgtactgattgatgtatt acatctccctaaaatgtacaaagcaagctgtacttcgaccactt tgggcacgtgtcctcaggacctcctgaggctgtatcatgggtgc atccttaaccttggcaaaaataaacttttctaaattgattgagact tgtctcagataatttttgggtttacaCAGCTTACACCTAGTGGGG GTAGTTTAAGAGACACTTAGGATACATTACCTGTTTAATCCTC ATAGCAAGTCTTAAAGGAAGTTGGAAGCTAAGAGATTTGCCCAA AGTGGTAAAGCCAAAC Y AGATGTGGTTTAAGAGGGAGGCCCTGGCACTGAGGTAGAAGTCT CAGTTTTCTGCTGCCACCTTCCAGAGAGCCAAGAGAACCCACTC TTGCCTTCCAAGGAATGTGCCCAAAAGTTAGAAGAGTTGGGAAG GTGACTACAGTATGCATCCATGAAGAGctctctgggatctaatac atgtctcagctttgccaactcattagctttgtaaccttgggccag ttactgagtcctcttgaccataagttttctcatcttgtttaaga gaataagaatGCTGTACTGAGCTTCCATTGTTTTTCTGAGTG ACACCCAGTCCCCTTCTGCTAACAGTGCCCAATTTTG	6
SLC22A7	rs4149178 (at position 401)	G/A	GGGGCTGACTGCACTGGTGGGCCGGCTGGGGGGCTCTTTGGCCC CACTGGCGGCCTTGCTGGATGGAGTGTGGCTGTCACTGCCCAAG CTTACTTATGGGGGGATCGCCCTGCTGGCTGCCGGCACC GCCCT CCTGCTGCCAGAGACGAGGCAGGCACAGCTGCCAGAGACCATCC AGGACGTGGAGAGAAAGAGGTGTGTGCACAGGACTGTGTCTGTG TACGTGTGATAACATGCATATGGATGCAGGTGCTCAGATACCTG ACACCTTAGGATTTCAGACAGGAACTATATCTGCACATGTGTAC TTGGTGCATGTATGTGTGCGTGCACAGGTGAGTGTTCAGGGTG CATGATTGGGGGCATGTAAACTGCAAGCATAAATGTGTACATGA GTGT R TGAACACAAATACATAAGGTTGTTTCAGGAAGATGTGTATGTGTG CACACTCAAGTATGCCTACACATCAGGGGTGTGTACAAGCATGT GTGTGAGTCACCTGTACAGGCGTTTGTATACCAGGAAGTGGGGT GAGCCCTGGGACAGCCCTGATGGGGCAGGCTGGGGTGGTAGGCT CGGGGATTAAACCCACCATTGCTCACAACCTCCTTCCCTCCCTAG TGCCCCAACCACTCTTCAGGAGGAAGAGATGCCCATGAAGCAGG TCCAGAACTAAGTGGGAGTGGAGGCAGGCCCTCCACAGAAGCTC TGCAGCAGGGGCTGGGAGAGCAGAAGGGCAGGCCCTGCAACTCA GGCTGGGAGTATCGAACCTCTGCCTAGGGCCGGAGTTGCTGCC AGTA	7
ABCB4	rs4148808 (at position 301)	A/G	TTACCTAGTACTTAATTCCTTATCCTCAACTCATTTATTCCTC GCAATAACCTGATAAGGGAGATGTTTTATCCTCATTTTACATA TAAGGAAACAGGCCTAGAGAAATGAGCACAGTGTCCAAAGTCAC ATAGTTAATAAGATGTGAAGCTCTGAGTTTGAAAGTCTCCGGTT TCAAAGCCATGAACTTATGGCTCCCYGTTTTAGACACTTCCTT TTGGGAAGAGTGTGGAGGAATTAATCAGAAAGAAGAAAGTCATA CTCAAATAGGTGGTAGGAGYAGAGACAATTCAATAC R GACAGAAGTCTTAGATGAGAGCAGTGAGCCAGGGCACTGGACTG GGACTCAGGAGGCTTCCCCTAGACTCTGGTTCCACMGATGCAGC CTCAGGCAGGACTTCACCTCTCTGGGCATCCGTTTCTTCATATG TTAAACATACGGGGTTTTAATTAGATGATCGCTGAAGACCCCTC TAGCCCTAAAACCTCTGTGTCTCTTAAGTGCTAAGAGGGCACCAA CAGCGTTCCCTCCTCCCCAAGGAGCATAATGTGATGGTTCCCTGCC GGCCCTGGCTGACTCTCGCCGTCCTTGGAGATAATT	8

ABCB4	rs1149222 (at position 201* AND 61)	C/A* or G/T (forward strand)	AGCACTTTTCTCAGACTTGGATCAATGAAAAAGGAAATTGCTT AGATTAAATGATCAAATTTCACTTAGCTACCTTGCAATATTTGA ATGCAGTGACTTACTTTTTTACGTCAAAACTTTTTCTTTTTCAC TGTAGTCAGTTTATTTAATGTAGTTTATTTAAAATCATACCAA CATGCAGTGATAATTTTTTGGAGA K GGGGAACCACAGCAAAATATACTGTCAATTTTGGTCCCACCGAT AATTTATCTACCCCCCATCCCATATAGAAATCATAAGATTCTCA GTCTGGCTACAACATGAAAGTGAAAAAGATAAAAACTAGGAAGA GAAACGTCTTTGCTCTGCACTAGTGCAGATGTCCTTTTCCAAAA TGAGATCAGtattattatattatctaaagatacatttaaaaaat attaataaGAGGTATTTTCTTTTCACTTATTACTAACAGGTCATT CATTTATATAAATTAAATGTCACTTTAGTCAGTACAACCTTATTC AATGTAGTTGATTCAAAAAAT (GGGGGGTAGATAAATTATCGGTGGGACCAAAATTGACAGTATA TTTTGCTGTGGTTCCCC M TCTCCAAAAAATTATCACTGCATGTTTGGTATGATTTTAAATAA ACTACATTAAATAAAC)	9
HNMT	rs17583889 (at position 301)	A/C	GCTTGACCTTCATCCATTTTATAGTCACTAGAAGATTTTAAAGCA GGGAAGTGATAGGTTTGCTTAGCAGGGTTTAAAATGGTTATCAT GGTAAGTGCTTTGAGGAATACAAGACCAGAGGTAGGCTGACCTT TTAGGATATTATTGCAGTATTCCAAATGTGAATAATCTTGACCT GAATAAGAGCACAGCTAGTAGGGACAGAATTGAGAAAGTAAACT TCTTGATGGCCAGGAAAGGAAGATAAAGGAAGAAGAACTTTGA GTGATGATGGATAGCAAACAGCATGGGTGAATGTTT M AAGAAAAGATAAATACTTTAAACTAAATGTTAAATGGTTAAACA AGACTATAATATGACAGATCTGAACTATTGACCCCCCTTTTTTT AAACAGGCAAGAGGATTATATACATCTTTTTGTTGGCCGGACGT GGTTACTCAGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGTA GGCAGATCACTTGAGGTCAAGAGTTCGAGACCAGCCTGGCCAAT ATGGTGAAACCCCGTCTCTACTAAAAATATGAAAATTAGCCTGG CATGGTGGTACGTGTCTGTAATCCCAGCTACTTGGG	10
SLC28A3	rs4877847 (at position 301)	A/C	CTCCTTCCCAAAAGCGAATTATTGTCATGTGTTTTTGTGTGTGT GTGTTTTAATTGAATATAGGAAGCTCCAGCCGAGCGCCTCTTGA GCAATGTGAAAGAGTACTCAATGTGATTCATATTGACTTCCCAG GCTAGTGGGTTAAGGAGAAACAATATTCTTCTTCCCTACATAAC ATAGTTACTTCTTACCTTAGAACTTTTGTATTCAAAGTACGTGC TTGACAAGAAGTCTACACCTGTACTCCAGCAGGTTTGGATCCCT GATGGACTTACACTTACATGATCTCCCAAAATTGTAA M CTCCATTTAACAACCTCTCCAAACCTGCTTGACTGGCTTACTCA TGGAGTAACTCAGCCTTTCTGTGTGTTGTGCAGAGTGAGGCC CATTTCCCCAGCCTAGTCCTCTCCAAGACTTTCGTAAGTGAGT TCCACCCAATAACTTCAGAGCAAGCTTGGGTTCTGTGCCTCTG AGAGCAGCTGTGGCAAGATTGATTGCTTATTCACAGGGAGGAGG AAACTTGAGTTGGGATTCCAAGAGATGCCTTTTAAATAAATGTGT TAAAAAACATTCTGATTTTCTCTTAGAAATTCT	11

SLC22A17	rs11625724 (at position 301)	A/T	TGGGAAGGTGACTACAGTATGCATCCATGAAGAGCTCTCTGGGA TCTAATCATGTCTCAGCTTTGCCACTCATTAGCTTTGTAACCTT GGGCCAGTTACTGAGTCTCTCTGACCATAAGTTTTCTCATCTTG TTAAGAGAATAAGAAATGCTGTACTGAGCTTCCATTGTTTTTTT CTGAGTGACACCCAGTCCCGCTTCTGCTAACAGTGCCCAAATTT TGCTTTACCTATTCCCAGTTGGGAGAAATAGGGGTGTGGGAGGC ATGTTACTCTAGCTAGCCAGTGACATGCTCTCTGGA W ATTAATCTTCAGTGGGATAACCCAAAGATTAAATTGCACTGGAG CTCATTCTCTAGAAAGAGCCCTGACAAAAATGTTGCTGGTTTT TGCTCCTCAGACAAGGAGCCTTCTCCGCTCCTGACTAGTTATTT TACGACGAGTTCTGCTTCTCCTTCATTCTGTGAGCTCCCACCAC AACCTGCCCCACCAAATCGAATAATTCCCTTTTTTGCTTAAGTTC ATTAGAGCCAGTTTTTATTGCTTACAACAACACAACAACAACA CAACAACAACAACCCCTAATTGATCCAAACACCTAC	12
SLC22A17	rs12882406 (at position 501)	C/G	AAGCTGGAGTGCCAGTCAAACAAATGCGCCACTTTATAGCCCTT TGTGGAGCATTTATTCGTGCTGATGGCAAAAGACTGTGGGCACT TCCCAATGACAGTGACTGGGCTAGAGAATTTCCACTTGAGGGGC ATTTACTGCCTTGCTGTGGAATGTTGACAGAAGCTATCTTATGC CAATGGAAATAATGGTGCCCAAAGGAGCACCATGACAAAATAAA AATGGTTTACATAAGATCTTGCTACCTGAGGATACAAGGAGGAG ATCCTTATGAGCAGGAAGCCTCTTTTTTTAAAAATTTTTTCCCT AGGACTGATTCTAACTATGTGAGGAGCTACTAGATTCTACAGTG CCTGATAGACAACCTCTCATTAGCTGTTTGGCTTGTGAATGGCAG TTCCAAGGTGAACAAACATCCTGTTTGAAAGCCCTGTTCTGGC TATAGAAGTGTCAAACAAATCCTTTTTCTTTGAGTTATTCACAG TTTAAAGCAACTGGGT S AAGTATGTGTTTGTAAAGCAAATTTGCCTTTCTATCTATCTGAGT TCTCCAAAATTTGGATTGTGAGTTCATGACAATACAGTCATTTG CATAAGCTTAGTAAGAGTCTTTTAGAACAGAGCAATTGGAGACG CTGGTTATTTTGCCAAGGCTTTGACTAGAATAACATATTTGTAG GTAAAGTTCCAGCAAAGCCAACTTAATAGAAGCCTATATGGCCA AACAATTCCTGCTGCACTTTATATGAATAATCAGGCAAAGTATA ATAAGCCTAAAATTTATTTTGTACACAAATTTAGCCCTGGCTT TTGTTTTTGTAGTGCAGATCGAAGAACGAATTATTTCTTGCTAC AATAATCCTCTGAAGAGTATCATTATAATTTTCTTCATATG TTTAGTCGGTACTCTAATAGAATAGCTCCCTTTTTCTGTTCTGA CATACAAATATTCTTTTGATTGTCAAATATTTATGTTATTTAT CTGTCCTTGTTTTATT	13
SLC22A17	rs12896494 (at position 301)	C/T	ACTGTGGGGAAGATAGGATGCTCACAATCCCACAGGCCAGCAG GACCAGGGAAGCAATGCCGGTAAGGGTCATGGAGAGAAGAAGGA TGCCCCGGCGGCCAAATCGGTCCACGGTGACCCCCAGGAAGACA CAGGCCAGGGCTGCCGTGCCGCTGGCCAGCAGAGAGCACAGGTA GAAGTCCGATGGGCTCCCTCCTCCTCCCACAGGCTGGTAGCAGT GGCGAATGGCATGGGCAATGAAGCTGTGAGAAGGGGTGGGGAAG CAACGAACAGAGCCTGAATCCCCTCCCGCAGCCTTC Y ACAGTGCTGATGGTCCCTCCGCAGGGGTCCCCGGGCCTTCCCACC CAGCCAGCCCCCATTCAGCTACTGGCCTGGCCCTCAGCCCAGA CAGCCAGGCTCACTTGGTGAAGCCCAGGATAAGCAGATTTTCC AGATGTTGCGGTAGTTGAGGAGGGAAGCAAAGGAAAAGGAGGAT GTTGCAGGGAGAGGGCAGGTATTCTCCAGGTCTTTGGGAGAGAG AGAGGAGCTGTGAGAAGAAAACCCTGAGATCCCAGGATCCTCCC ACATGGGTGGTGGGGACCCTGGGAGAAGGTGGCACA	14

SULT2B1	rs10426628 (at position 501)		GAGAGGGCTGGAGATAGAGTTAGAGGTGGGTGTGGGGCTGCAAT TATAAATAAGGAGGTCAAGGGGGCCCCAGTGAGTAAAAACCCG GGAGGTGATGCTGGGAGCCACAGAGGTTTCTAGAGGAAGAGCAT TTCAGGCAAGAGGGAAAAGCAGGTGCAAAGGCCCTGAGGTGGGT GTATCTGAGGTGCAAGAGGGAGGTGCGTGTGGCCGGAGCCAAGT GAGCAAGTGGGGAAAAGGAGTGGAGATGAGGTCCGGGTGGGGAGG CAACAGGGGCCAAAATGTGCAGGGCCGCGTGGGCCCGGTGCGGA CTTGAGCTCTGACCGAGTGACGAGGCAGCGCGGCAGGGATCGCA GCAGAGGAGGAGCCCGAGCGGGCTTAGGCTTCACCGGCGCCCTC TGGCGGTGAGATGGGGACCGACTGTGAGAGAGCAGAAGCCGGG AAGCCCGAGAGGCGGCGCTGGCCGGGGTCCAGGGGAGGGGACGG TGGCTGGACTAGGGTG N TGGTCACGGAGGTATTTGAAGGTGAGACCGGGAGGATTTGCTG ACAGACTGGATGTGGGTGTGAGAGAAGGGGATGAGTCAAGGGTG ACTCCAAGGTTTCGGCGGAAGCAACTGTCAGGGTGGGGCAACTG GGAAGGGGAGGAGGGAGGGGAGCAGGCGGGGAGGAAGACACGG GCTCTGCGGCGGCCATCCCACGTCCAGCAGAGGCTGCCGGTGG GCACAGAAATGCACCCATCGGGATCTGAGGGGAGAGACTTGGGC TGACACGTCCATTTGCAATTCCTTGGCACATATATTCTGTAATT AATCGAAGCTCCAGGACTGGGTGATATCGCCAAGGAATTGAGCG TGGACAGAGCAGAGAGGGGGGGCTGCAATCAGGAAGTCCCCAAG ACCACCCTCAGGCTCAACGACTTAGCAAACTCATAGTGTTTCAG CAAAGCTGCTGGACTCCCGGGTTATAGTCAATTACAATGAAAGC AAACGTGAAAATCAGG	15
		A/C/G/T		

It will be appreciated by a person of skill in the art that further linked polymorphic sites and combined polymorphic sites may be determined. A haplotype of the above genes can be created by assessing polymorphisms in normal subjects using a program that has an expectation maximization algorithm (for example PHASE). A constructed haplotype of these genes may be used to find combinations of SNPs that are in LD with the tag SNPs (tSNPs) identified herein. Accordingly, the haplotype of an individual could be determined by genotyping other SNPs or other polymorphisms that are in LD with the tSNPs identified herein. Single polymorphic sites or combined polymorphic sites in LD may also be genotyped for assessing subject risk of cardiotoxicity following anthracycline treatment.

It will be appreciated by a person of skill in the art that the numerical designations of the positions of polymorphisms within a sequence are relative to the specific sequence and the orientation of the strand being read (i.e. forward or reverse). Also the same positions may be assigned different numerical designations depending on the way in which the sequence is numbered and the sequence chosen. Furthermore, sequence variations within the population, such as insertions or deletions, may change the relative position and subsequently the numerical designations of particular nucleotides at and around a polymorphic site. For example, the sequences represented by accession numbers NC_000014.8, NM_016609, AJ243653.1 and BC111015.1 all comprise SLC22A17 nucleotide sequences, but may have some sequence differences and numbering differences between them. Furthermore, one of skill in the art

will appreciate that a variety of sequencing, amplification, extension, genotyping or hybridization primers or probes may be designed to specifically identify the polymorphisms described in TABLE 2, and the sequences flanking the various polymorphisms as provided herein are illustrative examples. One of skill in the art will also appreciate that a variety of sequencing, amplification, extension, genotyping or hybridization primers or probes adjacent to, complimentary to, or overlapping with the sequences provided in TABLE 2, may be developed or designed for the identification of the polymorphisms described herein, without going beyond the scope of various embodiments described herein.

The gene sequences described herein in association with SLC28A3, SLC28A3, UGT1A6, SULT2B1, SLC28A1, SLC22A17, SLC22A7, ABCB4, ABCB4, and HNMT are meant to include genomic sequences, cDNA sequences, mRNA sequences, and further may include 5' and 3' untranslated sequences, introns and the like. Sequence databases with this information, such as GenBank, operated by the National Centre for Biotechnology Information (NCBI) store such information in a retrievable format, and are publicly accessible. A person of skill in the art will appreciate the various methods and tools that may be used to access such information, in a context suitable to their particular application of aspects described herein. Furthermore, a person of skill in the art would appreciate that the sequences may appear in either orientation (either forward or reverse strand or both).

Polymorphic sites in SEQ ID NO:1-15 are identified by their variant designation (i.e. M, W, Y, S, R, K, V, B, D, H, N or by “-” for a deletion, a “+” or for example “G” etc. for an insertion).

An “rs” prefix designates a SNP in the database is found at the NCBI SNP database (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Snp>). The “rs” numbers are the NCBI | rsSNP ID form.

The Sequences given in TABLE 2 (SEQ ID NO:1-15) above may be useful to a person of skill in the art in the design of primers and probes or other oligonucleotides or PNAs for the identification of polymorphisms as described herein.

An “allele” is defined as any one or more alternative forms of a given gene. In a diploid cell or organism the members of an allelic pair (i.e. the two alleles of a given gene) occupy corresponding positions (loci) on a pair of homologous chromosomes and if these alleles are genetically identical the cell or organism is said to be “homozygous”, but if genetically different the cell or organism is said to be “heterozygous” with respect to the particular gene.

A “gene” is an ordered sequence of nucleotides located in a particular position on a particular chromosome that encodes a specific functional product and may include untranslated and untranscribed sequences in proximity to the coding regions (5’ and 3’ to the coding sequence). Such non-coding sequences may contain regulatory sequences needed for transcription and translation of the sequence or introns etc. or may as yet to have any function attributed to them beyond the occurrence of the SNP of interest.

A “genotype” is defined as the genetic constitution of an organism, usually in respect to one gene or a few genes or a region of a gene relevant to a particular context (i.e. the genetic loci responsible for a particular phenotype).

A “phenotype” is defined as the observable characters of an organism. In gene association studies, the genetic model at a given locus can change depending on the selection pressures (i.e., the environment), the population studied, or the outcome variable (i.e., the phenotype).

A similar observation would be seen in a gene association study with the hemoglobin, beta gene (HBB) with mortality as the primary outcome variable. A mutation in the HBB gene, which normally produces the beta chain subunit of hemoglobin (B allele), results in an abnormal beta chain called hemoglobin S (S allele; Allison A (1955) Cold Spring Harbor Symp. Quant. Biol. 20:239-255). Hemoglobin S results in abnormal sickle-shaped red blood cells which lead to anemia and other serious complications including death. In the absence of malaria, a gene association study with the HBB gene would suggest a codominant model (survival(BB) > survival (BS) > survival (SS)). However, in the presence of malaria, a gene association study with the HBB gene would suggest a heterozygote advantage model (survival(BB) < survival(BS) > survival(SS)).

A “single nucleotide polymorphism” (SNP) occurs at a polymorphic site occupied by a single nucleotide, which is the site of variation between allelic sequences. The site is usually preceded by and followed by highly conserved sequences of the allele (e.g., sequences that vary in less than 1/100 or 1/1000 members of the populations). A single nucleotide polymorphism usually arises due to substitution of one nucleotide for another at the polymorphic site. A “transition” is the replacement of one purine by another purine or one pyrimidine by another pyrimidine. A “transversion” is the replacement of a purine by a pyrimidine or vice versa. Single nucleotide polymorphisms can also arise from a deletion (represented by “-“ or “*del*”) of a nucleotide or an insertion (represented by “+“ or “*ins*” or “I”) of a nucleotide relative to

a reference allele. Furthermore, a person of skill in the art would appreciate that an insertion or deletion within a given sequence could alter the relative position and therefore the position number of another polymorphism within the sequence. Furthermore, although an insertion or deletion may by some definitions not qualify as a SNP as it may involve the deletion of or insertion of more than a single nucleotide at a given position, as used herein such polymorphisms are also called SNPs as they generally result from an insertion or deletion at a single site within a given sequence.

A “subject”, as used herein, refers to a patient or test subject, for example a human patient. The subject may have been previously diagnosed with a neoplastic disorder, or may be suspected of having a neoplastic disorder and thus may be a candidate for a chemotherapeutic regimen. The subject may be selected as part of a general population (for example a ‘control’ subject), or may be selected as part of a particular ethnic, gender, age or genetic subgroup of a population, or may be excluded from selection as part of a particular ethnic, gender, age or genetic subgroup of a population. Patients and test subjects, whether control or not, may be generally referred to as a subject.

As used herein, the terms “cancer” or “neoplastic condition” or “neoplastic disorder” or “neoplastic disease” refer to a proliferative disorder caused or characterized by the proliferation of cells which have lost susceptibility to normal growth control. A “cancer” or “neoplastic condition” or “neoplastic disorder” or “neoplastic disease” may include tumors and any other proliferative disorders. Cancers of the same tissue type usually originate in the same tissue, and may be divided into different subtypes based on their biological characteristics. Four general categories of cancers are carcinoma (epithelial tissue derived), sarcoma (connective tissue or mesodermal derived), leukemia (blood-forming tissue derived) and lymphoma (lymph tissue derived). Over 200 different types of cancers are known, and every organ and tissue of the body may be affected. Specific examples of cancers that do not limit the definition of cancer may include melanoma, leukemia, astrocytoma, glioblastoma, retinoblastoma, lymphoma, glioma, Hodgkins’ lymphoma and chronic lymphocyte leukemia. Examples of organs and tissues that may be affected by various cancers include pancreas, breast, thyroid, ovary, uterus, testis, prostate, thyroid, pituitary gland, adrenal gland, kidney, stomach, esophagus or rectum, head and neck, bone, nervous system, skin, blood, nasopharyngeal tissue, lung, urinary tract, cervix, vagina, exocrine glands and endocrine glands. Alternatively, a cancer may be multicentric or of unknown primary site (CUPS).

As used herein, a “therapeutic regimen” refers to a chemotherapeutic regimen or a radiotherapy regimen, or a combination thereof.

As used herein, a “chemotherapeutic regimen” or “chemotherapy” refers to the use of at least one chemotherapy agent to destroy cancerous cells. There are a myriad of such chemotherapy agents available for treating cancer. Chemotherapy agents may be administered to a subject in a single bolus dose, or may be administered in smaller doses over time. A single chemotherapeutic agent may be used (single-agent therapy) or more than one agent may be used in combination (combination therapy). Chemotherapy may be used alone to treat some types of cancer. Alternatively, chemotherapy may be used in combination with other types of treatment, for example, radiotherapy or alternative therapies (for example immunotherapy) as described herein. Additionally, a chemosensitizer may be administered as a combination therapy with a chemotherapy agent.

As used herein, a “chemotherapeutic agent” or “chemotherapeutic agent” refers to a medicament that may be used to treat cancer, and generally has the ability to kill cancerous cells directly. Examples of chemotherapeutic agents include alkylating agents, antimetabolites, natural products, hormones and antagonists, and miscellaneous agents. Examples of alternate names are indicated in brackets. Examples of alkylating agents include nitrogen mustards such as mechlorethamine, cyclophosphamide, ifosfamide, melphalan (L-sarcolysin) and chlorambucil; ethylenimines and methylmelamines such as hexamethylmelamine and thiotepa; alkyl sulfonates such as busulfan; nitrosoureas such as carmustine (BCNU), semustine (methyl-CCNU), lomustine (CCNU) and streptozocin (streptozotocin); DNA synthesis antagonists such as estramustine phosphate; and triazines such as dacarbazine (DTIC, dimethyl-triazenoimidazolecarboxamide) and temozolomide. Examples of antimetabolites include folic acid analogs such as methotrexate (amethopterin); pyrimidine analogs such as fluorouracil (5-fluorouracil, 5-FU, 5FU), floxuridine (fluorodeoxyuridine, FUdR), cytarabine (cytosine arabinoside) and gemcitabine; purine analogs such as mercaptopurine (6-mercaptopurine, 6-MP), thioguanine (6-thioguanine, TG) and pentostatin (2'-deoxycoformycin, deoxycoformycin), cladribine and fludarabine; and topoisomerase inhibitors such as amsacrine. Examples of natural products include vinca alkaloids such as vinblastine (VLB) and vincristine; taxanes such as paclitaxel and docetaxel (Taxotere); epipodophyllotoxins such as etoposide and teniposide; camptothecins such as topotecan or irinotecan; antibiotics such as dactinomycin (actinomycin D), bleomycin, mitomycin (mitomycin C); anthracycline antibiotics such as daunorubicin (daunomycin, rubidomycin), doxorubicin, idarubicin, epirubicin; enzymes such as L-asparaginase; and biological response modifiers such as interferon alpha and interleukin 2. Examples of hormones and antagonists include luteinising releasing hormone agonists such as buserelin; adrenocorticosteroids such as prednisone and related preparations; progestins such as hydroxyprogesterone caproate, medroxyprogesterone acetate and megestrol acetate; estrogens such as diethylstilbestrol and ethinyl estradiol and related preparations; estrogen antagonists such as tamoxifen

and anastrozole; androgens such as testosterone propionate and fluoxymesterone and related preparations; androgen antagonists such as flutamide and bicalutamide; and gonadotropin-releasing hormone analogs such as leuprolide. Examples of miscellaneous agents include thalidomide; platinum coordination complexes such as cisplatin (cis-DDP), carboplatin, oxaliplatin, tetraplatin, ormaplatin, iproplatin or satraplatin; anthracenediones such as mitoxantrone; substituted ureas such as hydroxyurea; methylhydrazine derivatives such as procarbazine (N-methylhydrazine, MIH); adrenocortical suppressants such as mitotane (o,p'-DDD) and aminoglutethimide; RXR agonists such as bexarotene; or tyrosine kinase inhibitors such as imatinib. Alternate names and trade-names of these and additional examples of chemotherapeutic agents, and their methods of use including dosing and administration regimens, will be known to an individual versed in the art, and may be found in, for example "The Pharmacological basis of therapeutics", 10th edition. HARDMAN HG., LIMBIRD LE. editors. McGraw-Hill, New York, or in "Clinical Oncology", 3rd edition. Churchill Livingstone/ Elsevier Press, 2004. ABELOFF, MD. editor.

2. General Methods

Once a subject is identified as a candidate for anthracycline administration, then genetic sequence information may be obtained from the subject to determine the risk of cardiotoxicity for the subject. Genetic sequence information may be obtained from a subject by any of several methods. For example, a biological sample comprising genetic material with a sequence or sequences of interest, may be obtained from the subject, for example a blood sample, a saliva sample, a hair sample including a follicle, skin scraping, such as a cheek scraping and the like. Or alternatively genetic sequence information may already have been obtained from the subject. For example, a subject may have already provided a biological sample for other purposes or may have even had their genetic sequence determined in whole or in part and stored for future use. Genetic sequence information may be obtained in numerous different ways and may involve the collection of a biological sample that contains genetic material, particularly, genetic material containing the sequence or sequences of interest. Many methods are known in the art for collecting biological samples and extracting genetic material from those samples. Genetic material can be extracted from blood, tissue, hair and other biological material. There are many methods known to isolate DNA and RNA from biological material. Typically, DNA may be isolated from a biological sample when first the sample is lysed and then the DNA is separated from the lysate according to any one of a variety of multi-step protocols, which can take varying lengths of time. DNA isolation methods may involve the use of phenol (Sambrook, J. *et al.*, "Molecular Cloning", Vol. 2, pp. 9.14-9.23, Cold Spring Harbor Laboratory Press (1989) and Ausubel, Frederick M. *et al.*, "Current Protocols in Molecular Biology", Vol. 1, pp. 2.2.1-2.4.5, John Wiley & Sons, Inc. (1994)). Typically, a biological sample is

lysed in a detergent solution and the protein component of the lysate is digested with proteinase for 12-18 hours. Next, the lysate is extracted with phenol to remove most of the cellular components, and the remaining aqueous phase is processed further to isolate DNA. In another method, described in Van Ness *et al.* (U.S. Pat. # 5,130,423), non-corrosive phenol derivatives are used for the isolation of nucleic acids. The resulting preparation is a mix of RNA and DNA.

Other methods for DNA isolation utilize non-corrosive chaotropic agents. These methods, which are based on the use of guanidine salts, urea and sodium iodide, involve lysis of a biological sample in a chaotropic aqueous solution and subsequent precipitation of the crude DNA fraction with a lower alcohol. The resulting nucleic acid sample may be used 'as-is' in further analyses or may be purified further. Additional purification of the precipitated, crude DNA fraction may be achieved by any one of several methods, including, for example, column chromatography (Analects, (1994) Vol 22, No. 4, Pharmacia Biotech), or exposure of the crude DNA to a polyanion-containing protein as described in Koller (U.S. Pat. # 5,128,247).

Yet another method of DNA isolation, which is described by Botwell, D. D. L. (Anal. Biochem. (1987) 162:463-465) involves lysing cells in 6M guanidine hydrochloride, precipitating DNA from the lysate at acid pH by adding 2.5 volumes of ethanol, and washing the DNA with ethanol.

Numerous other methods are known in the art to isolate both RNA and DNA, such as the one described by CHOMCZYNSKI (U.S. Pat. # 5,945,515), whereby genetic material can be extracted efficiently in as little as twenty minutes. EVANS and HUGH (U.S. Pat. # 5,989,431) describe methods for isolating DNA using a hollow membrane filter.

The level of expression of specific nucleic acids such as mRNAs or microRNAs, copy number of a gene, or the degree of heterozygosity for a polymorphism may also be determined once the nucleic acid sample has been obtained. Quantitative and semi-quantitative methods are known in the art, and may be found in, for example AUSUBEL, *supra*; SAMBROOK, *supra* or Harrison's Principles of Internal Medicine 15th ed. BRAUNWALD et al eds. McGraw-Hill.

Once a subject's genetic material has been obtained from the subject it may then be further be amplified by Reverse Transcription Polymerase Chain Reaction (RT-PCR), Polymerase Chain Reaction (PCR), Transcription Mediated Amplification (TMA), Ligase chain reaction (LCR), Nucleic Acid Sequence Based Amplification (NASBA) or other methods known in the art, and then further analyzed to detect or

determine the presence or absence of one or more polymorphisms or mutations in the sequence of interest, provided that the genetic material obtained contains the sequence of interest. Particularly, a person may be interested in determining the presence or absence of a polymorphism in a cardiotoxicity associated gene sequence, as described herein.

Detection or determination of a nucleotide identity, or the presence of one or more single nucleotide polymorphism(s) (SNP typing), may be accomplished by any one of a number methods or assays known in the art. Many DNA typing methodologies are useful for use in the detection of SNPs. The majority of SNP genotyping reactions or assays can be assigned to one of four broad groups (sequence-specific hybridization, primer extension, oligonucleotide ligation and invasive cleavage). Furthermore, there are numerous methods for analyzing/detecting the products of each type of reaction (for example, fluorescence, luminescence, mass measurement, electrophoresis, etc.). Furthermore, reactions can occur in solution or on a solid support such as a glass slide, a chip, a bead, etc.

In general, sequence-specific hybridization involves a hybridization probe, which is capable of distinguishing between two DNA targets differing at one nucleotide position by hybridization. Usually probes are designed with the polymorphic base in a central position in the probe sequence, whereby under optimized assay conditions only the perfectly matched probe target hybrids are stable and hybrids with a one base mismatch are unstable. A strategy which couples detection and sequence discrimination is the use of a "molecular beacon", whereby the hybridization probe (molecular beacon) has 3' and 5' reporter and quencher molecules and 3' and 5' sequences which are complementary such that absent an adequate binding target for the intervening sequence the probe will form a hairpin loop. The hairpin loop keeps the reporter and quencher in close proximity resulting in quenching of the fluorophor (reporter) which reduces fluorescence emissions. However, when the molecular beacon hybridizes to the target the fluorophor and the quencher are sufficiently separated to allow fluorescence to be emitted from the fluorophor.

Similarly, primer extension reactions (i.e. mini sequencing, nucleotide-specific extensions, or simple PCR amplification) are useful in sequence discrimination reactions. For example, in mini sequencing a primer anneals to its target DNA immediately upstream of the SNP and is extended with a single nucleotide complementary to the polymorphic site. Where the nucleotide is not complementary, no extension occurs.

Oligonucleotide ligation assays require two sequence-specific probes and one common ligation probe per SNP. The common ligation probe hybridizes adjacent to a sequence-specific probe and when there is a perfect match of the appropriate sequence-specific probe, the ligase joins both the sequence-specific and the common probes. Where there is not a perfect match the ligase is unable to join the sequence-specific and common probes. Probes used in hybridization can include double-stranded DNA, single-stranded DNA and RNA oligonucleotides, and peptide nucleic acids. Hybridization methods for the identification of single nucleotide polymorphisms or other mutations involving a few nucleotides are described in the U.S. Pat. 6,270,961; 6,025,136; and 6,872,530. Suitable hybridization probes for use in accordance with the invention include oligonucleotides and PNAs from about 10 to about 400 nucleotides, alternatively from about 20 to about 200 nucleotides, or from about 30 to about 100 nucleotides in length.

A unimolecular segment amplification method for amplifying nucleic acids is described in US patent 5854033. A rolling circle replication reporter system may be used for identification of polymorphisms or mutations.

An invasive cleavage method employs an “Invader™” (Applied Biosystems) probe and sequence-specific probes to hybridize with the target nucleic acid, usually DNA, with an overlap of one nucleotide. When the sequence specific probe is an exact match to the site of polymorphism, the overlapping probes form a structure that is specifically cleaved by a FLAP endonuclease, Release of the 5' end of the allele-specific probe may be detected by known methods as described. See for example, Lu, M., et al. J. Am. Chem. Soc. 2001, 124, 7924 – 7931; Lyamichev, et al. 1999. Nature Biotech. 17, 292 – 296; Landegren et al. 1998. Genome Research, 8, 769 – 776; Brookes, 1999. Gene 234, 177 – 186; Chen, et al 2004. J. Am. Chem. Soc. 126, 3016-3017; Wang, D.G., et al. Science 1998, 280, 1077 – 1082. The TaqMan™ assay (Applied Biosystems) exploits the 5' exonuclease activity of the Taq polymerase to displace and cleave an oligonucleotide probe hybridized to the target nucleic acid, usually DNA, generating a fluorescent signal. See, for example U.S. Patents 4,683,202, 4,683,195, and 4,965,188.

5' exonuclease activity or TaqMan™ assay (Applied Biosystems™) is based on the 5' nuclease activity of Taq polymerase that displaces and cleaves the oligonucleotide probes hybridized to the target DNA generating a fluorescent signal. It is necessary to have two probes that differ at the polymorphic site wherein one probe is complementary to the ‘normal’ sequence and the other to the mutation of interest. These probes have different fluorescent dyes attached to the 5' end and a quencher attached to the 3' end when the probes are intact the quencher interacts with the fluorophor by fluorescence resonance energy transfer (FRET) to quench the fluorescence of the probe. During the PCR annealing step the

hybridization probes hybridize to target DNA. In the extension step the 5' fluorescent dye is cleaved by the 5' nuclease activity of Taq polymerase, leading to an increase in fluorescence of the reporter dye. Mismatched probes are displaced without fragmentation. The presence of a mutation in a sample is determined by measuring the signal intensity of the two different dyes.

The Illumina Golden Gate™ Assay uses a combined oligonucleotide ligation assay/ allele-specific hybridization approach (SHEN R *et al Mutat Res* 2005573:70-82). The first series of steps involve the hybridization of three oligonucleotides to a set of specific target SNPs; two of these are fluorescently-labelled allele-specific oligonucleotides (ASOs) and the third a locus-specific oligonucleotide (LSO) binding 1-20 bp downstream of the ASOs. A second series of steps involve the use of a stringent polymerase with high 3' specificity that extends only oligonucleotides specifically matching an allele at a target SNP. The polymerase extends until it reaches the LSO. Locus-specificity is ensured by requiring the hybridization of both the ASO and LSO in order that extension can proceed. After PCR amplification with universal primers, these allele-specific oligonucleotide extension products are hybridized to an array which has multiple discretely tagged addresses (in this case 1536 addresses) which match an address embedded in each LSO. Fluorescent signals produced by each hybridization product are detected by a bead array reader from which genotypes at each SNP locus may be ascertained.

It will be appreciated that numerous other methods for sequence discrimination and detection are known in the art and some of which are described in further detail below. It will also be appreciated that reactions such as arrayed primer extension mini sequencing, tag microarrays and sequence-specific extension could be performed on a microarray. One such array based genotyping platform is the microsphere based tag-it high throughput genotyping array (BORTOLIN S. *et al. Clinical Chemistry* (2004) 50(11): 2028-36). This method amplifies genomic DNA by PCR followed by sequence-specific primer extension with universally tagged genotyping primers. The products are then sorted on a Tag-It array and detected using the Luminex xMAP system.

Mutation detection methods may include but are not limited to the following:

Restriction Fragment Length Polymorphism (RFLP) strategy – An RFLP gel-based analysis can be used to indicate the presence or absence of a specific mutation at polymorphic sites within a gene. Briefly, a short segment of DNA (typically several hundred base pairs) is amplified by PCR. Where possible, a specific restriction endonuclease is chosen that cuts the short DNA segment when one polymorphism is present but does not cut the short DNA segment when the polymorphism is not present, or vice versa. After incubation of the PCR amplified DNA with this restriction endonuclease, the reaction products are

then separated using gel electrophoresis. Thus, when the gel is examined the appearance of two lower molecular weight bands (lower molecular weight molecules travel farther down the gel during electrophoresis) indicates that the DNA sample had a polymorphism was present that permitted cleavage by the specific restriction endonuclease. In contrast, if only one higher molecular weight band is observed (at the molecular weight of the PCR product) then the initial DNA sample had the polymorphism that could not be cleaved by the chosen restriction endonuclease. Finally, if both the higher molecular weight band and the two lower molecular weight bands are visible then the DNA sample contained both polymorphisms, and therefore the DNA sample, and by extension the subject providing the DNA sample, was heterozygous for this polymorphism;

For example the Maxam-Gilbert technique for sequencing (MAXAM AM. and GILBERT W. Proc. Natl. Acad. Sci. USA (1977) 74(4):560-564) involves the specific chemical cleavage of terminally labelled DNA. In this technique four samples of the same labeled DNA are each subjected to a different chemical reaction to effect preferential cleavage of the DNA molecule at one or two nucleotides of a specific base identity. The conditions are adjusted to obtain only partial cleavage, DNA fragments are thus generated in each sample whose lengths are dependent upon the position within the DNA base sequence of the nucleotide(s) which are subject to such cleavage. After partial cleavage is performed, each sample contains DNA fragments of different lengths, each of which ends with the same one or two of the four nucleotides. In particular, in one sample each fragment ends with a C, in another sample each fragment ends with a C or a T, in a third sample each ends with a G, and in a fourth sample each ends with an A or a G. When the products of these four reactions are resolved by size, by electrophoresis on a polyacrylamide gel, the DNA sequence can be read from the pattern of radioactive bands. This technique permits the sequencing of at least 100 bases from the point of labeling. Another method is the dideoxy method of sequencing was published by SANGER *et al.* (Proc. Natl. Acad. Sci. USA (1977) 74(12):5463-5467). The Sanger method relies on enzymatic activity of a DNA polymerase to synthesize sequence-dependent fragments of various lengths. The lengths of the fragments are determined by the random incorporation of dideoxynucleotide base-specific terminators. These fragments can then be separated in a gel as in the Maxam-Gilbert procedure, visualized, and the sequence determined. Numerous improvements have been made to refine the above methods and to automate the sequencing procedures. Similarly, RNA sequencing methods are also known. For example, reverse transcriptase with dideoxynucleotides have been used to sequence encephalomyocarditis virus RNA (ZIMMERN D. and KAESBERG P. Proc. Natl. Acad. Sci. USA (1978) 75(9):4257-4261). MILLS DR. and KRAMER FR. (Proc. Natl. Acad. Sci. USA (1979) 76(5):2232-2235) describe the use of Q β replicase and the nucleotide analog inosine for sequencing RNA in a chain-termination mechanism. Direct chemical methods for

sequencing RNA are also known (PEATTIE DA. *Proc. Natl. Acad. Sci. USA* (1979) 76(4):1760-1764). Other methods include those of Donis-Keller *et al.* (1977, *Nucl. Acids Res.* 4:2527-2538), SIMONCSITS A. *et al.* (*Nature* (1977) 269(5631):833-836), AXELROD VD. *et al.* (*Nucl. Acids Res.*(1978) 5(10):3549-3563), and KRAMER FR. and MILLS DR. (*Proc. Natl. Acad. Sci. USA* (1978) 75(11):5334-5338). Nucleic acid sequences can also be read by stimulating the natural fluoresce of a cleaved nucleotide with a laser while the single nucleotide is contained in a fluorescence enhancing matrix (U.S. Pat. # 5,674,743); In a mini sequencing reaction, a primer that anneals to target DNA adjacent to a SNP is extended by DNA polymerase with a single nucleotide that is complementary to the polymorphic site. This method is based on the high accuracy of nucleotide incorporation by DNA polymerases. There are different technologies for analyzing the primer extension products. For example, the use of labeled or unlabeled nucleotides, ddNTP combined with dNTP or only ddNTP in the mini sequencing reaction depends on the method chosen for detecting the products.

Probes used in hybridization can include double-stranded DNA, single-stranded DNA and RNA oligonucleotides, and peptide nucleic acids. Hybridization methods for the identification of single nucleotide polymorphisms or other mutations involving a few nucleotides are described in the U.S. Pat. 6,270,961; 6,025,136; and 6,872,530. Suitable hybridization probes for use in accordance with the invention include oligonucleotides and PNAs from about 10 to about 400 nucleotides, alternatively from about 20 to about 200 nucleotides, or from about 30 to about 100 nucleotides in length.

A template-directed dye-terminator incorporation with fluorescent polarization-detection (TDI-FP) method is described by FREEMAN BD. *et al.* (*J Mol Diagnostics* (2002) 4(4):209-215) for large scale screening.

Oligonucleotide ligation assay (OLA) is based on ligation of probe and detector oligonucleotides annealed to a polymerase chain reaction amplicon strand with detection by an enzyme immunoassay (VILLAHERMOSA ML. *J Hum Virol* (2001) 4(5):238-48; ROMPPANEN EL. *Scand J Clin Lab Invest* (2001) 61(2):123-9; IANNONE MA. *et al.* *Cytometry* (2000) 39(2):131-40).

Ligation-Rolling Circle Amplification (L-RCA) has also been successfully used for genotyping single nucleotide polymorphisms as described in QI X. *et al.* *Nucleic Acids Res* (2001) 29(22):E116.

5' nuclease assay has also been successfully used for genotyping single nucleotide polymorphisms (AYDIN A. *et al.* *Biotechniques* (2001) (4):920-2, 924, 926-8.).

Polymerase proofreading methods are used to determine SNPs identities, as described in WO 0181631.

Detection of single base pair DNA mutations by enzyme-amplified electronic transduction is described in PATOLSKY F *et al.* Nat Biotech. (2001) 19(3):253-257.

Gene chip or microarray technologies are also known for single nucleotide polymorphism discrimination whereby numerous polymorphisms may be tested for simultaneously on a single array (for example: EP 1120646; and GILLES PN. *et al.* Nat. Biotechnology (1999) 17(4):365-70).

Matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectroscopy is also useful in the genotyping single nucleotide polymorphisms through the analysis of microsequencing products (HAFF LA. and SMIRNOV IP. Nucleic Acids Res. (1997) 25(18):3749-50; HAFF LA. and SMIRNOV IP. Genome Res. (1997) 7:378-388; SUN X. *et al.* Nucleic Acids Res. (2000) 28 e68; BRAUN A. *et al.* Clin. Chem. (1997) 43:1151-1158; LITTLE DP. *et al.* Eur. J. Clin. Chem. Clin. Biochem. (1997) 35:545-548; FEI Z. *et al.* Nucleic Acids Res. (2000) 26:2827-2828; and BLONDAL T. *et al.* Nucleic Acids Res. (2003) 31(24):e155).

Sequence-specific PCR methods have also been successfully used for genotyping single nucleotide polymorphisms (HAWKINS JR. *et al.* Hum Mutat (2002) 19(5):543-553). Alternatively, a Single-Stranded Conformational Polymorphism (SSCP) assay or a Cleavase Fragment Length Polymorphism (CFLP) assay may be used to detect mutations as described herein.

US 7,074,597 describes methods for multiplex genotyping using solid phase capturable dideoxynucleotides and mass spectrometry. Nucleotide identity is detected at a specific site of a nucleic acid sample by contacting DNA-primer complex with labeled dideoxynucleotides (ddNTPs) to generate labeled single base extended (SBE) primer. The identifying ddNTP may be within the SBE primer.

Multiplex analysis of PCR-amplified products may also be used to detect specific SNPs. Reporting DNA sequences comprising a fluorophore on a 5' end may be used to combine a multiplex PCR amplification reaction with microsphere based hybridization (US 7,083,951). Other multiplex detection methods include BeadArrayTM and similar hybridization-based methods, for example, those described in US Patent Nos. 6,429,027, 6,396,995, 6,355,431.

Microarray or 'gene chips' of oligonucleotides may be used for SNP discrimination. Oligonucleotides may be nucleic acids or modified nucleic acids, including PNAs, and may be 'spotted' onto a solid matrix, such as a glass or plastic slide. Alternatively, oligonucleotides may be synthesized in situ on the slide. See, for example, GAO et al 2004. Biopolymers 73:579-596; US 5,445,934; US 5,744,305, US 5,800,992, US 5,796,715.

Alternatively, if a subject's sequence data is already known, then obtaining may involve retrieval of the subject's nucleic acid sequence data (for example from a database), followed by determining or detecting the identity of a nucleic acid or genotype at a polymorphic site by reading the subject's nucleic acid sequence at the one or more polymorphic sites.

Once the identity of a polymorphism(s) is determined or detected an indication may be obtained as to the subject's risk of cardiotoxicity following anthracycline administration. Methods for predicting a subject's risk of cardiotoxicity following anthracycline administration may be useful in making decisions regarding the administration of anthracycline(s).

TREATMENT

Anthracycline compounds (for example, doxorubicin) may be used to treat a variety of cancers in children and adults. In a given therapeutic regimen, the anthracycline compound may be administered alone or in combination with other chemotherapeutic agents in various doses and compositions, depending on the type of cancer, age of subject, health of subject, body mass, etc. The choice of dose, chemotherapeutic agents or combinations, methods of administration and the like will be known to those skilled in the art. Further, methods of assessing response to treatment and side effects are also known. For example, heart function in a subject suspected of experiencing cardiotoxicity may be assessed by various methods including medical history, electrocardiogram (ECG) monitoring, endomyocardial biopsy, radionuclide angiography (MUGA scan) or LVEF monitoring with serial echo or exercise stress testing, or other methods that may be dependent on the age and condition of the subject, as are known in the art. Early signs of cardiotoxicity may include persistent reduction in the voltage of the QRS wave, prolongation of the systolic time interval, or reduction of LVEF as determined by echo or MUGA. A reduction of 10% to below the lower limit of normal, 20% at any level, or an absolute LVEF $\leq 45\%$ indicates deterioration of cardiac function.)

Response to a therapeutic regimen may be monitored. Tumor staging provides a method to assess the size and spread of a tumor in response to a treatment regimen. The TNM tumor staging system uses three

components to express the anatomic extent of disease: T is a measure of the local extent of tumor spread (size), N indicates the presence or absence of metastatic spread to regional lymph nodes, and M specifies the presence or absence of metastatic spread to distant sites. The combination of these classifications combine to provide a stage grouping. Clinical TNM (cTNM) defines the tumor based on clinical evidence. Pathologic TNM (pTNM) defines the tumor based on examination of a surgically resected specimen.

Changes in tumor size may be observed by various imaging methods known to physicians or surgeons in the field of oncology therapy and diagnostics. Examples of imaging methods include positron emission tomography (PET) scanning, computed tomography (CT) scanning, PET/CT scanning, magnetic resonance imaging (MRI), chemical shift imaging, radiography, bone-scan, mammography, fiberoptic colonoscopy or ultrasound. Contrast agents, tracers and other specialized techniques may also be employed to image specific types of cancers, or for particular organs or tissues, and will be known to those skilled in the art. Changes in rate of metastasis may also be observed by the various imaging methods, considering particularly the appearance, or frequency of appearance, of tumors distal to the primary site. Alternatively, the presence of tumor cells in lymph nodes adjacent and distal to the primary tumor site may also be detected and used to monitor metastasis.

A subject may be tested for a cardiotoxicity-associated polymorphism before undergoing a therapeutic regimen involving an anthracycline compound. If a subject's genotype includes a cardiotoxicity-associated polymorphism, this may indicate that the subject is at a risk for cardiotoxicity when an anthracycline compound is administered.

A subject at risk for cardiotoxicity may be administered a therapeutic regimen involving an anthracycline compound and the cardiac function monitored as described. If a decrease in cardiac function is identified, the therapeutic regimen may be altered to decrease the dose of the anthracycline compound, eliminate the dose of the anthracycline compound, or increase the dose of a second chemotherapeutic agent in the therapeutic regimen. Examples of chemotherapeutic agents that may be used in combination with an anthracycline compound in a therapeutic regimen may include, for example, cyclophosphamide, Ifosphamide, fluorouracil, Paclitaxel, vincristine, cisplatin, streptozocin, docetaxel, and the like.

A subject at risk for cardiotoxicity may also be administered a therapeutic regimen involving an anthracycline compound and the cardiac function monitored as described. The therapeutic regimen may be supplemented to include a cardioprotective agent. Examples of cardioprotective agents are known in

the art, and may include those described by Wouters et al 2005. Br. J Hematol 131:561-578). For example, Dexrazoxane is a cardioprotective agent and is approved for use in conjunction with doxorubicin to reduce the incidence and severity of cardiomyopathy associated with doxorubicin administration.

Alternatively, a subject at risk for cardiotoxicity may be administered a therapeutic regimen that does not involve an anthracycline compound and the cardiac function monitored as described.

GENES

Numerous genes are known to be involved in ADME (absorption, distribution, metabolism and elimination), for example UGT1A6, SULT2B1, SLC28A3, SLC28A1, SLC22A17, SLC22A7, ABCB4, and HNMT. Detailed information relating to the sequence, expression patterns, molecular biology, etc of these and related genes in both Homo sapiens and in other model species is known, and may be found at, for example Entrez Gene (<http://www.ncbi.nlm.nih.gov>) and references therein.

UDP glucuronosyltransferase 1 family, polypeptide A6 [Homo sapiens] (UGT1A6) (alternate names include GNT1, HLUGP, HLUGP1, MGC29860, UDPGT, UDPGT 1-6, UGT1, UGT1A6S, UGT1F, phenol-metabolizing UDP-glucuronosyltransferase) maps to chromosome 2q37. Examples of nucleic acid sequences comprising UGT1A6 include those found in the NCBI Entrez Gene database by accession number NM_001072, NC_000002.11. UGT1A6 encodes an enzyme of the glucuronidation pathway that transforms small lipophilic molecules, such as steroids, bilirubin, hormones, and drugs, into water-soluble, excretable metabolites. The UGT1A6 gene is part of a complex locus that encodes several UDP-glucuronosyltransferases and includes 13 unique alternate first exons followed by four common exons. Nine of the 5' exons may be spliced to the four common exons, resulting in nine proteins with different N-termini and identical C-termini and the remaining four alternate first exons are considered pseudogenes. Each first exon encodes the substrate binding site, and is regulated by its own promoter. The enzyme encoded by this gene is active on phenolic and planar compounds and alternative splicing in the unique 5' end of this gene results in two transcript variants.

Sulfotransferase family, cytosolic, 2B, member 1 [Homo sapiens] (SULT2B1) (alternate names include HSST2) maps to chromosome 19q13.3. Examples of nucleic acid sequences comprising SULT2B1 include NC_000019.9, NG_029063.1, and AK096418.1. Sulfotransferase enzymes catalyze the sulfate conjugation of many hormones, neurotransmitters, drugs, and xenobiotic compounds. SULT2B1 is a cytosolic enzyme and has variable tissue distributions and substrate specificities. This gene sulfates

dehydroepiandrosterone but not 4-nitrophenol, a typical substrate for the phenol and estrogen sulfotransferase subfamilies.

Solute carrier family 28 (sodium-coupled nucleoside transporter), member 3' [Homo sapiens] (SLC28A3) (alternate names include concentrative Na⁺-nucleoside cotransporter; concentrative nucleoside transporter 3; CNT3) maps to chromosome 9q22.2. The genomic region (chromosome) can be accessed in the NCBI Entrez Genome database by accession number NC_000009, about nucleotides (complement) 86082912-86173233 (in version NC_000009.10 GI:89161216, genome annotation build 36 version 3). Examples of nucleic acid sequences comprising SLC28A3 include those found in the NCBI Entrez Gene database by accession number NM_022127 (gene ID 64078), and the Ensembl database by gene ID ENSG00000197506. SLC28A3 shows broad specificity for pyrimidine and purine nucleosides.

Nucleoside transporters, such as SLC28A3, regulate multiple cellular processes, including neurotransmission, vascular tone, adenosine concentration in the vicinity of cell surface receptors, and transport and metabolism of nucleoside drugs.

Solute carrier family 28 (sodium-coupled nucleoside transporter), member 1' [Homo sapiens] (SLC28A1) (alternate names include human Organic Cation Transporter 1; hOCT1) maps to chromosome 6q26. The genomic region (chromosome) can be accessed in the NCBI Entrez Genome database by accession number NC_000006.10, about nucleotides (complement) 160462853-160499740. Examples of nucleic acid sequences comprising SLC28A1 include those found in the NCBI Entrez Gene database by accession number U77086 (gene ID 6580), and the Ensembl database by gene ID ENSG00000175003. SLC28A1 is one of three similar cation transporter genes located in a cluster on chromosome 6. Polyspecific organic cation transporters in the liver, kidney, intestine, and other organs are involved in elimination of many endogenous small organic cations as well as a wide array of drugs and environmental toxins. The encoded SLC28A1 protein contains twelve putative transmembrane domains and is a plasma integral membrane protein. Two transcript variants encoding two different isoforms have been found for this gene, but only the longer variant encodes a functional transporter.

Solute carrier family 22, member 17 [Homo sapiens] (SLC22A17) (alternate names include BOCT; BOIT; 24p3R; NGALR; hBOIT; NGALR2; NGALR3) maps to chromosome 14q11.2. Examples of nucleic acid sequences comprising SLC22A17 include NC_000014.8, NM_016609, AJ243653.1 and BC111015.1. Isoform b is encoded by transcript variant 2, solute carrier family 22 (organic cation transporter), member 17 and is reported to be a potent brain type organic ion transporter, NGAL receptor,

neutrophil gelatinase-associated lipocalin receptor, brain-type organic cation transporter, 24p3 receptor, and lipocalin-2 receptor.

Solute carrier family 22, member 7 [Homo sapiens] (SLC22A7) (alternate names include organic anion transporter) maps to chromosome 6p21.1. Examples of nucleic acid sequences comprising SLC22A7 include NM_153320, NM_006672.3, EU562669.1 and BC017963.1. SLC22A7 mediates the uptake of organic anions such as salicylate, acetylsalicylate, prostaglandin E2, dicarboxylates, and p-aminohippurate. Salicylate uptake is saturable and sodium-independent. SLC22A7 is thought to be a multispecific organic anion transporter of the liver.

ATP-binding cassette, sub-family B (MDR/TAP), member 4 [Homo sapiens] (ABCB4) (alternate names include P-GLYCOPROTEIN 3; PGY3, MULTIDRUG RESISTANCE 3; MDR3), maps to chromosome 7q21.12 and ABCB4 gene has been determined to have 28 exons over 74 kb. Examples of nucleic acid sequences comprising ABCB4 include NM_018849.2 (variant B), NM_000443.3(variant A), NM_018850.2 (variant C) and NG_007118.1. P-glycoproteins are thought to act as pumps for the extrusion of drugs from cells at the cost of ATP hydrolysis. P-glycoproteins belong to a class of vectorial transport proteins known as the ATP-binding cassette transporter proteins. These pumps are thought to be a defense against xenobiotic. The P-glycoproteins have 2 homologous halves, each with 6 hydrophobic segments adjacent to a consensus sequence for nucleotide binding. The hydrophobic segments are thought to form a membrane channel, whereas the nucleotide binding site may be involved in energization of drug transport.

Histamine n-methyltransferase [Homo sapiens] (HNMT) maps to chromosome 2q22.1. Examples of nucleic acid sequences comprising HNMT include NG_012966, NM_001024074.2 (variant 2), NM_001024075.1 (variant 3), NM_006895.2 (variant 1), and AH012839.2. Histamine is metabolized in most mammals by HNMT and diamine oxidase (DAO), but the relative contributions of these 2 enzymes to histamine metabolism is different depending on the tissue. HNMT plays the dominant role in histamine biotransformation in bronchial epithelium and has been associated with susceptibility to asthma. There are large individual variations of HNMT activity in human tissues. Biochemical genetic studies of red blood cell HNMT has demonstrated that 5-fold differences among individuals in levels of HNMT activity are due primarily to the effects of inheritance.

EXAMPLES

EXAMPLE 1 – multi-SNP Panel

Methods for EXAMPLE 1

Samples

Study participants were recruited through the Canadian Pharmacogenomics Network for Drug Safety (CPNDS), a multicenter active surveillance consortium studying adverse drug reactions in children (Carleton, B. *et al.* *Pharmacoepidemiol Drug Saf* (2009) 18:713-721). The current replication cohort for this example ($n=218$) consisted of a new Canadian CPNDS replication cohort that was recruited from pediatric oncology units and long-term follow-up clinics across Canada between February 2010 and April 2011 and the Dutch-EKZ cohort that was recruited at the Emma Children's Hospital/Academic Medical Centre in Amsterdam, the Netherlands, between July 2009 and April 2011 (van der Pal, H.J. *et al.* *Arch Intern Med* (2010) 170:1247-1255 ; and van Dalen, EC. *et al.* *Eur J Cancer* (2006) 42:3191-3198). This replication cohort was used to replicate the previous genetic findings and to test the prediction models (see below). The previous cohort, which consisted of the two initial Canadian cohorts (discovery and replication) (Visscher, H. *et al.* *J Clin Oncol* (2011) Epub 11 Oct 2011), was used in the combined analysis and training of the prediction model (see below) and has been described previously (Visscher, H. *et al.* *J Clin Oncol* (2011) Epub 11 Oct 2011).

Study cohorts consisted of patients who developed cardiotoxicity during or after treatment with anthracyclines for childhood cancer (cases) and patients who received anthracyclines, but did not show cardiotoxicity (controls). In the Dutch-EKZ cohort cases and controls were matched for age, gender and cumulative dose where possible and follow-up time in controls was matched with time to first available echocardiogram showing cardiotoxicity in cases.

ACT was defined as early- or late-onset left ventricular dysfunction measured by echocardiogram (shortening fraction, SF) and/or symptoms requiring intervention based on CTCAEv3 (Common Terminology Criteria for Adverse Events) (Cancer Therapy Evaluation Program - Common Terminology Criteria for Adverse Events - Version 3. In Edition 2003). To improve differentiation between cardiotoxicity cases and controls a more stringent SF threshold of $\leq 26\%$ at any time during or after anthracycline therapy was used to define cardiotoxicity. Only echocardiograms obtained ≥ 21 days after a dose of anthracyclines were used to exclude transient acute cardiotoxicity. Control patients were defined as those having normal echocardiograms ($SF \geq 30\%$) during and after therapy, with a follow-up of >5 years after completion of anthracycline therapy. Cumulative anthracycline doses were calculated using doxorubicin equivalents (Altman, A.J. *Children's Oncology Group. Supportive care of children with*

cancer: current therapy and guidelines from the Children's Oncology Group. Baltimore: Johns Hopkins University Press 2004).

Written informed consent or assent was obtained from each subject or their parents or legal guardians. The study was approved by the ethics committees of all participating universities and hospitals.

Genotyping

Twenty-three SNPs were selected that previously showed evidence of association ($P < 0.01$) with ACT (Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011). Genomic DNA was extracted from blood, saliva or buccal swabs using the QIAamp™ DNA purification system (Qiagen™, Canada). DNA samples were genotyped using a custom 96-plex Illumina Veracode GoldenGate™ SNP genotyping assay according to manufacturer's instructions (Illumina™, San Diego, USA). This assay included an additional 63 non-study SNPs used for quality control (QC) purposes only. All SNP genotypes were manually clustered using Illumina GenomeStudio™ software. One SNP in *FMO2* (rs2020870) could not be reliably clustered and was therefore removed from further analyses. Two non-study SNPs with a SNP call rate of $< 95\%$ were also removed. Sixteen samples (2 cases and 14 controls) with a call rate of $< 95\%$ were removed. The remaining 202 samples had an average call rate of 99.7%. All 22 remaining study SNPs had a SNP call rate of $> 98\%$ (mean 99.7%) and were in Hardy-Weinberg equilibrium ($P > 0.05$).

Power calculation and statistical analysis

We calculated *a priori* to have 61-92% (mean 72%) statistical power to detect an association based on the previous effect sizes and allele frequencies and a two-sided type I error rate of 0.05 using QUANTO™ v1.2.4 software (Gauderman, W. and Morrison, J. QUANTO 1.1™: A computer program for power and sample size calculations for genetic-epidemiology studies, <http://hydra.usc.edu/gxe>. In Edition 2006).

Clinical variables in cases and controls were compared using Fisher's exact test for proportional variables and Wilcoxon-Mann-Whitney rank-sum test for continuous variables. Hardy-Weinberg Equilibrium of the polymorphisms was tested using Fisher's Exact test in controls. The primary association test was a single SNP test assuming an additive genetic model using logistic regression with cumulative anthracycline dose, age, gender and radiation therapy to the heart included as covariates. As we had previously noted an effect of *SLC28A3* rs7853758 only in patients who received doxorubicin and/or daunorubicin (Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011), we decided *a priori* to conduct our primary analysis in the 192 (out of 218) patients who received doxorubicin and/or daunorubicin – of

these, 177 were successfully genotyped. In order to have sufficient power to replicate the previous associations, the Dutch-EKZ and new Canadian CPNDS replication cohort were analyzed together. We also combined the initial cohorts (Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011) and the current new cohorts to show the overall significance. Secondary association tests were stratified analyses that included clinical covariates not used to stratify (e.g. stratification by age-group did not include age as a covariate). For continuous variables, the median was used to define the subgroups. Heterogeneity between groups was assessed using Cochran's Q-statistic. A p-value of <0.05 was considered to be statistically significant.

Multivariate logistic regression models including multiple genetic variants and/or clinical variables were all trained in the initial Canadian discovery and replication cohort combined (previous cohort) (Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011) and tested in the current replication cohort (Dutch and new Canadian patients combined). Risk scores were calculated by multiplying each variable with the estimated beta (log odds ratio) from the training cohort. The previous full model was constructed using step-wise regression with forward selection, where 9 SNPs with $P<0.01$ were retained in the final model (Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011). The revised model was constructed to include the 5 SNPs that showed an effect in the same direction in the current replication cohort and that were more significant in the combined analysis than before as well as the clinical variables gender, age, anthracycline dose and radiation to the heart. One SNP in SLC28A3 (rs885004) was not included in this model as it was in high linkage disequilibrium (LD) with another SNP (rs7853758) as previously shown (Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011). Models were assessed by constructing Receiver Operating Characteristic (ROC) curves and calculating the c-statistic (Area Under the Curve - AUC) using the risk scores from the model and the actual value (case or control).

Statistical analyses were conducted using SNP and Variation Suite 7.4.5 (Golden Helix™, Bozeman, USA) and R 2.13.0 (R Development Core Team).

Patient baseline characteristics are provided in Table 3 for both the Dutch-EKZ and new Canadian replication cohort separately. Cumulative anthracycline doses were significantly different between cases and controls in both cohorts ($P=0.0071$ and $P=0.00017$, respectively) – particularly in the Canadian cohort in which all cases received a higher dose than the median dose in controls. In addition, controls were significantly younger in the Canadian cohort ($P=0.013$). Furthermore, there were more acute myeloid leukemia among Canadian cases ($P=0.0069$), although the overall numbers were low. Finally, follow-up was significantly longer in cases in the Dutch-EKZ cohort (21.3 versus 16.8 years, $P=0.012$).

TABLE 3. Subject Demographics Replication Cohorts

	Dutch-EKZ (n=128)			Canadian-CPNDS (n=90)		
	Cardiotoxicity (n=44)	Controls (n=84)	P-value	Cardiotoxicity (n=12)	Controls (n=78)	P-value
Age in yrs, median (range)	9.1 (0.5-16.8)	11.2 (1.8-17.7)	0.30	12.6 (0.9-17.0)	4.9 (0.5-16.0)	0.013
Gender, no. female (%)	21 (48%)	40 (48%)	1.00	4 (33%)	47 (60%)	0.12
Dose in mg/m ² , median (range) ^a	360 (100-720)	280 (50-720)	0.0071	300 (175-550)	150 (50-540)	0.00017
Anthracycline type ^b , no. (%)						
Doxorubicin	28 (64%)	46 (55%)	0.35	7 (58%)	59 (76%)	0.29
Daunorubicin	2 (5%)	6 (7%)	0.71	2 (17%)	9 (12%)	0.64
Doxorubicin plus daunorubicin	2 (5%)	6 (7%)	0.71	1 (8%)	7 (9%)	1.00
Doxorubicin plus other	3 (7%)	8 (10%)	0.75	0 (0%)	1 (1%)	1.00
Daunorubicin plus other	0 (0%)	1 (1%)	1.00	2 (17%)	2 (3%)	0.08
Doxorubicin, daunorubicin plus other	0 (0%)	0 (0%)	1.00	0 (0%)	0 (0%)	1.00
Epirubicin	6 (14%)	14 (17%)	0.80	0 (0%)	0 (0%)	1.00
Epirubicin plus other	3 (7%)	2 (2%)	0.34	0 (0%)	0 (0%)	1.00
Other	0 (0%)	1 (1%)	1.00	0 (0%)	0 (0%)	1.00
Tumor type, no. (%)						
Acute Lymphoblastic Leukemia	10 (23%)	14 (17%)	0.48	3 (25%)	29 (37%)	0.53
Acute Myelogenous Leukemia	0 (0%)	7 (8%)	0.09	3 (25%)	1 (1%)	0.0069
Other Leukemia	0 (0%)	1 (1%)	1.00	1 (8%)	2 (3%)	0.35
Hodgkin's Lymphoma	4 (9%)	10 (12%)	0.77	0 (0%)	9 (12%)	0.60
Non-Hodgkin's Lymphoma	10 (23%)	19 (23%)	1.00	0 (0%)	7 (9%)	0.59
Osteosarcoma	3 (7%)	11 (13%)	0.38	0 (0%)	2 (3%)	1.00
Rhabdomyosarcoma	4 (9%)	4 (5%)	0.45	0 (0%)	3 (4%)	1.00
Ewing's sarcoma	6 (14%)	5 (6%)	0.19	1 (8%)	2 (3%)	0.35
Other sarcoma	2 (5%)	1 (1%)	0.27	0 (0%)	0 (0%)	1.00
Nephroblastoma	3 (7%)	11 (13%)	0.38	3 (25%)	12 (15%)	0.41
Hepatoblastoma	0 (0%)	0 (0%)	1.00	1 (8%)	1 (1%)	0.25
Neuroblastoma	0 (0%)	0 (0%)	1.00	0 (0%)	10 (13%)	0.35
Carcinoma	2 (5%)	0 (0%)	0.12	0 (0%)	0 (0%)	1.00
Germ Cell Tumor	0 (0%)	1 (1%)	1.00	0 (0%)	0 (0%)	1.00
Radiotherapy involving heart, no. (%)	9 (20%)	19 (23%)	0.83	4 (33%)	18 (23%)	0.48
Follow-up in yrs, median (range)	21.3 (7.4-28.5)	16.8 (5.0-31.6)	0.012	6.8 (0.4-27.2)	7.4 (5.0-23.1)	0.81

Patient characteristics for the Dutch-EKZ and new Canadian replication cohort separately. For age, dose and follow-up, the Wilcoxon-Mann-Whitney test with normal approximation was used. For gender, anthracycline type, tumor type and radiotherapy involving the heart region, the Fisher exact test was used. In bold are statistically significant values at $p < 0.05$.^aCumulative anthracycline dose in doxorubicin isotoxic equivalent doses. ^bOther anthracycline type included idarubicin, epirubicin or mitoxantrone.

TABLE 4. Association between SNPs and Anthracycline-Induced Cardiotoxicity

SNP rs-ID	Gene	Allele ^a	Previous cohort (n=344) 78 cases, 266 controls		Current replication (n=177) 46 cases, 131 controls		Combined (n=521) 124 cases, 397 controls	
			OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value
rs7853758	SLC28A3	A/G	0.31 (0.16 - 0.60)	0.00010	0.46 (0.20 - 1.08)	0.058	+ 0.36 (0.22 - 0.60)	1.6E-05

rs885004	SLC28A3	A/G	0.31 (0.15 - 0.62)	0.00021	0.42 (0.16 - 1.10)	0.058 + 0.34 (0.20 - 0.60)	3.0E-05
rs17863783	UGT1A6	A/C	3.68 (1.45 - 9.30)	0.0059	7.98 (1.85 - 34.4)	0.0062 + 4.30 (1.97 - 9.36)	2.4E-04
rs10426377	SULT2B1	A/C	0.54 (0.34 - 0.86)	0.0071	0.52 (0.26 - 1.04)	0.054 + 0.56 (0.38 - 0.81)	0.0015
rs2305364	SLC28A1	A/G	1.76 (1.20 - 2.58)	0.0033	1.48 (0.88 - 2.51)	0.14 + 1.60 (1.18 - 2.17)	0.0020
rs4148350	ABCC1	A/C	3.44 (1.65 - 7.15)	0.0012	1.29 (0.48 - 3.47)	0.61 + 2.40 (1.33 - 4.33)	0.0040
rs17645700	HNMT	G/A	0.46 (0.26 - 0.82)	0.0053	0.70 (0.36 - 1.36)	0.29 + 0.56 (0.37 - 0.86)	0.0054
rs9514091	SLC10A2	A/G	0.43 (0.23 - 0.78)	0.0033	0.77 (0.41 - 1.46)	0.42 + 0.57 (0.38 - 0.87)	0.0063
rs4148808	ABCB4	A/G	1.86 (1.17 - 2.96)	0.0093	1.41 (0.72 - 2.77)	0.33 + 1.67 (1.15 - 2.43)	0.0072
rs17583889	HNMT	A/C	1.91 (1.21 - 3.02)	0.0057	1.26 (0.65 - 2.46)	0.50 + 1.67 (1.15 - 2.41)	0.0073
rs2290271	SLC28A1	C/A	0.56 (0.37 - 0.83)	0.0035	0.80 (0.46 - 1.39)	0.43 + 0.66 (0.48 - 0.91)	0.0098
rs1736557	FMO3	A/G	0.33 (0.13 - 0.81)	0.0060	0.67 (0.25 - 1.79)	0.41 + 0.47 (0.25 - 0.87)	0.011
rs2019604	SPG7	C/A	0.39 (0.20 - 0.76)	0.0021	0.84 (0.39 - 1.80)	0.64 + 0.56 (0.35 - 0.90)	0.012
rs7319981	SLC10A2	A/G	0.51 (0.32 - 0.81)	0.0029	0.93 (0.54 - 1.59)	0.78 + 0.66 (0.47 - 0.93)	0.016
rs4261716	UGT1A6	A/C	1.76 (1.19 - 2.59)	0.0043	1.02 (0.60 - 1.74)	0.93 + 1.44 (1.06 - 1.95)	0.018
rs6759892	UGT1A6	C/A	1.77 (1.20 - 2.61)	0.0038	0.99 (0.58 - 1.69)	0.96 1.43 (1.05 - 1.94)	0.022
rs2235047	ABCB1	C/A	2.92 (1.31 - 6.49)	0.0087	1.34 (0.51 - 3.49)	0.56 + 1.79 (1.05 - 3.04)	0.036
rs4877847	SLC28A3	A/C	0.60 (0.41 - 0.89)	0.0092	1.06 (0.61 - 1.84)	0.83 0.73 (0.54 - 0.98)	0.037
rs729147	ADH7	G/A	1.86 (1.18 - 2.93)	0.0072	0.94 (0.50 - 1.74)	0.83 1.43 (1.02 - 2.01)	0.041
rs1149222	ABCB4	C/A	1.87 (1.20 - 2.92)	0.0054	0.89 (0.49 - 1.61)	0.69 1.36 (0.97 - 1.90)	0.075
rs2108623	CYP4F11	G/A	0.57 (0.38 - 0.86)	0.0055	1.17 (0.68 - 1.99)	0.57 0.77 (0.57 - 1.04)	0.084
rs316019	SLC22A2	A/C	0.40 (0.20 - 0.81)	0.0049	1.90 (0.89 - 4.05)	0.10 0.75 (0.46 - 1.21)	0.23

Results from initial cohorts from previous study combined (previous cohort) compared to the current replication

cohort as well as all cohorts combined. In the current replication cohort, only patients that received doxorubicin and/or daunorubicin were included. Odds ratios are per copy of the minor allele. Plus (+) sign indicates result in same direction in replication cohort. Star (*) sign indicates result more significant in combined cohort than in previous cohort. ^aSNP alleles assayed; minor allele is mentioned first; P-values in bold are significant after multiple testing correction in combined cohort; SNP, Single Nucleotide Polymorphism; repl, replication; OR, Odds Ratio; CI, Confidence Interval.

The present results confirm the previous association between rs17863783 in UGT1A6 and ACT (P=0.0062; odds ratio (OR) 7.98 - Table 4). Furthermore, two SNPs in SLC28A3 (rs7853758 and rs885004) and one in SULT2B1 (rs10426377) were close to being significantly associated in the current replication cohort (P=0.058, P=0.058 and P=0.054, respectively). The two SLC28A3 SNPs were also significantly associated with ACT in a combined analysis of the initial and current cohorts after applying the same threshold for multiple testing as previously defined (P<0.00015) (Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011), while rs17863783 showed a trend (P corrected = 0.078). In total, 16 out of 22 SNPs showed an effect in the same direction (Table 4), which is significantly higher than expected by chance alone (P=0.026). Six SNPs were also more significantly associated in all cohorts combined than in the previous cohort (Table 4). Including patients that did not receive doxorubicin or

daunorubicin, but received other anthracyclines instead, yielded similar, though slightly less significant results (see Table 8).

We assessed whether our definition of cardiotoxicity using a threshold of $SF \leq 26\%$ influenced our results by performing a subgroup analysis of patients with more severe cardiotoxicity ($SF < 24\%$ or symptoms, CTCAE grade 2-4). Again, similar results were obtained with comparable effect sizes (results not shown). Interestingly, in the combined cohort of severe cardiotoxicity, UGT1A6 rs17863783 now was significantly associated with ACT after correcting for multiple testing ($P_{\text{corrected}} = 0.036$).

Next, we explored whether the effects of these variants were influenced by gender, age or cumulative anthracycline dose (effect-size heterogeneity) which could explain some of the differences in susceptibility to ACT. In females, two variants in ABCB4 (rs4148808 and rs1149222) were associated with increased risk for ACT ($P=0.00067$ and $P=0.0024$ respectively), while these variants had no effect in males ($P=0.81$ and $P=0.65$) which was significantly different ($P_{\text{het}}=0.028$ and $P_{\text{het}}=0.012$, respectively – Table 5). Conversely, SULT2B1 variant rs10426377 only had an effect in males ($P=0.00021$), but not in females ($P=0.42$; $P_{\text{het}}=0.033$). None of the other variants showed significant gender-specific effects ($P>0.05$). Similarly, in younger children (<5.3 years), HNMT variant rs17583889 was associated with increased ACT ($P=0.00025$), while this effect was not detected in older children ($P=0.59$; $P_{\text{het}}=0.0022$ – Table 6). A similar pattern was observed for rs316019 in SLC22A2 ($P=0.0021$ versus $P=0.59$, $P_{\text{het}}=0.016$), though this SNP did not replicate in the current replication cohort (Table 4). No statistically significant different effect-sizes were detected for any genetic variant between higher and lower cumulative dose groups (results not shown).

TABLE 5. Association between SNPs and anthracycline-induced cardiotoxicity by gender

SNP rs-ID	Gene	Allele ^a	Female (n=248)		Male (n=273)		P _{het} -value
			59 cases, 189 controls		65 cases, 208 controls		
			OR (95% CI)	P-value	OR (95% CI)	P-value	
rs1149222	ABCB4	C/A	2.18 (1.31-3.60)	0.0024	0.89 (0.56-1.44)	0.65	0.012
rs4148808	ABCB4	A/G	2.53 (1.47-4.36)	0.00067	1.07 (0.62-1.85)	0.81	0.028
rs10426377	SULT2B1	A/C	0.82 (0.50-1.34)	0.42	0.35 (0.20-0.64)	0.00021	0.033
rs1736557	FMO3	A/G	0.76 (0.33-1.76)	0.52	0.25 (0.08-0.72)	0.0024	0.10
rs4148350	ABCC1	A/C	1.33 (0.47-3.80)	0.60	3.18 (1.48-6.82)	0.0027	0.19
rs17583889	HNMT	A/C	1.34 (0.76-2.33)	0.31	2.19 (1.30-3.69)	0.0029	0.20
rs2108623	CYP4F11	G/A	0.92 (0.60-1.41)	0.70	0.64 (0.41-0.99)	0.041	0.25
rs885004	SLC28A3	A/G	0.48 (0.22-1.04)	0.048	0.27 (0.12-0.62)	0.00028	0.33
rs2019604	SPG7	C/A	0.69 (0.34-1.38)	0.27	0.45 (0.23-0.87)	0.012	0.39

rs4261716	UGT1A6	A/C	1.27 (0.82-1.96)	0.28	1.62 (1.05-2.49)	0.029	0.44
rs7853758	SLC28A3	A/G	0.45 (0.22-0.92)	0.019	0.31 (0.15-0.64)	0.00034	0.45
rs6759892	UGT1A6	C/A	1.26 (0.82-1.95)	0.29	1.59 (1.03-2.46)	0.036	0.47
rs729147	ADH7	G/A	1.54 (0.97-2.44)	0.069	1.19 (0.71-2.00)	0.51	0.47
rs316019	SLC22A2	A/C	0.65 (0.33-1.28)	0.19	0.91 (0.46-1.81)	0.79	0.49
rs2235047	ABCB1	C/A	2.18 (1.01-4.72)	0.053	1.54 (0.73-3.25)	0.26	0.53
rs9514091	SLC10A2	A/G	0.50 (0.27-0.93)	0.022	0.62 (0.35-1.09)	0.089	0.61
rs17863783	UGT1A6	A/C	5.00 (1.56-16.0)	0.0047	3.62 (1.22-10.7)	0.025	0.69
rs4877847	SLC28A3	A/C	0.74 (0.48-1.15)	0.18	0.68 (0.45-1.04)	0.074	0.79
rs17645700	HNMT	G/A	0.60 (0.33-1.08)	0.078	0.54 (0.30-0.99)	0.037	0.83
rs2290271	SLC28A1	C/A	0.69 (0.44-1.09)	0.10	0.65 (0.41-1.02)	0.055	0.85
rs7319981	SLC10A2	A/G	0.67 (0.41-1.11)	0.11	0.65 (0.40-1.04)	0.068	0.91
rs2305364	SLC28A1	A/G	1.59 (1.02-2.47)	0.038	1.58 (1.04-2.41)	0.032	0.99

Results from all cohorts combined. Odds ratios are per copy of the minor allele. In bold are statistically significant values at $P < 0.05$. *SNP alleles assayed; minor allele is mentioned first; SNP, Single Nucleotide Polymorphism; OR, Odds Ratio; CI, Confidence Interval; P_{het} , P-value for heterogeneity.

TABLE 6. Association between SNPs and anthracycline-induced cardiotoxicity by age-group

SNP rs-ID	Gene	Allele ^a	<5.3 years (n=257)		≥5.3 years (n=264)		P _{het} -value
			45 cases, 212 controls		79 cases, 185 controls		
			OR (95% CI)	P-value	OR (95% CI)	P-value	
rs17583889	HNMT	A/C	3.53 (1.79-6.93)	0.00025	1.13 (0.72-1.79)	0.59	0.0022
rs316019	SLC22A2	A/C	0.16 (0.03-0.71)	0.0021	1.16 (0.68-2.00)	0.59	0.016
rs2235047	ABCB1	C/A	1.97 (0.81-4.76)	0.14	1.42 (0.73-2.74)	0.31	0.069
rs4148350	ABCC1	A/C	4.66 (1.85-11.7)	0.0015	1.57 (0.74-3.33)	0.24	0.089
rs17645700	HNMT	G/A	0.36 (0.16-0.81)	0.0062	0.67 (0.40-1.11)	0.11	0.13
rs1736557	FMO3	A/G	0.85 (0.28-2.52)	0.76	0.39 (0.18-0.84)	0.0087	0.15
rs9514091	SLC10A2	A/G	0.41 (0.19-0.90)	0.017	0.69 (0.42-1.14)	0.14	0.30
rs1149222	ABCB4	C/A	1.44 (0.82-2.54)	0.21	1.22 (0.80-1.87)	0.36	0.34
rs6759892	UGT1A6	C/A	1.26 (0.77-2.05)	0.35	1.50 (1.01-2.23)	0.046	0.38
rs4261716	UGT1A6	A/C	1.27 (0.78-2.07)	0.33	1.53 (1.03-2.27)	0.036	0.40
rs4148808	ABCB4	A/G	1.43 (0.77-2.67)	0.26	1.66 (1.04-2.66)	0.036	0.52
rs2290271	SLC28A1	C/A	0.53 (0.29-0.97)	0.033	0.73 (0.50-1.08)	0.11	0.52
rs2108623	CYP4F11	G/A	0.63 (0.37-1.07)	0.084	0.87 (0.60-1.25)	0.45	0.52
rs2019604	SPG7	C/A	0.49 (0.19-1.22)	0.099	0.61 (0.35-1.08)	0.077	0.60
rs7319981	SLC10A2	A/G	0.61 (0.34-1.08)	0.080	0.72 (0.47-1.10)	0.12	0.66
rs4877847	SLC28A3	A/C	0.63 (0.37-1.07)	0.080	0.75 (0.52-1.10)	0.14	0.66
rs7853758	SLC28A3	A/G	0.34 (0.13-0.89)	0.014	0.37 (0.20-0.68)	0.00043	0.68
rs10426377	SULT2B1	A/C	0.62 (0.34-1.14)	0.11	0.55 (0.35-0.87)	0.0080	0.74
rs17863783	UGT1A6	A/C	2.63 (0.68-10.2)	0.18	5.51 (1.93-15.7)	0.00070	0.78
rs885004	SLC28A3	A/G	0.44 (0.16-1.20)	0.080	0.31 (0.16-0.62)	0.00018	0.82
rs729147	ADH7	G/A	1.41 (0.80-2.49)	0.24	1.46 (0.95-2.24)	0.087	0.92

rs2305364 SLC28A1 A/G 1.81 (1.07-3.06) 0.025 1.45 (1.00-2.11) 0.051 0.98

Results from all cohorts combined. Odds ratios are per copy of the minor allele. In bold are statistically significant values at $P < 0.05$. ^aSNP alleles assayed; minor allele is mentioned first; SNP, Single Nucleotide Polymorphism; OR, Odds Ratio; CI, Confidence Interval; P_{het} , P-value for heterogeneity.

An earlier risk prediction model incorporated multiple genetic variants as well as clinical risk factors. To assess the ability of this model to discriminate between cases and controls, we applied this model to the current replication cohort and performed ROC analyses (Table 7). We revised the previous model to include only SNPs that replicated in the same direction and were more significant in the combined cohort (Table 4). The revised model, that included 5 SNPs and clinical variables, was trained in the previous cohort and then tested in the current replication cohort. In the training set, the full (clinical plus genetic) revised model performed better than the clinical-only model (AUC 0.771 versus 0.685; $P = 0.0031$; Table 7). In the test set (current replication cohort), similar metrics were obtained with the full model discriminating better between cases and controls than the clinical-only model (AUC 0.767 versus 0.688) which was close to being significant ($P = 0.060$).

TABLE 7. AUCs of different models in training and test sets

Model	AUC (95% CI)	
	Training set	Test set
	Previous cohort	Current Replication
<i>Previous model</i>		
Clinical only	0.679 (0.614-0.745)	0.670 (0.584-0.755)
Genetic only (9 SNPs)	0.814 (0.764-0.865)	0.570 (0.473-0.666)
Full (clinical + genetic)	0.867 (0.820-0.913)	0.671 (0.579-0.763)
<i>Revised model</i>		
Clinical only	0.685 (0.619-0.751)	0.688 (0.604-0.772)
Genetic only (5 SNPs)	0.711 (0.646-0.776)	0.651 (0.561-0.741)
Full (clinical + genetic)	0.771 (0.710-0.831)	0.767 (0.688-0.846)

AUCs for different models in training set (previous cohort) and test set (current replication cohort). Clinical only model includes age at start of treatment, cumulative dose, gender, radiation therapy involving the heart region. Full model includes the clinical variables as well as the genetic variants, while the genetic-only model contains only the genetic variants. AUC, Area Under the Curve; CI, Confidence Interval; disc., discovery; repl., replication; SNP, Single Nucleotide Polymorphism.

It is shown herein that combining the replicated variants with clinical risk factors in an optimized risk prediction model allowed for better discrimination between cases and controls than clinical risk factors alone in both the training as well as the test cohort.

TABLE 8. Association between SNPs and Anthracycline-Induced Cardiotoxicity Including All Anthracycline Types

SNP rs- ID	Gene	Allele ^a	Original cohort (n=344) 78 cases, 266 controls		Replication cohort (n=202) 54 cases, 148 controls		Combined (n=546) 132 cases, 414 controls	
			OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value
rs885004	SLC28A3	A/G	0.31 (0.15 - 0.62)	0.00021	0.54 (0.23 - 1.26)	0.14 +	0.38 (0.22 - 0.65)	9.4x10⁻⁵ *
rs17863783	UGT1A6	A/C	3.68 (1.45 - 9.30)	0.0059	7.47 (1.77 - 31.5)	0.0071+	4.29 (1.97 - 9.33)	2.3x10⁻⁴ *
rs10426377	SULT2B1	A/C	0.54 (0.34 - 0.86)	0.0071	0.81 (0.45 - 1.46)	0.47 +	0.64 (0.45 - 0.91)	0.012
rs2305364	SLC28A1	A/G	1.76 (1.20 - 2.58)	0.0033	1.31 (0.81 - 2.13)	0.26 +	1.54 (1.15 - 2.07)	0.0037
rs4148350	ABCC1	A/C	3.44 (1.65 - 7.15)	0.0012	1.60 (0.64 - 3.99)	0.31 +	2.53 (1.42 - 4.49)	0.0017 *
rs17645700	HNMT	G/A	0.46 (0.26 - 0.82)	0.0053	0.82 (0.44 - 1.50)	0.51 +	0.61 (0.41 - 0.91)	0.013
rs9514091	SLC10A2	A/G	0.43 (0.23 - 0.78)	0.0033	0.87 (0.49 - 1.57)	0.65 +	0.61 (0.41 - 0.91)	0.012
rs4148808	ABCB4	A/G	1.86 (1.17 - 2.96)	0.0093	1.46 (0.79 - 2.70)	0.23 +	1.69 (1.18 - 2.42)	0.0049
rs17583889	HNMT	A/C	1.91 (1.21 - 3.02)	0.0057	1.18 (0.63 - 2.20)	0.60 +	1.62 (1.13 - 2.33)	0.0096
rs2290271	SLC28A1	C/A	0.56 (0.37 - 0.83)	0.0035	0.84 (0.50 - 1.40)	0.50 +	0.67 (0.49 - 0.92)	0.012
rs1736557	FMO3	A/G	0.33 (0.13 - 0.81)	0.0060	0.87 (0.37 - 2.06)	0.75 +	0.53 (0.30 - 0.96)	0.026
rs2019604	SPG7	C/A	0.39 (0.20 - 0.76)	0.0021	0.65 (0.31 - 1.34)	0.23 +	0.51 (0.32 - 0.81)	0.0030 *
rs7319981	SLC10A2	A/G	0.51 (0.32 - 0.81)	0.0029	1.07 (0.65 - 1.77)	0.79	0.72 (0.52 - 1.00)	0.047
rs4261716	UGT1A6	A/C	1.76 (1.19 - 2.59)	0.0043	1.05 (0.64 - 1.71)	0.86 +	1.43 (1.06 - 1.92)	0.019
rs6759892	UGT1A6	C/A	1.77 (1.20 - 2.61)	0.0038	1.01 (0.62 - 1.66)	0.96 +	1.42 (1.05 - 1.91)	0.022
rs2235047	ABCB1	C/A	2.92 (1.31 - 6.49)	0.0087	1.35 (0.52 - 3.47)	0.54 +	1.80 (1.06 - 3.05)	0.034
rs4877847	SLC28A3	A/C	0.60 (0.41 - 0.89)	0.0092	1.24 (0.74 - 2.06)	0.41	0.70 (0.52 - 0.94)	0.016
rs729147	ADH7	G/A	1.86 (1.18 - 2.93)	0.0072	0.94 (0.53 - 1.68)	0.84	1.39 (1.00 - 1.94)	0.053
rs1149222	ABCB4	C/A	1.87 (1.20 - 2.92)	0.0054	1.02 (0.59 - 1.75)	0.95 +	1.40 (1.01 - 1.94)	0.046
rs2108623	CYP4F11	G/A	0.57 (0.38 - 0.86)	0.0055	1.35 (0.83 - 2.19)	0.22	0.83 (0.62 - 1.11)	0.22
rs316019	SLC22A2	A/C	0.40 (0.20 - 0.81)	0.0049	1.68 (0.82 - 3.43)	0.16	0.74 (0.46 - 1.19)	0.20

Results from original cohort compared to replication cohort and all cohorts combined. In the replication cohort, also patients that did not receive doxorubicin and/or daunorubicin, but other anthracyclines instead, were included. ORs are per copy of the minor allele. Plus (+) sign indicates result in same direction in replication cohort. Star (*) sign indicates result more significant in combined cohort than in original cohort. ^aSNP alleles assayed; minor allele mentioned first; P-values in bold are significant after multiple testing correction in combined cohort; SNP, Single Nucleotide Polymorphism; OR, Odds Ratio; CI, Confidence Interval.

When patients that did not receive doxorubicin or daunorubicin were included, but received other anthracyclines, results were found (see Table 8).

An assessment was also made to determine whether the definition of cardiotoxicity using a threshold of $SF \leq 26\%$ influenced the results by performing a subgroup analysis of patients with more severe cardiotoxicity ($SF < 24\%$ or symptoms, CTCAE grade 2-4). Again, similar results were obtained with comparable effect sizes (Table 9). Interestingly, in the combined cohort of severe cardiotoxicity, UGT1A6 rs17863783 now was significantly associated with ACT after correcting for multiple testing ($P = 1.1 \times 10^{-4}$; OR 6.22; P -corrected = 0.036).

TABLE 9. Association between SNPs and Anthracycline-Induced Cardiotoxicity Grade 2-4

			Combined ($n=462$) 65 cases, 397 controls	
SNP rs-ID	Gene	Allele ^a	OR (95% CI)	P-value
rs7853758	SLC28A3	A/G	0.37 (0.19 - 0.71)	9.6×10^{-4}
rs885004	SLC28A3	A/G	0.36 (0.17 - 0.75)	0.0021
rs17863783	UGT1A6	A/C	6.22 (2.52 - 15.4)	1.1×10^{-4}
rs10426377	SULT2B1	A/C	0.58 (0.36 - 0.95)	0.024
rs2305364	SLC28A1	A/G	1.99 (1.34 - 2.97)	5.6×10^{-4}
rs4148350	ABCC1	A/C	1.89 (0.87 - 4.11)	0.12
rs1764570	HNMT	G/A	0.77 (0.47 - 1.26)	0.29
rs9514091	SLC10A2	A/G	0.70 (0.42 - 1.17)	0.16
rs4148808	ABCB4	A/G	1.79 (1.12 - 2.87)	0.018
rs17583889	HNMT	A/C	1.88 (1.19 - 2.96)	0.0077
rs2290271	SLC28A1	C/A	0.54 (0.35 - 0.84)	0.0046
rs1736557	FMO3	A/G	0.18 (0.06 - 0.60)	3.5×10^{-4}
rs2019604	SPG7	C/A	0.46 (0.23 - 0.90)	0.014
rs7319981	SLC10A2	A/G	0.72 (0.47 - 1.11)	0.13
rs4261716	UGT1A6	A/C	1.67 (1.12 - 2.47)	0.011
rs6759892	UGT1A6	C/A	1.65 (1.11 - 2.45)	0.013
rs2235047	ABCB1	C/A	1.91 (0.99 - 3.67)	0.064
rs4877847	SLC28A3	A/C	0.58 (0.39 - 0.87)	0.0078
rs729147	ADH7	G/A	1.59 (1.03 - 2.44)	0.038
rs1149222	ABCB4	C/A	1.23 (0.80 - 1.90)	0.35
rs2108623	CYP4F11	G/A	0.77 (0.52 - 1.15)	0.20
rs316019	SLC22A2	A/C	0.71 (0.37 - 1.35)	0.28

Results from all cohorts combined. OR are per copy of the minor allele. ^aSNP alleles assayed; minor allele is mentioned first; P-values in bold are significant after multiple testing correction; SNP, Single Nucleotide Polymorphism; OR, Odds Ratio; CI, Confidence Interval.

A SNP replicated in this study was rs17863783, a synonymous variant (Val209Val) in UDP glucuronosyltransferase 1A6 (UGT1A6), which is known to glucuronidate several different substrates (Nagar, S. *et al.* Pharmacogenetics (2004) 14:487-499). This variant tags a specific haplotype (*4) in Europeans and East-Asians (not shown) that has been shown to have altered enzyme activity (Nagar, S. *et*

al. Pharmacogenetics (2004) 14:487-499), although this effect might be substrate specific (Krishnaswamy, S. *et al. J Pharmacol Exp Ther* (2005) 313:1340-1346). Even though the doxorubicin and daunorubicin parent compounds are likely not glucuronidated, it has been shown that certain metabolites do undergo glucuronidation (Andrews, P.A. *et al. Drug Metab Dispos* (1980) 8:152-156). Thus altered glucuronidation might lead to accumulation of toxic anthracycline metabolites. Similarly, the replication of rs10426377 in SULT2B1 or sulfotransferase 2B1, which catalyzes the sulfate conjugation of many compounds (Ji, Y. *et al. J Pharmacol Exp Ther* (2007) 322:529-540), might be explained by altered sulfonation of anthracycline metabolites (Andrews, P.A. *et al. Drug Metab Dispos* (1980) 8:152-156).

Previously, we showed the highly significant association of rs7853758 in SLC28A3 with ACT and further replication in a subset of the Dutch-EKZ cohort used in this study (Visscher, H. *et al. J Clin Oncol* (2011) Epub 11 Oct 2011). Here, we further substantiated this finding as well as that of another SNP in SLC28A3 (rs885004) in an extended cohort. In addition, we now found additional, although not statistically significant (P-replication = 0.14; P-combined = 0.0020), evidence for association of rs2305364 in SLC28A1, which is structurally similar to SLC28A3 and has substantial substrate overlap (Pastor-Anglada, M. *et al. Xenobiotica* (2008) 38:972-994). Both SLC28A genes encode for concentrative nucleoside transporters, which can putatively transport several anthracyclines into cells (Nagasawa, K. *et al. Curr Drug Metab* (2001) 2:355-366).

Evidence is presented herein for association of two variants in the ATP-binding cassette transporter B4 (ABCB4) in females only. ABC-transporters such as ABCB4 are known to efflux a variety of drugs including anthracyclines (Smith, A.J. *et al. J Biol Chem* (2000) 275:23530-23539). Reduced function or expression will lead to intracellular accumulation of anthracyclines. The strongest associated variant in ABCB4 (rs4148808) is located in the promoter region of the gene (Lang, T. *et al. Drug Metab Dispos* (2006) 34:1582-1599), potentially affecting expression. Female sex hormones likely reduce the expression of ABCB4 as well, as women with heterozygous mutations in ABCB4 can develop intrahepatic cholestasis during pregnancy when such hormones are high (Jacquemin, E. *et al. Lancet* (1999) 353:210-211), providing an explanation why an association was only seen in females.

Similarly, the effect of SULT2B1 rs10426377 was observed in males only. While no differences in SULT2B1 expression were found between male and female mice (Alnouti, Y. and Klaassen, C.D. *Toxicol Sci* (2006) 93:242-255), they do respond differently in SULT2B1 up-regulation after treatment with certain enzyme inducers (Alnouti, Y. and Klaassen, C.D. *J Pharmacol Exp Ther* (2008) 324:612-621).

rs17583889, and to a lesser extent of rs17645700, in HNMT or histamine N-methyltransferase in younger children (<5.3 years) only. HNMT catalyzes the N-methylation of histamine thereby terminating its activity (Verburg, K.M. and Henry, D.P. Histamine N-Methyltransferase. In Boulton AA, Baker GB, Yu PH (eds): Neurotransmitter Enzymes, Edition Humana Press 1986; 147-204). In an experimental rat model of hypotension, central inhibition of HNMT, leads to activation of the histaminergic system and mobilization of compensatory cardiovascular mechanisms (Jochem, J. Inflamm Res (2004) 53:316-323). HNMT activity gradually increases in both mouse and rat brain after birth and in mouse kidney the activity is about 60 times higher at maturation than at birth (Laduron, P. *et al.* Naunyn Schmiedeberg Arch Pharmacol (1975) 286:379-387). In human red blood cells HNMT activity is correlated with age with younger age having lower activities (Scott, M.C. *et al.* Clin Pharmacol Ther (1988) 43:256-262), though no children under 5 years were included and no such correlation was seen in liver or renal samples (De Santi, C. *et al.* Xenobiotica (1998) 28:571-577). Nevertheless, in younger patients, in whom activity may be lower, the effects of variants might therefore have a greater impact. The exact mechanisms by which these variants affect ACT as well as validation of these potential heterogeneous effects will require future studies.

The present studies were powered to find similar effect sizes, but often the effects are smaller in replication studies (Chanock, S.J. *et al.* Nature (2007) 447:655-660). Even though we have tried to keep the replication cohorts similar to the original cohorts, small differences might exist, for example in ethnicity, (supportive) treatment or follow-up and could potentially lead to non-replication due to different effects of the variants in specific populations or subgroups (Chanock, S.J. *et al.* Nature (2007) 447:655-660).

In our analyses we corrected for the effects of several important clinical risk factors. Not unexpectedly, cumulative doses were statistically significant higher in cases compared to controls in both the Dutch-EKZ and CPNDS cohort. Age was higher in CPNDS cases, whereas younger age is usually considered a risk factor (Kremer, L.C. *et al.* Ann Oncol (2002) 13:503-512). This is likely in part due to our requirement of controls to have at least 5 year follow-up, selecting for relatively younger controls in the CPNDS cohort.

Population stratification could not be assessed in the current replication cohort as the number of SNPs included were insufficient to use principal component analysis to assess population structure (Visscher, H. *et al.* Pharmacogenomics J (2009) 9:362-372). However, in the original study cohorts, we calculated

the genomic inflation factor to be 1.0 before principal component correction, indicating no population stratification and suggesting no or little influence on the results.

The current risk prediction model based on replicated genetic variants and clinical factors improved the ability to discriminate between cases and controls compared to clinical factors alone (AUC 0.77 versus AUC 0.69). More importantly, this optimized model was replicated in our test cohort with similar metrics. Including these genetic factors to predict patients at high and low risk for ACT could therefore inform treatment options such as administering cardioprotective agents (e.g. dexrazoxane) or using alternative anthracycline dosing or formulations as well as change monitoring decisions which could lead to improved and safer anthracycline treatment.

EXAMPLE 2 – EXTENDED PANEL CONTAINING 4,500+ SNPS

Methods for EXAMPLE 2

Samples

Study participants were recruited through the Canadian Pharmacogenomics Network for Drug Safety (CPNDS), a multicenter active surveillance consortium studying adverse drug reactions in children (Carleton, B. *et al.* *Pharmacoepidemiol Drug Saf* (2009) 18:713-721). The discovery cohort ($n=344$) in this study consisted of patients recruited from pediatric oncology units and long-term follow-up clinics across Canada between February 2005 and January 2010 and comprised of the two Canadian cohorts combined that were used previously to identify variants associated with ACT (Visscher, H. *et al.* *J Clin Oncol* (2011) Epub 11 Oct 2011). The replication cohort ($n=218$) consisted of additional patients recruited from across Canada between February 2010 and April 2011 and patients recruited at the Emma Children's Hospital/Academic Medical Centre in Amsterdam, the Netherlands, between July 2009 and April 2011 (van der Pal, H.J. *et al.* *Arch Intern Med* (2010) 170:1247-1255 ; and van Dalen, EC. *et al.* *Eur J Cancer* (2006) 42:3191-3198). This replication cohort was used previously to replicate earlier genetic findings and to validate the prediction model (Visscher H, Ross CJ, Rassekh SR *et al.* Validation of SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. Submitted to Cancer 2011).

The study cohorts and case-control definitions have been described in detail elsewhere (Visscher, H. *et al.* *J Clin Oncol* (2011) Epub 11 Oct 2011; and Visscher H, Ross CJ, Rassekh SR *et al.* Validation of SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. Submitted to Cancer 2011). In short, ACT cases were defined by early- or late-onset left ventricular dysfunction, during or after anthracycline treatment, on echocardiogram and/or symptoms requiring

intervention based on CTCAEv3 (Cancer Therapy Evaluation Program - Common Terminology Criteria for Adverse Events - Version 3. In Edition 2003). To improve differentiation between cases and controls, a more stringent shortening fraction (SF) threshold of $\leq 26\%$ was used for cases, while controls were required to have normal echocardiograms with $SF \geq 30\%$ during and at least 5 years after completion of anthracycline therapy. Transient acute cardiotoxicity was excluded by using only echocardiograms obtained ≥ 21 days after an anthracycline dose. Cumulative anthracycline doses were calculated using doxorubicin equivalents (Altman, A.J. Children's Oncology Group. Supportive care of children with cancer: current therapy and guidelines from the Children's Oncology Group. Baltimore: Johns Hopkins University Press 2004).

Written informed consent or assent was obtained from each subject or their parents or legal guardians. The study was approved by the ethics committees of all participating universities and hospitals.

Genotyping

Genomic DNA was extracted from blood, saliva or buccal swabs using the QIAamp™ DNA purification system (Qiagen™, Canada). DNA samples were genotyped for 4536 SNPs using a customized Illumina GoldenGate™ SNP genotyping assay (Illumina™, San Diego, USA), which was designed to capture the genetic variation of approximately 300 key drug biotransformation genes (i.e. phase I and II drug metabolism enzymes, drug transporters, drug targets, drug receptors, transcription factors, ion channels and other specific genes known to be related to the pathophysiological pathway of ADRs. This ADME (absorption, distribution, metabolism and elimination)-toxicity panel consisted of functional SNPs – that had been identified primarily by literature review and from public databases – that cause non-synonymous amino-acid changes or could be associated with changes in enzyme activity or function. In addition, tagSNPs were included that were identified using the ldSelect™ algorithm to select a maximally informative set of tagSNPs to assay the candidate genes (Carlson, C.S. *et al.* Am J Hum Genet (2004) 74:106-120). TagSNP selection was performed using data from phase II of the International HapMap project that included all four populations (CEU, CHB, JPT and YRI) (International HapMap Consortium. A haplotype map of the human genome. Nature (2005) 437:1299-1320) with a threshold for the linkage disequilibrium (LD) statistic r^2 of 0.8 and a minor allele frequency of >0.05 . The current SNP panel is an updated version of the panel that was used previously (Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011; Visscher, H. *et al.* Pharmacogenomics J (2009) 9:362-372; and Ross, C.J. *et al.* Nat Genet (2009) 41:1345-1349), which was extended to include additional genes and further optimized by replacing previously unsuccessful SNPs with others where possible or by optimizing the design of specific oligonucleotides. In addition, the main SNP panel was supplemented by a custom 96-SNP

Illumina™ Veracode GoldenGate™ genotyping assay. This assay was designed specifically to include both functional and tagSNPs in genes involved in the metabolism of anthracyclines into alcohol metabolites (AKRs and CBRs) (Blanco, J.G. *et al.* Cancer (2008) 112:2789-2795; Blanco, J.G. *et al.* J Clin Oncol (Meeting Abstracts) (2010) 28:9512; Bains, O.S. *et al.* Drug Metab Dispos (2008) 36:904-910; and Bains, O.S. *et al.* J Pharmacol Exp Ther (2010) 335:533-545) as well as other SNPs possibly related to ACT not included in the main panel.

All SNP genotypes were manually clustered using Illumina™ GenomeStudio™ software. Fifty-four SNPs were assayed in duplicate, so a total of 4578 unique SNPs were included. Furthermore, 374 SNPs that could not be clustered, or were non-polymorphic or had a completion rate of <95% as well as 51 ancestry-informative markers were excluded, leaving a total of 4153 SNPs for further analysis. The average call rate for the included SNPs was 99.8%. Concordance between replicate samples ($n=34$) was >99.9%. Sixteen samples (3 cases and 13 controls) with a call rate <95% were removed. The remaining 546 samples had an average call rate of 99.8%.

Statistical analysis

Hardy-Weinberg Equilibrium (HWE) tests were conducted using Fisher's Exact test in controls only. Twenty-nine SNPs had $P < 1.7 \times 10^{-5}$ in the HWE test. These SNPs were marked, but retained in the analysis. All of the top associated SNPs were in HWE. To reduce possible false-positive associations due to multiple testing, we applied a tiered analysis to identify SNPs associated at $P < 0.01$ in the larger discovery cohort that remained associated in the smaller replication cohort at $P < 0.05$. A more conservative overall Bonferroni corrected significance threshold was calculated at $P < 1.7 \times 10^{-5}$ using the effective number of independent tests (M_{effG}) (Gao, X. *et al.* Genet Epidemiol (2008) 32:361-369). No duplicate or (cryptic) related samples were found by calculating the average identity-by-state for each subject-pair. Population structure was assessed by principal component analysis.

Our primary analysis was a case-control association test using logistic regression assuming an additive genetic model. Cumulative anthracycline dose was included, age at start of treatment, gender and radiation therapy to the heart as important clinical covariates and the first two principal components to correct for potential population stratification. As our primary aim was to identify additional variants, we also included the previously (Visscher H, Ross CJ, Rassekh SR *et al.* Validation of SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. Submitted to Cancer 2011) validated variants in *SLC28A3*(rs7853758) and *UGT1A6* (rs17863783) as covariates to adjust for the effect of these variants. Finally, the primary analysis was conducted in the 536 patients

(125 cases and 411 controls) who received doxorubicin and/or daunorubicin as the effect of rs7853758 had previously only been observed in these patients (Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011) – of these, 520 (122 cases and 398 controls) were successfully genotyped. Haplotypes were inferred using the expectation-maximization algorithm. Haplotype association tests were done using logistic regression and included the same covariates.

Multivariate logistic regression models including multiple genetic variants and/or clinical variables were all trained in the discovery cohort and tested in the replication cohort as described previously (Visscher H, Ross CJ, Rassekh SR *et al.* Validation of SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. Submitted to Cancer 2011). Risk scores were calculated by multiplying each variable with the estimated beta (log odds ratio) from the training cohort. To assess whether adding the newly identified variants to our previous model – that consisted of 5 SNPs as well as the clinical variables gender, age, anthracycline dose and radiation to the heart (Visscher H, Ross CJ, Rassekh SR *et al.* Validation of SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. Submitted to Cancer 2011) – would improve prediction of ACT, Receiver Operating Characteristic (ROC) curves were constructed and calculated the *c*-statistic (Area Under the Curve- AUC) using the risk scores from the model and the actual value (case or control).

Statistical analyses were conducted using SNP and Variation Suite™ 7.4.5 (Golden Helix™, Bozeman, USA) and R 2.13.0 (R Development Core Team™) with package pROC 1.4.3 (Robin, X. *et al.* BMC Bioinformatics 2011; 12: 77).

SLC22A17 and SLC22A7

In a tiered analysis, two SNPs were identified, one located in the 3' region of solute carrier family 22, member 17 (SLC22A17;rs4982753) and another one in SLC22A7 (rs4149178), that were significantly associated with ACT in the discovery cohort ($P=0.0078$ and $P=0.0034$, respectively) and were confirmed in the replication cohort ($P=0.0071$ and $P=0.047$, respectively – TABLE 10). Both SNPs had similar effect sizes in both discovery and replication cohort with combined odds ratios (OR) of 0.50 (95% CI 0.33-0.75) and 0.45 (95% CI 0.26-0.75) respectively.

Nearly identical results were obtained when we included patients that received other anthracyclines but not doxorubicin or daunorubicin (combined $P=6.7 \times 10^{-4}$ and $P=9.4 \times 10^{-4}$; OR 0.53 and OR 0.45, respectively). Again, in an analysis of only patients of European ancestry (99 cases, 314 controls), as determined by the first two principal components (Visscher, H. *et al.* Pharmacogenomics J (2009) 9:362-

372), results were similar with a combined OR for SLC22A17 rs4982753 of 0.57 (95% CI 0.37-0.88) and of 0.45 (95% CI 0.25-0.79) for SLC22A7 rs4149178. Finally, in a subgroup analysis of patients (64 cases) with more severe cardiotoxicity (SF<24% or symptoms, CTCAE grade 2-4), the result for SLC22A7 rs4149178 was comparable (OR 0.49 [95% CI 0.25-0.95]), while the effect for SLC22A17 rs4982753 seemed less strong in this analysis [OR 0.73 (95% CI 0.44-1.20)].

We also genotyped several other variants in both genes; of these one other SNP in SLC22A17 (rs11625724) showed a marginal significant association with ACT ($P=0.020$, OR 1.63 in the combined cohort). To see whether this was due to LD between rs4982753 and this marker ($r^2=0.067$), we adjusted for the effect of rs4982753 in SLC22A17 in the regression analysis. However, the effect of rs11625724 did not completely disappear ($P=0.11$, OR 1.40), suggesting an independent effect of this SNP. Interestingly, two other SNPs in SLC22A17 (rs12882406 and rs12896494) also became marginal significant after adjusting for the effect of rs4982753 ($P=0.042$ and $P=0.031$; OR 1.52 and OR 0.65, respectively). Haplotype analysis including these 4 SNPs did not reveal more significant results.

Additional variants

In addition to the variants in SLC22A17 and SLC22A7, suggestive evidence ($P<0.005$) was found for association of several other variants with ACT (see TABLE 11). Despite the fact that these variants were not identified in the tiered analysis were they significant after Bonferroni correction, it is likely that there are true associations between these SNPs and ACT.

The strongest associated variant was rs10426628 (combined $P=3.2\times 10^{-4}$, OR 1.92), which is located in sulfotransferase 2B1 (SULT2B1). As we had previously found evidence for association of another variant in SULT2B1 with ACT (rs10426377) see EXAMPLE 1, the analysis was re-run adjusting for the effect of rs10426377. In this analysis, rs10426628 remained associated with ACT (combined $P=0.0013$, OR 1.80), suggesting an independent effect of this variant. Haplotype analysis revealed that the protective effect of the minor allele of rs10426377 (A) in carriers of haplotype AC (OR 0.41 [95% CI 0.26-0.66], $P=7.4\times 10^{-5}$) and, to a lesser extent, the risk effect of the minor A-allele of rs10426628 in CA-carriers (OR 2.04 [95% CI 1.38-3.03], $P=4.4\times 10^{-4}$) was offset by the other allele in haplotype AA-carriers (OR 1.50 [95% CI 0.72-3.11], $P=0.29$ – TABLE 12).

Interestingly, many of the other variants with suggestive evidence ($P<0.005$) for association with ACT were found in genes related to oxidative stress (TABLE 11), which is important in ACT, indicating that these might be true associations.

Predictive model

Next, we assessed whether adding the newly identified and replicated variants in SLC22A17 and SLC22A7 to the risk prediction model that was previously created discussed in EXAMPLE 1, would improve the ability to discriminate between cases and controls. ROC analyses showed that adding the two variants to the previous model did improve the AUC in the discovery cohort, in which the models were constructed, as well as in the replication cohort, in which the models were tested (TABLE 13). The AUC for the extended model in the combined cohort was 0.781 compared to 0.755 for the previous model ($P=0.029$). Similarly, in a genetic-only model (without the clinical variables included), adding the 2 SNPs improved the AUC significantly (0.716 vs. 0.671, $P=0.0081$; TABLE 14).

TABLE 10. Discovery and replication of SNPs associated with anthracycline-induced cardiotoxicity

SNP rs-ID	Gene	Chr	Position ^a	Type ^b	Alleles ^c	Discovery ($n=335$)		Replication ($n=185$)		Combined ($n=520$)	
						78 cases, 257 controls		44 cases, 141 controls		122 cases, 398 controls	
						OR (95% CI)	<i>P</i>	OR (95% CI)	<i>P</i>	OR (95% CI)	<i>P</i>
rs4982753	SLC22A17	14	22,884,409	flanking/3'-UTR	A/G	0.52 (0.31-0.85)	0.0078	0.39 (0.19-0.81)	0.0071	0.50 (0.33-0.75)	4.4×10^{-4}
rs4149178	SLC22A7	6	43,380,166	intronic	G/A	0.41 (0.21-0.77)	0.0034	0.39 (0.14-1.05)	0.047	0.45 (0.26-0.75)	0.0013

SNPs associated with anthracycline-induced cardiotoxicity in discovery cohort at $P<0.01$ and validated in replication cohort at $P<0.05$. Results are from logistic regression analysis that included important clinical variables as well as the previously validated variants rs7853758 in SLC28A3 and rs17863783 in UGT1A6.[14] Only patients that did receive doxorubicin and/or daunorubicin were included.

Odds ratios are per copy of the minor allele. ^aPosition based on NCBI Build 36.3; ^bRelative to gene of interest; ^cSNP alleles assayed, minor allele mentioned first; SNP, Single Nucleotide Polymorphism; OR, Odds Ratio; CI, Confidence Interval.

TABLE 11. Other associations between SNPs and anthracycline-induced cardiotoxicity

SNP rs-ID	Gene	Chr	Position ^a	Type ^b	Alleles ^c	Discovery (<i>n</i> =335)		Replication (<i>n</i> =185)		Combined (<i>n</i> =520)	
						78 cases, 257 controls	P-value	44 cases, 141 controls	P-value	122 cases, 398 controls	P-value
rs2294950	CYP2J2	1	60,087,084	flanking/3'-UTR	C/A	0.41 (0.19-0.90)	0.015	0.40 (0.13-1.18)	0.071	0.39 (1.34-2.73)	0.0014
rs12059276	GSTM3	1	110,075,064	flanking/3'-UTR	A/G	0.37 (0.14-0.96)	0.027	0.30 (0.09-1.05)	0.035	0.36 (0.33-0.75)	0.0031
rs4407290	XDH	2	31,460,174	synon (V279V)	A/G	0.26 (0.06-1.16)	0.035	0.13 (0.00-1.10)	0.065	0.18 (1.40-3.71)	0.0039
rs2236168	XDH	2	31,460,632	intronic	G/A	1.42 (0.95-2.14)	0.089	2.78 (1.47-5.24)	8.6 x10 ⁻⁴	1.68 (1.50-4.76)	0.0017
rs10497346	ABCB11	2	169,479,442	flanking/3'-UTR	G/A	2.29 (1.16-4.54)	0.018	2.17 (0.91-5.20)	0.089	2.23 (0.22-0.73)	0.0033
rs2600834	SLCO4C1	5	101,633,540	intronic	A/G	2.01 (1.28-3.16)	0.0022	1.42 (0.74-2.75)	0.30	1.80 (1.26-2.57)	0.0011
rs12658397	SLCO6A1	5	101,779,552	intronic	G/A	1.83 (1.20-2.80)	0.0048	1.35 (0.71-2.58)	0.36	1.68 (0.26-0.75)	0.0025
rs2233302	GPX3	5	150,395,291	flanking/3'-UTR	C/G	0.27 (0.11-0.65)	7.4x10 ⁻⁴	0.65 (0.27-1.57)	0.33	0.40 (0.21-0.74)	0.0011
rs1214763	ABCC10	6	43,468,239	flanking/5'-UTR	A/G	0.34 (0.15-0.75)	0.0031	0.63 (0.23-1.71)	0.35	0.43 (1.24-2.57)	0.0035
rs2180314	GSTA2	6	52,725,690	nonsynon (S112T)	G/C	0.75 (0.50-1.13)	0.17	0.44 (0.25-0.79)	0.0038	0.62 (0.15-0.72)	0.0036
rs7754103	SOD2	6	159,981,080	flanking/3'-UTR	A/G	0.30 (0.10-0.94)	0.020	0.29 (0.09-0.98)	0.035	0.32 (1.20-2.35)	0.0024
rs42524	COL1A2	7	93,881,175	nonsynon (P549A)	G/C	1.78 (1.11-2.88)	0.018	1.78 (0.95-3.33)	0.072	1.79 (0.17-0.76)	0.0020
rs3887137	ABCA1	9	106,738,433	flanking/5'-UTR	A/G	2.33 (1.31-4.15)	0.0041	2.31 (0.92-5.82)	0.078	2.28 (1.32-3.77)	9.5 x10 ⁻⁴
rs11046217	ABCC9	12	21,908,424	intronic	C/G	4.48 (2.10-9.57)	7.1 x10 ⁻⁵	0.92 (0.31-2.71)	0.88	2.67 (0.24-0.79)	9.9 x10 ⁻⁴
rs8001466	SLC15A1	13	98,153,602	intronic	C/G	1.81 (0.99-3.32)	0.057	2.32 (0.99-5.47)	0.054	2.02 (0.45-0.86)	0.0042
rs10144771	SERPINA6	14	93,848,406	intronic	A/G	2.23 (1.39-3.58)	9.0 x10 ⁻⁴	1.11 (0.58-2.13)	0.75	1.72 (0.04-0.79)	0.0042
rs10426628	SULT2B1	19	53,784,242	intronic	A/G	1.60 (1.03-2.48)	0.037	2.95 (1.51-5.78)	0.0012	1.92 (1.19-2.50)	3.2 x10 ⁻⁴
rs2425886	SLC13A3	20	44,693,250	intronic	G/A	1.56 (0.98-2.49)	0.063	2.07 (1.02-4.20)	0.042	1.75 (1.25-3.26)	0.0037

Other SNPs associated with anthracycline-induced cardiotoxicity in combined cohort at *P*<0.005. Results are from logistic regression analysis that included important clinical variables as well as the previously validated variants rs7853758 in *SLC28A3* and rs17863783 in *UGT1A6*. [14] The replication cohort included only patients that did receive doxorubicin and/or daunorubicin.

Odds ratios are per copy of the minor allele. ^aPosition based on NCBI Build 36.3; ^bRelative to gene of interest; ^cSNP alleles assayed, minor allele mentioned first; SNP, Single Nucleotide Polymorphism; UTR, Untranslated Region; OR, Odds Ratio; CI, Confidence Interval.

TABLE 12. SULT2B1 haplotype analysis and anthracycline-induced cardiotoxicity

Haplotype		Combined (n=520) 122 cases, 398 controls			
		Freq. cases	Freq. controls	OR (95% CI)	P-value
rs10426377-A	rs10426628-G	0.14	0.25	0.41 (0.26-0.66)	7.4x10 ⁻⁵
rs10426377-C	rs10426628-A	0.26	0.17	2.04 (1.38-3.03)	4.4 x10 ⁻⁴
rs10426377-A	rs10426628-A	0.07	0.06	1.50 (0.72-3.11)	0.29
rs10426377-C	rs10426628-G	0.53	0.52	0.97 (0.68-1.37)	0.85

Haplotype analysis of variants rs10426377 and rs10426628 in *SULT2B1*. Protection of rs10426377-A is offset by risk of rs10426628-A in AA-carriers and vice versa. Haplotypes were inferred by EM-algorithm. Results are from logistic regression including clinical covariates. Odds ratios are per copy of the haplotype compared to all other haplotypes. Freq., haplotype frequency; OR, Odds Ratio; CI, Confidence Interval.

TABLE 13. Comparison of different genetic only predictive models

Model	Discovery (n=335)		Replication (n=185)		Combined (n=520)		P-value vs. 5 SNPs only
	AUC	95% CI	AUC	95% CI	AUC	95% CI	
5 SNPs only	0.697	(0.630-0.763)	0.615	(0.524-0.705)	0.671	(0.617-0.724)	--
5 SNPs only + SLC22A17	0.723	(0.658-0.787)	0.652	(0.564-0.740)	0.700	(0.648-0.752)	0.018
5 SNPs only + SLC22A17 + SLC22A7	0.740	(0.678-0.803)	0.664	(0.579-0.750)	0.716	(0.666-0.766)	0.0081

TABLE 14. Comparison of different predictive models

Model	Discovery (n=335)		Replication (n=185)		Combined (n=520)		P-value	
	AUC	95% CI	AUC	95% CI	AUC	95% CI	vs. Clin-only	vs. Previous
Clinical only	0.681	(0.615-0.747)	0.663	(0.578-0.747)	0.676	(0.624-0.728)	--	--
Previous model (Clinical + 5 SNPs)	0.762	(0.701-0.823)	0.740	(0.659-0.821)	0.755	(0.706-0.804)	6.8x10 ⁻⁴	--
Previous + SLC22A17	0.774	(0.716-0.833)	0.751	(0.673-0.828)	0.766	(0.719-0.813)	1.9 x10 ⁻⁴	0.16
Previous + SLC22A17 + SLC22A7	0.793	(0.737-0.849)	0.756	(0.684-0.829)	0.781	(0.737-0.826)	1.7 x10 ⁻⁵	0.029

Comparison of different models that were trained in the discovery and then tested in the replication cohort. P-values calculated for the combined cohort. The clinical only model included age, cumulative dose, gender, radiation therapy involving the heart region and the first two principal components. The previous model includes the same clinical variables as well as five SNPs (rs7853758, *SLC28A3*; rs17863783, *UGT1A6*; rs10426377, *SULT2B1*; rs2305364, *SLC28A1*; and rs4148808, *ABCB4*) from Visscher et al. [14]. The SNPs in *SLC22A17* (rs4982753) and *SLC22A7* (rs4149178) were subsequently added. AUC, Area Under the Curve; CI, Confidence interval.

Comparison of different genetic only models that were trained in the discovery and then tested in the replication cohort. P-values calculated for the combined cohort. The 5 SNPs only model includes five SNPs (rs7853758, *SLC28A3*; rs17863783, *UGT1A6*; rs10426377, *SULT2B1*; rs2305364, *SLC28A1*; and rs4148808, *ABCB4*) from Visscher et al. [14]. The SNPs in *SLC22A17* (rs4982753) and *SLC22A7* (rs4149178) were subsequently added. AUC, Area Under the Curve; CI, Confidence interval.

Previously, several genetic variants were found to be associated with anthracycline-induced cardiotoxicity in children, though few have been replicated to date (Blanco, J.G. *et al.* Cancer (2008) 112:2789-2795; Rajic, V. *et al.* Leuk Lymphoma (2009) 50:1693-1698; Visscher, H. *et al.* J Clin Oncol (2011) Epub 11 Oct 2011); and Visscher H, Ross CJ, Rassekh SR *et al.* Validation of *SLC28A3* and *UGT1A6* as genetic markers predictive of anthracycline-induced cardiotoxicity in children. Submitted to Cancer 2011). In the present example, we report the identification and replication of two additional variants in *SLC22A17* and *SLC22A7* as predictive markers of ACT. Additionally, evidence for association of variants in *SULT2B1* are shown as well as in several antioxidant genes. Addition of the replicated variants to our ACT risk prediction model, further improved the ability of this model to predict patients at risk.

The two identified and replicated variants were both found in a member of the solute carrier family 22. This large gene family currently consists of 26 members that encode for organic cation transporters (OCTs), organic cation/carnitine transporters (OCTNs), and organic anion transporters (OATs) (Koepsell, H. and Endou, H. Pflugers Arch (2004) 447:666-676; and Cano-Soldado, P. and Pastor-Anglada, M. Transporters that translocate nucleosides and structural similar drugs: structural requirements for substrate recognition. Med Res Rev (2011) published online:1 FEB DOI: 10.1002/med.20221). Interestingly, two members of this family, *SLC22A16*(OCT6) and *SLC22A4* (OCTN1), have recently been identified as importers of doxorubicin (Okabe, M. *et al.* Biochem Biophys Res Commun (2005) 333:754-762; and Okabe, M. *et al.* Mol Cancer Ther (2008) 7:3081-3091). *SLC22A17* or brain-type OCT, as it was first identified in brain, is an orphan transporter without a known endogenous substrate and is expressed in a variety of tissues including the heart as well as liver and kidney (Bennett, K.M. *et al.* Mol Cell Biochem (2011) 352:143-154). It shows significant sequence similarity to many other OCTs, though differences do exist (Bennett, K.M. *et al.* Mol Cell Biochem (2011) 352:143-154). *SLC22A7* or OAT2, which has been studied in more detail, is widely expressed including in cardiac tissue (Cropp, C.D. *et al.* Mol Pharmacol (2008) 73:1151-1158). It is able to transport naturally occurring nucleobases, nucleosides and nucleotides, with a preference for guanine analogs, and several nucleoside-based drugs and has considerable substrate overlap with concentrative nucleoside transporters such as *SLC28A3* (Cropp, C.D. *et al.* Mol Pharmacol (2008) 73:1151-1158; and Errasti-Murugarren, E. and Pastor-Anglada, M. Pharmacogenomics (2010) 11:809-841). Therefore, it seems plausible to that *SLC22A17* and *SLC22A7* are transporting anthracyclines.

Furthermore, additional support was found for the involvement of SULT2B1 (rs10426628) in ACT. In combination with the independent association of rs10426377 set out in EXAMPLE 1, this provides compelling evidence for a role of genetic variants in SULT2B1 in ACT risk. Interestingly, inspection of these two intronic variants in the UCSC genome browser (<http://genome.ucsc.edu/>) revealed that they are located less than 200bp apart in a region with specific histone marks, DNase hypersensitivity and transcription factor binding which suggests putative enhancer activity (Heintzman, N.D. *et al.* Nat Genet (2007) 39:311-318), suggesting that these SNPs might affect SULT2B1 expression. This sulfotransferase, which catalyzes the sulfate conjugation of many compounds (Ji, Y. *et al.* J Pharmacol Exp Ther (2007) 322:529-540), may influence anthracycline metabolite sulfonation (Andrews, P.A. *et al.* Drug Metab Dispos (1980) 8:152-156), thereby affecting excretion of the drug and its toxic metabolites.

Using a tiered analysis approach, the likelihood of finding false-positives (type I error) is reduced. Clinical risk factors for ACT, such as cumulative dose and concomitant cardiac irradiation, will not have influenced our results, as they were included as covariates in the analysis. Furthermore, corrections for potential population stratification were made by including principal components in our analysis to reduce false-positives. Furthermore, we excluded cases with mild cardiotoxicity ($SF \geq 27-30\%$) as well as controls with less than 5 years follow-up to ensure optimal separation between cases and controls and increasing the chance to find true associations.

Adding the variants in SLC22A17 and SLC22A7 improved the ACT risk prediction model significantly.

Risk stratification for anthracycline-induced cardiotoxicity based on a model that includes genetic risk variants, and not solely clinical risk factors, {COG Guidelines, 2008} might prove useful and could inform monitoring frequency to detect early damage {COG Guidelines, 2008} as well as indicate benefit of preventive measures such as use of cardio-protective agents (e.g. dexrazoxane) or alternative anthracycline dosing or formulations (Lipshultz, S.E. *et al.* Heart (2008) 94:525-533; van Dalen, E.C. *et al.* Cochrane Database Syst Rev (2011) 6:CD003917; van Dalen, E.C. *et al.* Cochrane Database Syst Rev (2009) CD005008; and van Dalen, E.C. *et al.* Cochrane Database Syst Rev (2010) CD005006). Before clinical implementation of stratification based on genetic and clinical risk factors, future prospective studies will likely be needed to validate the current risk prediction model and to establish the benefit of such measures in these patients. The current results help to further unravel genetic risk of ACT, thereby providing means to improve cancer therapy safety in children and potentially adults as well.

WHAT IS CLAIMED IS:

1. A method of screening a subject having a neoplastic disease for cardiotoxicity risk, the method comprising: determining the identity for one or more of the following single
5 nucleotide polymorphisms (SNPs): rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628 or one or more polymorphic sites in linkage disequilibrium thereto, for the subject, wherein the subject is a candidate for anthracycline administration.
- 10 2. A method for diagnosing a predisposition for cardiotoxicity risk in a human subject from anthracycline administration, the method comprising: a) determining an identity for one or more of the following single nucleotide polymorphisms (SNPs) in a biological sample from the subject: rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178;
15 rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628 or one or more polymorphic sites in linkage disequilibrium thereto from the sample; and b) making a cardiotoxicity risk determination based on the prevalence of risk alleles in the subject sample.
- 20 3. The method of claim 1 or 2, further comprising determining the identity of rs17863783 or one or more polymorphic sites in linkage disequilibrium thereto.
4. The method of claim 1, 2 or 3, wherein the anthracycline is selected from one or more of the following: anthracycline antibiotics such as daunorubicin (daunomycin, rubidomycin),
25 doxorubicin, idarubicin, epirubicin, mitoxantrone, carminomycin, esorubicin, quelamycin, aclarubicin, esorubicin, zorubicin, pirarubicin, amrubicin, iododoxorubicin, detorubicin, marcellomycin, rodorubicin, and valrubicin.
5. The method of any one of claims 1-4, wherein the method further comprises administering
30 the anthracycline in accordance with the subject's risk of developing cardiotoxicity.
6. The method of any one of claims 1-5, wherein the subject has a cardiotoxicity risk allele selected from one or more of: rs7853758G; rs885004G; rs17863783A (reverse);
rs17863783T (forward); rs10426377C; rs2305364A; rs4982753G (reverse); rs4982753C
35 (forward); rs4149178A; and rs10426628A.

7. The method of any one of claims 1-5, wherein the subject has a reduced cardiotoxicity risk allele selected from one or more of: rs7853758A; rs885004A; rs17863783C (reverse); rs17863783G (forward); rs10426377A; rs2305364G; rs4982753A (reverse); rs4982753T (forward); rs4149178G; and rs10426628G.
- 5
8. The method of any one of claims 1-7, wherein the identity of a single nucleotide polymorphism is determined by one or more of the following techniques:
- (a) restriction fragment length analysis;
 - (b) sequencing;
 - 10 (c) micro-sequencing assay;
 - (d) hybridization;
 - (e) invader assay;
 - (f) gene chip hybridization assays;
 - (g) oligonucleotide ligation assay;
 - 15 (h) ligation rolling circle amplification;
 - (i) 5' nuclease assay;
 - (j) polymerase proofreading methods;
 - (k) allele specific PCR;
 - (l) matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectroscopy;
 - 20 (m) ligase chain reaction assay;
 - (n) enzyme-amplified electronic transduction;
 - (o) single base pair extension assay; and
 - (p) reading sequence data.
- 25
9. A method of determining cardiotoxicity risk from anthracycline administration, the method comprising: determining the identity of the single nucleotide polymorphism (SNP) at each of the following polymorphic sites: rs7853758; rs885004; rs17863783; rs10426377; and rs2305364; or a polymorphic sites in linkage disequilibrium thereto, for a subject receiving or about to receive one or more anthracyclines.
- 30
10. A method of determining cardiotoxicity risk from anthracycline administration, the method comprising: determining the identity of the single nucleotide polymorphism (SNP) at each of the following polymorphic sites: rs7853758; rs4148808; rs17863783; rs10426377; and rs2305364; or a polymorphic sites in linkage disequilibrium thereto, for a subject receiving or about to receive one or more anthracyclines.
- 35

11. A method of determining cardiotoxicity risk from anthracycline administration, the method comprising: determining the identity of the single nucleotide polymorphism (SNP) at each of the following polymorphic sites: rs7853758; rs17863783; rs10426377; and rs2305364; or a polymorphic sites in linkage disequilibrium thereto, for a subject receiving or about to receive one or more anthracyclines.
12. The method of claim 9, 10, or 11, further comprising determining the identity of the following two SNPs: rs4982753; and rs4149178; or one or more polymorphic sites in linkage disequilibrium thereto.
13. The method of any one of claims 9-12, wherein the anthracycline is selected from one or more of the following: anthracycline antibiotics such as daunorubicin (daunomycin, rubidomycin), doxorubicin, idarubicin, epirubicin, mitoxantrone, carminomycin, esorubicin, quelamycin, aclarubicin, esorubicin, zorubicin, pirarubicin, amrubicin, iododoxorubicin, detorubicin, marcellomycin, Rodorubicin, and valrubicin.
14. The method of any one of claims 7, wherein the method further comprises administering the anthracycline in accordance with the subject's risk of developing cardiotoxicity.
15. The method of any one of claims 9-12, wherein the subject has a cardiotoxicity risk allele selected from one or more of: rs7853758G; rs885004G; rs17863783A (reverse); rs17863783T (forward); rs10426377C; rs2305364A; rs4982753G (reverse); rs4982753C (forward); rs4149178A; and rs10426628A.
16. The method of any one of claims 9-12, wherein the subject has a reduced cardiotoxicity risk allele selected from one or more of: rs7853758A; rs885004A; rs17863783C (reverse); rs17863783G (forward); rs10426377A; rs2305364G; rs4982753A (reverse); rs4982753T (forward); rs4149178G; and rs10426628G.
17. The method of any one of claims 9-16, wherein the identity of a single nucleotide polymorphism is determined by one or more of the following techniques:
- (a) restriction fragment length analysis;
 - (b) sequencing;
 - (c) micro-sequencing assay;
 - (d) hybridization;

- 5 (e) invader assay;
 (f) gene chip hybridization assays;
 (g) oligonucleotide ligation assay;
 (h) ligation rolling circle amplification;
 (i) 5' nuclease assay;
 (j) polymerase proofreading methods;
 (k) allele specific PCR;
 (l) matrix assisted laser desorption ionization time of flight (MALDI-TOF)
 mass spectroscopy;
 10 (m) ligase chain reaction assay;
 (n) enzyme-amplified electronic transduction;
 (o) single base pair extension assay; and
 (p) reading sequence data.
- 15 18. Two or more oligonucleotides or peptide nucleic acids of about 10 to about 400 nucleotides
 that hybridize specifically to a sequence contained in a human target sequence consisting of
 a subject's toxicity associated gene sequence, a complementary sequence of the target
 sequence or RNA equivalent of the target sequence and wherein the oligonucleotides or
 peptide nucleic acids are operable in determining the presence or absence of two or more
 20 polymorphism(s) in the toxicity associated gene sequence selected from of the following
 polymorphic sites: rs7853758; rs885004; rs17863783; rs10426377; rs2305364; rs4982753;
 rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406;
 rs12896494; and rs10426628.
- 25 19. Two or more oligonucleotides or peptide nucleic acids selected from the group consisting
 of:
 (a) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency
 conditions to a nucleic acid molecule comprising SEQ ID NO:1 having an A at position 301
 but not to a nucleic acid molecule comprising SEQ ID NO:1 having a G at position 301;
 30 (b) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency
 conditions to a nucleic acid molecule comprising SEQ ID NO:1 having a G at position 301
 but not to a nucleic acid molecule comprising SEQ ID NO:1 having an A at position 301;
 (c) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency
 conditions to a nucleic acid molecule comprising SEQ ID NO:2 having a G at position 201
 35 but not to a nucleic acid molecule comprising SEQ ID NO:2 having an A at position 201;

- (d) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:2 having an A at position 201 but not to a nucleic acid molecule comprising SEQ ID NO:2 having an G at position 201;
- (e) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:3 having an A at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:3 having a C at position 301;
- (f) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:3 having a C at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:3 having an A at position 301;
- (g) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:4 having an A at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:4 having a C at position 501;
- (h) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:4 having a C at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:4 having an A at position 501;
- (i) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:5 having an A at position 251 but not to a nucleic acid molecule comprising SEQ ID NO:5 having a G at position 251;
- (j) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:5 having a G at position 251 but not to a nucleic acid molecule comprising SEQ ID NO:5 having an A at position 251;
- (k) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:6 having an A at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:6 having a G at position 501;
- (l) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:6 having a G at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:6 having an A at position 501;
- (m) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:7 having an A at position 401 but not to a nucleic acid molecule comprising SEQ ID NO:7 having a G at position 401;
- (n) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:7 having a G at position 401 but not to a nucleic acid molecule comprising SEQ ID NO:7 having an A at position 401;
- (o) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:8 having an A at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:8 having a G at position 301;

- (p) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:8 having a G at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:8 having an A at position 301;
- 5 (q) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:9 having an A at position 201 but not to a nucleic acid molecule comprising SEQ ID NO:9 having a C at position 201;
- (r) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:9 having a C at position 201 but not to a nucleic acid molecule comprising SEQ ID NO:9 having an A at position 201;
- 10 (s) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:10 having an A at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:10 having a C at position 301;
- (t) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:10 having a C at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:10 having an A at position 301;
- 15 (u) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:11 having an A at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:11 having a C at position 301;
- 20 (v) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:11 having a C at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:11 having an A at position 301;
- (w) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:12 having an A at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:12 having a T at position 301;
- 25 (x) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:12 having a T at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:12 having an A at position 301;
- 30 (y) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:13 having a C at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:13 having a G at position 501;
- (z) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:13 having a G at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:13 having a C at position 501;
- 35

- (aa) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:14 having a C at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:14 having a T at position 301;
- (bb) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:14 having a T at position 301 but not to a nucleic acid molecule comprising SEQ ID NO:14 having a C at position 301;
- (cc) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having an A at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having a C, G, or T at position 501;
- (dd) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having a C at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having an A, G, or T at position 501;
- (ee) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having an G at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having a C, A, or T at position 501; and
- (ff) an oligonucleotide or peptide nucleic acid that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO:15 having a T at position 501 but not to a nucleic acid molecule comprising SEQ ID NO:15 having an A, G, or C at position 501.
20. An array of oligonucleotides or peptide nucleic acids attached to a solid support, the array comprising two or more of the oligonucleotides or peptide nucleic acids of claim 18 or 19.
21. A composition comprising an addressable collection of two or more oligonucleotides or peptide nucleic acids, the two or more oligonucleotides or peptide nucleic acids consisting essentially of two or more nucleic acid molecules set out in SEQ ID NO:1-15 or compliments, fragments, variants, or analogs thereof.
22. The oligonucleotides or peptide nucleic acids of claim 18 or 19, further comprising one or more of the following: a detectable label; a quencher; a mobility modifier; a contiguous non-target sequence situated 5' or 3' to the target sequence or 5' and 3' to the target sequence.
23. A use of an anthracycline compound having a cardiotoxicity risk in the manufacture of a

- medicament for the treatment of a subject, where the subject is a candidate for anthracycline administration, and wherein the subject treated has a reduced cardiotoxicity risk genotype at one or more of the following polymorphic sites: rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto.
24. A use of an anthracycline compound having a cardiotoxicity risk for the treatment of a subject, wherein the subject treated has a reduced cardiotoxicity risk genotype at one or more of the following polymorphic sites: rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto; for the subject, where the subject is a candidate for anthracycline administration.
25. The use of claim 23 or 24, wherein the subject has a reduced cardiotoxicity risk genotype at rs17863783 or one or more polymorphic sites in linkage disequilibrium thereto.
26. Anthracyclines for use in a method of treating a neoplastic disease in a subject in need thereof, the method comprising:
- (a) selecting a subject having a reduced risk of developing cardiotoxicity, wherein cardiotoxicity is based on the identity of a single nucleotide polymorphism (SNP) at one or more of the following polymorphic sites: rs7853758; rs885004; rs10426377; rs2305364; rs4982753; rs4149178; rs4148808; rs1149222; rs17583889; rs4877847; rs11625724; rs12882406; rs12896494; and rs10426628; or one or more polymorphic sites in linkage disequilibrium thereto; and
- (b) administering said subject one or more anthracyclines.
27. The anthracycline of claim 26, wherein the subject also has a reduced cardiotoxicity risk genotype at rs17863783 or one or more polymorphic sites in linkage disequilibrium thereto.
28. The anthracycline of claim 26 or 27, wherein the anthracycline is selected from one or more of the following: anthracycline antibiotics such as daunorubicin (daunomycin, rubidomycin), doxorubicin, idarubicin, epirubicin, mitoxantrone, carminomycin, esorubicin, quelamycin, aclarubicin, esorubicin, zorubicin, pirarubicin, amrubicin, iododoxorubicin, detorubicin, marcellomycin, Rodorubicin, and valrubicin.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2012/000529

A. CLASSIFICATION OF SUBJECT MATTER

IPC: **C40B 40/06** (2006.01) , **A61K 31/65** (2006.01) , **A61P 35/00** (2006.01) , **C07H 21/00** (2006.01) , **C12Q 1/68** (2006.01) , **C40B 30/04** (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC (2006.01): C40B 40/06, A61K 31/65, A61P 35/00, C07H 21/00, C12Q 1/68, C40B 30/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)

Databases: EPODOC, PubMed, Canadian Patent Database

Keywords: anthracycline, rs7853758, rs885004, rs10426377, rs2305364, rs4982753, rs4149178, rs4148808, rs1149222, rs17583889, rs4877847, rs11625724, rs12882406, rs12896494, rs10426628, rs17863783, SLC28A3, SULT2B1, SLC28A1, SLC22A17, SLC22A7, ABCB4, HNMT, UGT1A6, polymorphism, SNP, toxicity, cancer

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2008/058394 A1 (HAYDEN et al.) 22 May 2008 (22-05-2008) *Abstract, page 2 line 33 - page 3 line 3; page 6 lines 5-11; page 12 lines 1-8; page 12 line 35 - page 13 line 28; page 20 lines 20-24; Tables 1 and 2; page 34 line 32 - page 36 line 22; page 40 lines 20-28; page 41 lines 4-16; Example; claims*	1-5, 8 and 18-28
A	INNOCENTI et al. Epirubicin glucuronidation is catalyzed by human UDP-glucuronosyl transferase 2B7. Drug Metabolism and Disposition. 2001, Vol. 29, No. 5, Pages 686-692, ISSN 0090-9556. *Whole document*	1-28
A	RAJIC et al. Influence of the polymorphism in candidate genes on late cardiac damage in patients treated due to acute leukemia in childhood. Leukemia & Lymphoma. October 2009 (10-2009), Vol. 50, No. 10, Pages 1693-1698, ISSN 1042-8194. *Whole document*	1-28

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

08 August 2012 (08-08-2012)

Date of mailing of the international search report

29 August 2012 (29-08-2012)

Name and mailing address of the ISA/CA
Canadian Intellectual Property Office
Place du Portage I, C114 - 1st Floor, Box PCT
50 Victoria Street
Gatineau, Quebec K1A 0C9
Facsimile No.: 001-819-953-2476

Authorized officer

Pascal Lachance (819) 994-8889

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/CA2012/000529**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of the first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons :

1. ☒ Claim Nos. : 5 and 14

because they relate to subject matter not required to be searched by this Authority, namely :

Claims 5 and 14 are directed to a method of medical treatment of the human body by surgery or therapy which the international authority is not required to search under **Rule 39.1 (iv) of the PCT**. However, this authority has carried out a search based on the alleged effect or purpose/use of administering anthracycline according to a subject's risk of developing cardiotoxicity, said risk being determined by the method of screening using one of more of the polymorphisms of the instant claims.

2. ☒ Claim Nos. : 1-3, 9-12 and 23-26 (all partially)

because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically :

See extra sheet.

3. ☐ Claim Nos. :

because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows :

See extra sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claim Nos. :
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim Nos. :

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2012/000529

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ROSSI et al. Analysis of the host pharmacogenic background for prediction of outcome and toxicity in diffuse large B-cell lymphoma treated with R-CHOP21. Leukemia. 29 January 2009 (29-01-2009), Vol. 23, Pages 1118-1126, ISSN 0887-6924. *Whole document*	1-28
A	BLANCO et al. Genetic polymorphisms in the carbonyl reductase 3 gene CBR3 and the NAD(P)H:Quinone oxidoreductase 1 gene NQO1 in patients who developed anthracycline-related congestive heart failure after childhood cancer. Cancer. 5 May 2008 (05-05-2008), Vol. 112, No. 12, Pages 2789-2795, ISSN 0008-543X. *Whole document*	1-28
X	MINOTTI et al. Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. Pharmacological Reviews. 2004, Vol. 56, No. 2, Pages 185-229, ISSN 0031-6997. *Whole document*	26-28
P, X	VISSCHER et al. Pharmacogenomic prediction of anthracycline-induced cardiotoxicity in children. Journal of Clinical Oncology. 1 May 2012 (01-05-2012), Vol. 30, No. 13, Pages 1422-1428, ISSN 0732-183X. *Whole document*	1-2, 4-5, 8 and 18-28

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CA2012/000529

Patent Document Cited in Search Report	Publication Date	Patent Family Member(s)	Publication Date
WO2008058394A1	22 May 2008 (22-05-2008)	AU2007321678A1	22 May 2008 (22-05-2008)
		CA2669131A1	22 May 2008 (22-05-2008)
		EP2102392A1	23 September 2009 (23-09-2009)
		EP2102392A4	14 July 2010 (14-07-2010)
		EP2312024A1	20 April 2011 (20-04-2011)
		JP2010508858A	25 March 2010 (25-03-2010)
		US2011312907A1	22 December 2011 (22-12-2011)

Continuation of Box No. II.2:

Claims 1-3, 9-12 and 23-26 feature polymorphic sites defined by their desired property of being “in linkage disequilibrium” with the polymorphisms listed in said claims. The putative polymorphic sites encompassed by this desired property are not identified in the description. Consequently, they are not supported by the description and fail to comply with the requirements of **Article 5 of the PCT** and **Article 6 of the PCT** to such an extent that a meaningful search over the entire claimed scope is not possible.

The scope of the search and examination was therefore limited to the subject matter of claims 1-3, 9-12 and 23-26 (all partially) insofar as they feature the polymorphisms specified in the claims, namely rs7853758, rs885004, rs10426377, rs2305364, rs4982753, rs4149178, rs4148808, rs1149222, rs17583889, rs4877847, rs11625724, rs12882406, rs12896494, rs10426628 and rs17863783.

Continuation of Box No. III:

Claims 1-28 are directed to a plurality of separate inventions, wherein each group of invention features one of the specific combinations of one or more diagnostic polymorphisms for assessing the susceptibility of a subject to the development of cardiotoxicity in response to receiving anthracycline compounds for the treatment of a neoplastic disease, said polymorphisms comprising one or more of rs7853758, rs885004, rs10426377, rs2305364, rs4982753, rs4149178, rs4148808, rs1149222, rs17583889, rs4877847, rs11625724, rs12882406, rs12896494, rs10426628 and rs17863783.

The polymorphisms have different structural features, therefore the only technical feature which might serve to unify the claims is the use of combinations of one or more polymorphisms for assessing the susceptibility of a subject to the development of cardiotoxicity in response to receiving anthracycline compounds. However, the use of combinations of one or more polymorphisms for assessing the susceptibility of a subject to the development of cardiotoxicity in response to receiving anthracycline compounds is known in the art, as disclosed for example in each of HAYDEN et al, INNOCENTI et al., RAJIC et al., ROSSI et al., and BLANCO et al. The claims thus do not relate to a single inventive concept *a posteriori* under **Rule 13.1 of the PCT**. Consequently, the claims are directed to a plurality of alleged inventions wherein each specific combination of one or more of the polymorphisms featured in claims 1-28 represents separate inventions.

The claims must be limited to one inventive concept as set out in **Rule 13 of the PCT**.