Genetic susceptibility for environmental carcinogenesis : a biomarker approach

Citation for published version (APA):

Ketelslegers, H. (2007). *Genetic susceptibility for environmental carcinogenesis : a biomarker approach*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20071004hk>

Document status and date:

Published: 01/01/2007

DOI:

[10.26481/dis.20071004hk](https://doi.org/10.26481/dis.20071004hk)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

**Genetic susceptibility for environmental
carcinogenesis:**

A biomarker approach

Hans Bernardus Ketelslegers

ISBN: 978-90-9022236-3

© Hans B. Ketelslegers, Maastricht 2007

Omslagillustratie en lay-out: Antje Schöne en Hans Ketelslegers

Omslagrealisatie en druk: QuickPrint, Nijmegen

The studies presented in this thesis were performed at the Nutrition and Toxicology Research Institute Maastricht (NUTRIM), which participates in the Graduate School VLAG (Food Technology, Agrobiotechnology, Nutrition and Health Sciences), accredited by the Royal Netherlands Academy of Arts and Sciences.

Genetic susceptibility for environmental carcinogenesis:

A biomarker approach

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. mr. G.P.M.F. Mols,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen op
donderdag 4 oktober 2007 om 14:00 uur

door

Hans Bernardus Ketelslegers

geboren te Leunen op 15 december 1978

Promotor:

Prof. dr. J.C.S. Kleinjans

Copromotor:

Dr. J.H.M. van Delft

Beoordelingscommissie:

Prof. dr. J.P.M. Geraedts (voorzitter)

Prof. dr. P.A. van den Brandt

Prof. dr. N.A. van Larebeke (Universiteit Gent, België)

Prof. dr. E.C.M. Mariman

Dr. W.K.W.H. Wodzig

The research presented in this thesis was conducted as part of and therefore financed by the Flanders Environment and Health Survey (Steunpunt Milieu en Gezondheid), currently known as the Flemish Center of Expertise of Environment and Health.

Contents

<i>Chapter 1</i>	General Introduction	11
	1.1 Genetic cancer susceptibility	
	1.1.1 Genetic variation: Single Nucleotide Polymorphisms	
	1.1.2 Implications for toxicology: Toxicogenetics	
	1.2 The role of SNPs in molecular cancer epidemiology	
	1.2.1 Biomarkers of exposure and SNPs	
	1.2.2 Biomarkers of effect and SNPs	
	1.3 Genotyping as biomarker for genetic susceptibility	
	1.3.1 Monitoring genetic susceptibility: a historic point of view	
	1.3.2 Genotyping techniques	
	1.3.3 Genotyping techniques: selection issues	
	1.4 Aim and outline of the thesis	
<i>Chapter 2</i>	Simultaneous genotyping of nine polymorphisms in xenobiotic metabolizing enzymes using multiplex PCR amplification and Single Base Extension	41
<i>Chapter 3</i>	Inter individual variations in DNA adduct levels assessed by analysis of multiple genetic polymorphisms in smokers	59
<i>Chapter 4</i>	Potential role of CYP1B1 in the metabolic activation of 4-AminoBiPhenyl in humans	77
<i>Chapter 5</i>	Multiplex genotyping as a biomarker for susceptibility to identify subgroups at increased risk in carcinogenic exposure-effect relationships in the FLEHS biomonitoring study	93
<i>Chapter 6</i>	The prevalence of an at-risk genotype for environmental carcinogenesis decreases with age; Survival of the fittest?	117
<i>Chapter 7</i>	Summary and General discussion	133
	Samenvatting en Algemene discussie	143
	Dankwoord	153
	Curriculum vitae	157
	List of publications	161

Abbreviations

(d)dNTP	(di)DeoxyNucleotideTriPhosphate
1-OHP	1-Hydroxypyrene
4-ABP	4-AminoBiPhenyl
8-OhdG	8-Hydroxy-2'-deoxyGuanosine
AA	Aromatic Amine
ASO	Allele Specific Oligonucleotide
BER	Base Excision repair
CA	Chromosome Abberations
DNA	DeoxyriboNucleic Acid
DSB	Double Strand Break repair
EDTA	Ethylene Diamine Tetraacetic Acid
FLEHS	Flanders Environment and Health Study
FRET	Fluorescence Resonance Energy Transfer
GCMS	Gas Chromatography Mass-Spectrometry
GSH	Glutathione
HapMap	Haplotype Map
Hb	Hemoglobin
HPRT	Hypoxantine-guanine PhosphoRibosyl Transferase
IARC	International Agency for Research on Cancer
<i>i.e.</i>	<i>id est</i> (that is)
MMR	Mis Match Repair
MN	MicroNuclei
mRNA	messenger RiboNucleic Acid
NER	Nucleotide Excision repair
OLA	Oligonucleotide Ligation Assay
PAH	Polycyclic Aromatic Hydrocarbon
PCR	Polymerase Chain Reaction
RFLP	Restriction Fragment Length Polymorphism
ROS	Reactive Oxygen Species
rs#	RefSNP accession ID
SBE	Single Base Extension
SCE	Sister Chromotid Exchange
SNP	Single Nucleotide Polymorphism
T _m	(DNA) Melting Temperature
tt-MA	<i>trans-trans</i> Muconic Acid
XP	Xeroderma Pigmentosum

"You're innocent when you dream"

Tom Waits

Chapter 1

General Introduction

Chapter I

When Charles Darwin published “The Origin of Species” in 1859, he observed that no individual within a population is the same, although these differences tend to fluctuate around a certain mean. Darwin first described these observations in animals and plants. However, although he could not read the “book of life (1,2)”, and thus had no knowledge of the underlying genetic processes, he soon saw that humans could not be withdrawn from the evolutionary process he described. Twelve years later, he revealed his controversial theory in which he described that man were nothing more than a descendent of former mammals (The Descent of Man, 1871). He described this process of evolution in terms of “natural selection”; slow and gradual changes in heritable factors as reaction to different environments. Within a population, individuals who had adapted better to the environment would survive.

“No one supposes that all individuals of the same species are cast in the very same mould. These individual differences are highly important to us”.

Charles Darwin, 1859

After further studies on inheritance by Gregor Mendel in the second half of the 19th century and two big milestones in biology in the twentieth century, namely the discovery of the DNA double helix (3) and the mapping of the complete genome (1, 4), it is now clear that the described heritable factors are genes, encoded in the DNA, of which different forms exist (alleles) resulting from common variations in this genetic code between individuals. The largest part of these common variations consists of polymorphisms, which reflect past mutations implemented during evolution. Each individual carries a “genetic passport”, which comprises a unique combination of alleles inherited from both parents, the genotype. As a result, no individual is the same (except for homozygous twins), and because of the information encoded in this genotype, each individual reacts differently to exogenous factors. Therefore, interest is increasing in the role which these variations in the human genome play in the inter-individual susceptibility in disease development, among others as a result of exposure to hazardous environmental pollutants: an alternative way of explaining Darwin’s natural selection theory.

1.1 Genetic Cancer Susceptibility

Although cancer is a genetic disease, only a very small portion of all cancers is a result of solely genetic factors (5). The majority of all cancers results from an interplay between genetic and environmental factors like, for example, tobacco smoking and pollutants: often referred to as gene-environment interactions. This is an optimistic message for cancer prevention policies, as these exposures are theoretically preventable (6). However, regarding the continued growth and ageing of the world population, it is estimated that by 2030 there will be 27 million new cancer cases diagnosed, 17 million deaths of cancer annually worldwide (more than 45000 every day) and 75 million persons alive with cancer (7). Therefore, the pressure is great to translate this current

knowledge into public health interventions. Prevention has to be carried out at two levels, namely in modifying the hazardous lifestyles of individuals (e.g. cigarette smoking and dietary consumptions such as alcohol and burned meat) and in reducing involuntary exposures to carcinogens (e.g. industrial exposures) usually through regulation (8). Historically, policy makers assumed that all individuals in a population have the same biological response to a certain dose of environmental carcinogens. Today, policy makers supporting biomonitoring studies, such as the Flanders Environment & Health Study (FLEHS, www.milieu-en-gezondheid.be), are considering the possibility to account for inter-individual variations in susceptibility for the response to (carcinogenic) environmental exposures. Investigating genetic variation can provide this information about susceptibility for diseases such as cancer.

The development of cancer induced by genotoxic environmental components, so called environmental carcinogenesis, is a slow and gradual complex process that takes many years or even decades before the actual disease can be clinically diagnosed. It comprises several steps, shortly initiation, promotion, progression and transformation, which are characterized by genetic changes (9). The initial step, the transformation of a normal “healthy cell” to an “initiated cell”, is assumed to be generally induced by the covalent binding of highly reactive metabolic intermediates to the DNA, which results in DNA damage. The latter may eventually lead to irreversible changes, causing mutations in critical genes involved in cell-cycle regulation processes. Several pathways are involved in this initial process, from modulation of the initial exposure to the modulation of DNA damage. These pathways, and moreover how genetic variation between individuals may explain why at approximately the same exposure level to carcinogenic compounds, one individual develops cancer whereas another one does not, will be described in this paragraph.

1.1.1 Genetic variation: Single Nucleotide Polymorphisms

The sequencing of the human genome revealed that 99.9% of the three billion base pairs of DNA is the same between individuals. However, the information encoded in this 0.1% that is different has profound effects on biological makeup. The most dramatic examples are so called “highly penetrant mutations” in genes directly involved in disease development, as is seen with inherited disorders (8, 10, 11). These mutations are rare, and for cancer development generally interfere directly with e.g. cell cycle regulation or differentiation without interacting with external exposures. A well known example of a “high risk” genetic disorder is the development of breast cancer in woman carrying inherited defects in the so called breast cancer genes *BRCA1* and *BRCA2* (both related to a deficiency in the repair of DNA damage, due to the loss/inactivation of these tumor-suppressor genes) (12-14). Because these highly penetrating polymorphisms are directly associated with disease state, they can give information about the *mechanistic basis* for disease. In contrast, functional polymorphisms that are of lower impact, for instance alter the effect of environmental exposures that may lead to disease states or

Chapter 1

their precursors, and hence can give information about the *susceptibility* for environmental disease. The individual risk associated with the latter is often low, but the population attributable risk is greater because of their high population frequency, and is therefore potentially of greater public health relevance.

Investigation of the human genome revealed that the largest part of the variability in the human genome, about 90% of the sequence variation, consists of Single Nucleotide Polymorphisms (SNPs) (2, 4). SNPs are single base pair positions in genomic DNA at which different sequence alternatives (gene variants, or alleles) exist in normal, healthy individuals. These alleles comprise the differences in phenotype between individuals, and may therefore also explain an individual's susceptibility for certain diseases or health effects. A human SNP map was published in 2001 by a public consortium, who found that SNPs occur, on average, about once every 1000-2000 bp (2). The largest part of these polymorphisms are biologically silent (about 99%), and the biological relevance of the remainder is dependent on the position where it occurs within the genome. The different ways in which polymorphisms can alter biological function, is shown in Figure 1 (15). If a SNP occurs within a coding region of a gene, it can result in a protein that has an amino acid substitution, which may alter catalytic activity, enzyme stability and/or substrate specificity. The same effects can be observed in splice site variants, which can give rise to alternatively spliced or truncated proteins.

On the other hand, SNPs in non-coding, regulatory regions of a gene can also have a significant impact on protein function. Polymorphisms near intron-exon junctions can affect mRNA stability or mRNA splicing, whereas those that fall in promoter regions may change the regulation and level of a protein. Last, polymorphisms can also involve larger segments of the genome and lead to gene-deletions, hence no gene product and loss of protein function, or on the contrary (multi-)duplicated genes which leads to higher enzyme levels.

1.1.2 Implications for toxicology: Toxicogenetics

As was described in the last paragraph, it is important to understand that susceptibility genes are neither necessary nor sufficient to cause disease, they *modify risk* (16). In other words: they provide the genetic basis for inter-individual differences in response to potential toxicants, a study field that has evolved in the past decade as *Toxicogenetics* (15). How genetic factors relate to environmental exposures, was strikingly described by Olden and Guthrie (16):

“Conceptually, the relationship between genes and the environment is similar to that of a loaded gun and its trigger. A loaded gun itself causes no harm; it is only when the trigger is pulled that the potential for harm is released or initiated. Likewise, one can inherit a predisposition for a devastating disease, yet never develop the disease unless exposed to the environmental trigger(s)”.

Figure 1. The effects of polymorphisms in the genome on protein function. SNPs that fall in the coding region of a gene may alter protein activity resulting from an amino acid substitution, altered splice variant or truncated protein. Complete loss of protein function or, in contrast, an increase in protein activity is caused by gene deletions or duplications respectively (adapted from Orphanides and Kimber (15)).

Chapter 1

Thus, most health effects are the result of an unfavorable combination of an individual's genotype and environmental factors, which is particularly true at low doses of the environmental toxicant, as was discussed previously (17). Yet, how can genetic factors influence the classic principles of toxicology, i.e. how is the trigger pulled?

Ironically, most compounds that are carcinogenic to humans (classic examples are Polycyclic Aromatic Hydrocarbons (PAH), present e.g. in cigarette smoke, traffic exhaust and dietary factors such as burned meat, and Aromatic Amines (AA), also present in cigarette smoke and several environmental media) need to be metabolically activated to exert their carcinogenic effect. These ultimate carcinogens, the reactive metabolites produced during biotransformation, can bind covalently to the DNA thereby forming DNA damage. All living mammalian cells have evolved efficient mechanisms to cope with the presence and to modulate the adverse influences of exposure to toxic compounds and accumulation of the possible induced DNA damage. These mechanisms are genetically controlled, and it is thought that polymorphisms in genes involved in these mechanisms may modify environmental disease risk (18). The main pathways in the study field of toxicogenetics are biotransformation, DNA repair and oxidative stress.

BIOTRANSFORMATION

The process of biotransformation, that is the chemical alteration of the xenobiotic compound to facilitate excretion, can be divided into 2 phases (Figure 2).

Metabolic activation of the carcinogenic compound to the ultimate carcinogen occurs in phase I, during which the involved enzymes introduce new (or modify existing) functional chemical entities to xenobiotics (19). These reactions are primary metabolized by cytochrome P450s (CYPs), which catalyze so called oxidation reactions, the addition of molecular oxygen into a substrate. The highly reactive metabolites that are produced, are subsequently *detoxified* in phase II by the conjugation of endogenous ligands to the activated compounds, which increases the hydrophilic nature of it and thereby facilitates urinary excretion. The main groups of phase II enzymes include glutathione *S*-transferases (GSTs), mainly involved in the detoxification of PAH, and *N*-acetyl-transferases (NAT1 and NAT2) which deactivate aromatic amines by acetylation.

An optimal balance between phase I and II biotransformation reactions is necessary for an appropriate detoxification and elimination of the xenobiotic compound. As can be seen in Figure 2, an imbalance may lead to DNA damage, and eventually health effects, induced by the electrophilic metabolites formed during biotransformation. Because many of the genes encoding for Phase I and II enzymes, are polymorphic, functional SNPs in genes involved in these reactions may explain inter-individual differences in health risk as a result of exposure to toxicants. Indeed, several metabolic polymorphisms have been associated with cancer risk. Genes that code for the CYP450 isoenzymes CYP1A1, CYP1A2 and CYP1B1 have been associated with an increased risk for lung cancer (20-24). Moreover, these enzymes have been shown to be induced by exposure to cigarette smoke, and a number of studies conducted in smokers has described an association between polymorphisms in *CYP* genes and several types of can-

Figure 2. Implementation of genetic polymorphisms as a biomarker for susceptibility in molecular epidemiology and environmental carcinogenesis. The initial step in environmental carcinogenesis is a mutation in critical cancer related genes. The process from environmental carcinogen exposure to the initial effect, can be measured at several intermediate endpoints (molecular epidemiological biomarkers). Genetic variants in genes (alleles) involved in biotransformation, DNA repair and oxidative stress pathways may affect every stage of the process, and may thereby explain why one individual develops cancer, whereas another one doesn't after exposure to approximately the same level of carcinogens. Examples for biomarkers of exposure and effect are shown in bold and explained in de text.

cer risk (25-30), reviewed by (31) and (32). A classic example is *CYP1A1*, of which the gene product, aryl hydrocarbon hydroxylase (AHH), catalyzes the first step of the metabolism of PAH to electrophilic compounds. *CYP1A1* is induced by cigarette smoking and environmental exposures such as dioxins and PAH (33), and the level of *CYP1A1* inducibility varies among individuals. Several polymorphisms have been described in *CYP1A1* (34, 35), but the first mutation (the so called *MspI* restriction site) was found by Kawajiri (22) and linked to a polymorphism in exon 7 leading to a *Iso/Val* aminoacid change (36). The valine-type protein is almost twice as active as the wildtype in metabolizing PAH, and has been associated with an increased risk for lung cancer in a Japanese population (37, 38).

A large number of studies have reported an association between functional polymorphisms in genes involved in phase II detoxification reactions and cancer risk. Functional polymorphisms in isoenzymes of GST have been linked to the risk for several types of cancers (39, 40). For example, the *GSTM1* gene is deleted in 50% of the population and individuals carrying this polymorphism have no *GSTM1* activity. This *GSTM1 null* variant has been linked with an increased risk for bladder and lung cancers (41-44). Regarding gene-environment interactions, Bell et al (44) showed that in individuals carrying the *GSTM1 null* variant the risk for bladder cancer was small in the absence of environmental exposure. This risk was dramatically increased upon exposure to tobacco smoke. At least two *NAT* genes are known, *NAT1* and *NAT2*, of which *NAT2* is the most well known as it characterizes the so called slow and fast acetylator phenotype (45). More than 50% of the Caucasian population and 30-40% of African/Americans are slow acetylators (46, 47). Association studies investigating *NAT2* polymorphisms show that slow acetylators have an increased risk for developing bladder cancer due to environmental and occupational exposures to aromatic amines such as 4-ABP (46, 48-50), whereas they have a lower risk for colorectal cancer (51). For more information is referred to some reviews (52-56).

DNA REPAIR

As was already brought up in the last paragraph, the highly reactive intermediate products produced by metabolic activation may bind covalently, and thereby induce damage, to the DNA. Ironically, several endogenous processes can also cause DNA damage (57). The major type this endogenous DNA damage results from oxidation of DNA bases, caused by reactive oxygen species which are continuously produced during respiration in all living cells. Fortunately, the genome carries information to protect its own integrity (58). Several DNA repair pathways have evolved to reduce the mutation frequency of cancer related genes (59), predominantly base- and nucleotide excision repair (BER and NER respectively), mis-match repair (MMR) and double strand break (DSB) repair (60). DNA repair enzymes involved in these pathways, continuously monitor the chromosomes for undesirable changes in the DNA sequence. Normally, the cell cycle will be delayed after toxic exposures which allows the cell to repair the damage DNA and normal cell functions will be maintained (61). If DNA repair is absent or insufficient,

division of these cells possessing the damaged DNA will proceed which eventually leads to irreversible damage, i.e. mutations. If these mutations occur in critical genes related to carcinogenesis, such as (proto-)oncogenes and tumor suppressor genes, damaged cells may lead to initiation of the carcinogenic process and/or stimulate progression (59). Functional polymorphisms in genes involved in these repair pathways may affect an individual's capacity to repair DNA damage and therefore contribute to cancer susceptibility (62). The most well known examples are seen in rare genetic disorders (63), such as functional polymorphism in the double strand break repair genes *BRCA1* and *BRCA2*, which are related to breast cancer (13, 64). Another example is the severe reduction in nucleotide excision repair capacity due to polymorphisms in the *XPD* genes, displayed in xeroderma pigmentosum (XP) patients, which is related to an increased sensitivity to UV radiation and development of skin cancer (65). In contrast, more common low penetrance polymorphisms in DNA repair genes have also been associated with cancer risk, such as *XRCC1*, an important enzyme involved in BER, which comprises three SNPs ($R^{194}W$, $R^{399}Q$ and $R^{280}H$). Studies investigating the *XRCC1* $R^{194}W$ polymorphism described a reduced risk for cancer in individuals carrying the variant allele, possibly related to cigarette smoke and/or alcohol exposure (66-69). Regarding NER, polymorphisms in the *XPD* gene are intensively studied. Although mutations in the *XPD* gene are associated with disease (XP), the most commonly studied polymorphisms $L^{751}Q$ and $D^{312}N$ in the *XPD* gene have been associated with an increased risk of lung cancer, for $D^{312}N$ variant also after exposure to cigarette smoke (70, 71). A commonly investigated DSB repair gene in relation to cancer risk is *XRCC3*. The $T^{241}M$ polymorphism in *XRCC3* is associated with a decrease in repair capacity which results in a higher risk for bladder cancer and melanoma (72, 73).

OXIDATIVE STRESS

Humans require oxygen for energetic processes within the cell. One major side effect of the use of oxygen, is that it can produce reactive oxygen species in its conversion to water in the cell. The first of these so called reactive oxygen species (ROS) that is produced is the superoxide anion, which in turn generated other toxic metabolites including hydrogen peroxide, hydroxyl radical and peroxynitrite. These ROS are continuously produced during normal and essential endogenous reactions, such as respiration in the mitochondria (74). Moreover, free radicals produced as side effects of xenobiotic metabolism or released by macrophages during inflammatory response are also regarded as endogenous, without the presence of an exogenous insult (75). Although endogenous reactions are regarded as the main source of ROS production, several exogenous factors may also contribute to the level of ROS within the cell, for example radiation and exposure to environmental pollutants, including emission from traffic and industries, cigarette smoking and alcohol consumption. Irrespective of their origin, ROS may interact with DNA leading to oxidative DNA damage. A steady state level of damage representing a minimal risk to the cell or organism, is maintained by a balance between the rate at which the oxidative damage is produced and by which it is efficiently repaired and

removed. The former is dependent on how fast ROS are generated and subsequently inactivated by anti-oxidant defense mechanisms. Several endogenous defense mechanisms have evolved, comprising a number of antioxidant enzymes.

The primary defenses against the toxic effects of ROS are superoxide dismutase (SOD), the only enzymatic system which catalyzes the conversion of the superoxide anion to hydrogenperoxide, catalase (CAT) and glutathione peroxidase (GPX) which reacts with the powerful reductant glutathione (GSH). Both CAT and GPX catalyze the conversion of hydrogenperoxide to water, and without this enzymatic conversion hydrogenperoxide may form the toxic hydroxyl radical or lead to cell toxicity itself. In spite of these defense mechanisms, DNA damage from ROS still occurs and the rate at which it is removed is dependent on DNA repair enzymes. Several repair pathways, mainly the BER and NER pathways mentioned above, are involved in the repair of oxidative DNA damage (76). If protection against ROS or repair of oxidized DNA is insufficient, the balance is shifted in favor of the inducement of oxidative DNA damage, called oxidative stress. Oxidative stress has been linked to aging and various (degenerative) diseases, including heart disease and cancer (75, 77-79).

It is believed that functional polymorphisms in genes coding for enzymes involved in anti-oxidant, DNA repair and also xenobiotic metabolism pathways may contribute to susceptibility for oxidative stress and thereby disease risk (77). For example, Moscow et al. described a polymorphism in *GPX1* at position 198 leading to a *Pro/Leu* amino acid change which is associated with lung cancer risk among Caucasian exposed to cigarette smoke (80). A higher risk for lung cancer was also observed in individuals carrying the *Val/Ala* aminoacid substitution in *codon 16* of *mnSOD*, as described in a study conducted in 1100 lung cancer patients (81). Last, a common polymorphism in the promoter region of *Catalase* at position -262 was associated with a reduction in Catalase activity and an increase in Breast cancer risk (82). The thus protective effect on breast cancer of the *catalase wildtype* variant, was more pronounced in women who did not take vitamin supplements but who had a high fruit intake. This implies an important role for diet in breast cancer prevention, influenced by the *Catalase* polymorphism. Most studies investigating the repair of oxidative DNA damage, have focused on the frequently observed missense change in *codon 326* at exon 7 of *OGG1*, which results in a *Ser³²⁶Cys* amino acid change (60). Individuals carrying this polymorphism are associated with the risk for developing several types of cancer, including lung, prostate and esophageal cancer (83-85).

It may be obvious, that it is unlikely that a single polymorphism, or a single pathway, determines an individual's integrative response to a toxic environmental exposure and susceptibility for disease risk. All described pathways are related to each other, and polymorphisms in genes involved in these pathways may contribute simultaneously to the ultimate effect in response to an (environmental) exposure. To investigate the interaction between the involved pathways, and also regarding a specific environmental

pollutant, it is considered useful to study toxicogenetics in a molecular epidemiology approach.

1.2 The role of SNPs in Molecular Cancer Epidemiology

The first step of the carcinogenic process, described in the last paragraph, can be monitored at several intermediate levels between exposure to the environmental compound and the ultimate health effect, studied by Molecular Epidemiology, a research field first introduced by Perera and Weinstein ((86), Figure 2). There is no general accepted definition of molecular epidemiology, but from a biologists point of view, it comprises the application of molecular biology tools to measure biological variables (biomonitoring) relevant for epidemiology, into epidemiological research (87, 88). These variables are called biological markers or biomarkers, which can be any measurable substance, structure or process in the body or its products that influences health effects. As can be seen in Figure 2, these biomarkers are intermediate endpoints in the molecular biology framework and can be classified into markers of exposure, effect and susceptibility.

Markers for susceptibility refer to the inter-individual differences which determine whether an individual is particularly sensitive at a certain stage of the carcinogenic process between exposure and the ultimate health effect. A distinction can be made between genetic and “acquired” susceptibility. The latter refers to factors influencing susceptibility such as age, gender, ethnicity, health- and nutritional status. For molecular epidemiology, the most interesting marker for susceptibility is genetic predisposition. As was discussed in the last paragraph, SNPs involved in pathways which contribute to the environmental carcinogenic process, from exposure to the carcinogenic compound to the modulation of DNA damage, may explain the individual susceptibility for cancer development. Implementation of the role of these pathways in the molecular epidemiology framework, can explain how genetic predisposition as marker for susceptibility may affect crucial analyses in biological monitoring, and may thus explain inter-individual differences in biomarker levels. The role which SNPs play in gene-environment interactions will be discussed for each biomarker separately.

1.2.1 Biomarkers of exposure and SNPs

Biomarkers of exposure can be subdivided into markers of external exposure dose, internal dose and biologically effective dose. The external exposure dose is the determination and measurement of the concentration of the compound in the external medium in which it is contained. This is done in a study field called environmental monitoring, which is not influenced by genetic predisposition. However, a more accurate measurement of the uptake of a (carcinogenic) compound in an exposed individual, is by measuring the internal exposure of a compound, which is done in biological monitoring or biomonitoring. Biological markers of exposure comprise the measurement of the amount of chemicals effectively absorbed by an individual, i.e. the internal dose, and

the amount of chemicals bound to target sites, i.e. the biologically effective dose. The most commonly used biological materials are whole blood, plasma and urine.

INTERNAL DOSE

Classical biomarkers in urine for the assessment of exposure to carcinogens reflecting the internal dose include the measurement of 1-hydroxy-pyrene (1-OHP), as marker of exposure to PAH, and tt-muconic acid (tt-MA), for benzene exposure (89, 90). Both biomarkers are so called “urinary metabolites”, which are excreted products in the urine resulting from biotransformation of the original compound. It is therefore postulated that SNPs in genes involved in biotransformation processes may affect these urinary biomarker levels (53).

A method to measure 1-OHP in urine was pioneered by Jongeneelen et al (91). Since then, variations on this method have been described and numerous studies have reported a relationship between urinary 1-OHP and environmental, occupational and individual PAH exposures, such as cigarette smoking and roasted meat consumption (92, 93). Exposure to PAH has been associated with higher levels of DNA damage (the formation of so called PAH DNA adducts, which are used as a biomarker for biological effective dose) and cancer risk in humans (94). Phase I and Phase II metabolizing enzymes are involved in the biotransformation of PAH, and several studies described an influence of SNPs in *CYP1A1* and *GSTM1 null* variants (35, 95). For example, Alexandrie et al (96) reported that the observed variations in 1-OHP levels in the urine of occupational PAH exposed aluminum smelter workers could partly be explained by polymorphisms in *CYP1A1*, *GSTM1* and *GSTT1* enzymes. Moreover, the highest levels of urinary 1-OHP were detected in individuals carrying the *CYP1A1 Ile/Val* genotype, leading to a higher Phase 1 enzyme activity, who also lacked *GSTM1* (Phase 2) activity. More recently, Cocco et al (93) found in a non-occupationally PAH exposed study population, that variations in background urinary 1-OHP levels could be explained by living in a heavily trafficked urban area and several dietary habits, namely the frequent intake of grilled meat and alcohol intake. Changes in 1-OH excretion were reported to be the highest in subjects lacking the detoxifying activity of *GSTT1*, after adjusting for smoking. This effect could not be observed in a healthy, non-occupational exposed Brazilian population. However, a clear effect of SNPs in the Phase1 activating enzyme *CYP1B1*, and to some degree polymorphisms in *CYP1A1* and *CYP1A2* enzymes and the Aryl Hydrocarbon (Ah) receptor (which up-regulates the expression of CYPs after PAH binding) on urinary 1-OHP variations(97) was observed.

Benzene is one of the most widely used industrial chemicals and reported to be carcinogenic to humans (98). Moreover, benzene is present in a high concentration in mainstream cigarette smoke and urban air resulting from traffic exhaust (99, 100). tt-MA, a urinary metabolite of benzene, has been widely used as a biomarker for occupational and environmental benzene exposure (101-103). It has been related to DNA damage and risk for cancer development. Variations in urinary tt-MA levels have been reported to be explained by several acquired and genetic susceptibility factors (104-106).

Significantly higher tt-MA levels in the urine after exposure to urban air were observed in individuals lacking GSTM1 activity (107). Moreover, in a study conducted in bus drivers professionally exposed to benzene from vehicle exhaust, a higher excretion of tt-MA was found in subjects with the *GSTT1 null* genotype (105). More recently, in a study conducted in Thai workers occupationally exposed to benzene, higher levels of tt-MA in urine were observed in individuals carrying the *CYP2E1**5 polymorphism which is associated with a higher Phase 1 activity (108).

BIOLOGICALLY EFFECTIVE DOSE

Typical biomarkers for the measurement of the biologically effective dose are adducts formed in DNA or proteins. Metabolites resulting from the biotransformation of environmental carcinogens, such as PAH and AA, are able to bind covalently to DNA and proteins to form adducts (109, 110). Because these markers are thought to be closely related to cancer risk, e.g. the formation of DNA adducts is an initial event in carcinogenesis as it may reflect mutations in critical genes involved in the carcinogenic process, they are sometimes referred to as markers of early effect. However, the number of cells bearing DNA adducts that undergo transformation after exposure is minimal, as it is dependent on a number of other processes such as DNA repair. It is therefore suggested that variations in DNA adduct levels as a biomarker of effect are partly due to polymorphisms in DNA repair genes, whereas if DNA and protein adducts are regarded as markers of biologically effective dose, polymorphisms in biotransformation enzymes can provide information about differences in these biomarker levels after toxic exposures (53).

Several studies have described an effect of polymorphisms in CYP, NAT and GST biotransformation enzymes on variations in PAH-DNA adducts. For example, adduct levels were found to be higher in individuals carrying the commonly investigated *CYP1A1 Ile/Val* polymorphism (111). Moreover, smoking individuals carrying this polymorphism together with a deficiency in GSTM1 activity were associated with high DNA adduct levels related to benzo(a)pyrene exposure (112). In a study conducted in smokers, the level of aromatic DNA adducts adjusted for the amount of cigarettes smoked per day was higher in individuals lacking the *GSTM1* gene and higher in individuals bearing slow NAT1 or NAT2 acetylator status (113). Moreover, they also found an effect of *GSTT1* together with the *NAT2* genotype, and the highest adduct levels were found in individuals without GSTM1 activity who were also NAT1 and NAT2 slow acetylators. More recently, an increase in DNA adduct levels was found associated with polymorphisms in *GSTM1*, *mEH* and *GPX* together with smoking status (114). In addition, DNA adduct levels were almost 2-fold higher in individuals with a relatively high number of risk alleles compared to wild type individuals for these three genes. Commonly investigated polymorphisms in DNA repair in relation to aromatic DNA adduct levels comprise SNPs in *XRCC* and *XPB* (115-117). For example, a dose response relationship was observed between the number of risk alleles for *XPB Lys⁷⁵¹Gln*, *XRCC1 Arg³⁹⁹Gln* and *XRCC3 Thr²⁴¹Met* polymorphisms and the level of DNA adducts.

Chapter 1

A classic example of a human carcinogen the metabolites of which form protein adducts, is the aromatic amine 4-ABP, which forms 4-ABP-hemoglobin (Hb) adducts (109). 4-ABP-Hb adducts have been widely used as a marker for the internal dose of aromatic amines. Because protein adducts will not be repaired, variations in protein adduct levels will largely depend on polymorphisms in biotransformation related genes. The main enzymes involved in 4-ABP metabolism are NATs, as was described in the last paragraph. Higher 4-ABP-Hb adduct levels were found in NAT slow acetylators, and the effect of the polymorphism is expected to be more pronounced at low exposure doses (49, 113, 118). Moreover, in low dose smokers who are NAT2 slow acetylators and also have the rapid phase 1 *CYP1A2* genotype, a higher level of 4-ABP-Hb adducts was detected as compared to other smokers (50). There is some debate about the main phase I enzyme involved in 4-ABP metabolism, which was thought to be *CYP1A2* although some studies suggest that other CYP enzymes, such as *CYP1A1* or *CYP1B1*, are involved (119, 120).

1.2.2. Biomarkers of effect and SNPs

Biomarkers of effect refer to the measurement of in general irreversible biological and biochemical changes that affect the structure or function of a cell and are therefore directly or indirectly involved in the carcinogenic process. These biomarkers can be subdivided into markers for early biological effect and early response markers. In contrast to the latter, which determine alterations in genes directly involved in the multistep carcinogenesis model, namely mutations in oncogenes and tumor suppressor genes, early biological effect markers measure alterations in events that are not directly involved in the carcinogenic process. That is, they exhibit an (early) effect which is related to cancer risk, as a reaction to a certain level of exposure. These methods are sometimes referred to as markers of genotoxic exposure (121).

EARLY BIOLOGICAL EFFECT

Well known examples of early effect markers are the Comet assay, urinary 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-OH-dG), cytogenetic markers, such as chromosomal aberrations (CA), micronuclei (MN) and sister chromatid exchange (SCE) and tests at gene level such as the hypoxanthine-guanine phosphoribosyl transferase (HPRT) mutation test (122-125). More recently, it was shown that gene-expression patterns may be applicable as a biomarker for early biological effect (126). A number of studies has investigated polymorphisms in biotransformation enzymes in relation to these biomarkers after occupational and environmental exposures, describing an important role for GSTM and GSTT and a possible role for *CYP1A1* and NAT2 (53). In addition, a recent paper suggests that the oxidative damage induced after PAH exposure may be modulated by polymorphisms in *mnSOD* and *MPO* (127).

EARLY RESPONSE

Markers for early response refer to the detection of altered gene products and gene expression patterns in genes responsible for cancer development, namely (proto)oncogenes and tumorsuppressor genes, which are mainly involved in DNA repair, cell cycle regulation, cell growth and differentiation pathways (128). A mutation in these genes may lead to initiation or progression of a normal cell to a cancerous cell.

1.3 Genotyping as biomarker for Genetic Susceptibility

As was described in the former paragraphs, monitoring genetic polymorphisms provides important information on the inter-individual susceptibility in (environmental) cancer development. Several techniques for analyzing SNPs, so called genotyping, have been developed over the past decades and can be used as biomarkers to identify these genetic risk factors. From the analysis of an individual's phenotype as a "surrogate marker" for determination of the genotype in the mid 1970's (129), to the first genotyping techniques after Karry Mullis' invention of the polymerase chain reaction (PCR) in 1984 (130) and the high throughput genotyping methods developed in the past decade, genotyping still is a rapidly evolving field searching for innovative solutions to fulfill the need for cost effective higher throughput solutions and for ultimately enabling genome wide SNP genotyping.

1.3.1 Monitoring Genetic Susceptibility: a historic point of view

The oldest and most classical techniques for analyzing an individual's genotype, are based on the examination of the "phenotype", which is the expressed genotype, or the observable features of an individual which result from an interaction between the genotype and environmental factors. These methods depend on the administration of test substances and the subsequent collection of their metabolites in the urine. The most well known example, is the association between the *N*-Acetyltransferase phenotype and the risk for urinary bladder cancer (131, 132). This type of cancer is the first human neoplasm for which precisely defined chemicals, namely arylamines, were described as causal factors and for which chronic exposure to various environmental or occupational arylamines was related to a percentage of cancer incidence. These carcinogenic compounds are generally known to undergo metabolic activation and detoxification reactions, the latter mainly by acetylation (133, 134). Several studies contributed to the hypothesis, that the "genetically distinct slow acetylation phenotype" may explain urinary bladder cancer risk in arylamine exposed populations (135). To identify the acetylator phenotype, individuals were administrated orally with sulfamethazine, a free acid which is acetylated in the liver. Free (unacetylated) and total (unacetylated and acetylated) sulfamethazine were quantitated on filter paper discs from both blood and urine samples, and subsequent phenotypic categorization was accomplished by plotting the percentage of acetylated sulfamethazine in blood versus urine (129).

Chapter 1

After the development of Sanger's DNA sequencing (136) and the invention of Mullis' PCR method (130), the first SNP typing methodologies came available. The PCR technique, which is the enzymatic amplification of (specific) DNA sequences based on temperature cycling (130), provided copies of DNA, or so called PCR products, which are accessible for genotyping. PCR-RFLP (PCR Restriction Fragment Length Polymorphism) is a classical approach based on this technique. After PCR, the amplified products are digested with a specific restriction enzyme (endonuclease), which cuts at specific restriction sites in the DNA sequence. This produces *restriction fragments* of different lengths, which can be separated and visualized by gel-electrophoresis. If a SNP occurs within the restriction site, the specific endonuclease will not react and the sequence stays intact, leading to a different pattern of restriction fragments. Several similar approaches have been developed based on the PCR-RFLP method, reviewed by d'Surney et al (137). However, due to a limited number of restriction enzymes, these methods are not well suitable for high throughput purposes.

1.3.2 Genotyping techniques

More recently, the knowledge provided by genotyping studies in the last three decades has been used to develop more high-throughput genotyping techniques. For a better understanding of the methodologies, it is useful to divide the techniques according to their characteristics, namely the allelic discrimination reaction, the assay environment and the detection method (Figure 3). Regarding the method used to discriminate between different alleles, the majority of the SNP genotyping assays can be classified into one of four groups based on molecular mechanism: allele specific hybridization, primer extension, oligonucleotide ligation and invasive cleavage.

The first one, *hybridization*, is based on the fact that two allele specific oligonucleotides (ASO probes) are hybridized to the target sequence containing the SNP. Only fully complementary probe-target hybridizations are stable. An example is presented by molecular beacons, which use a fluorescent dye/quencher system (138, 139). In this case, the centre of the beacon is complementary to one allele of the SNP, while the two ends, which contain the dye on one and the quencher on the other end, are complementary to each other, thus forming a stem-loop structure. The quencher now reacts with the fluorescent dye, quenching its fluorescence by means of fluorescence resonance energy transfer (FRET). If the PCR product is fully complementary to the central sequence, it hybridizes and the two ends of the beacon are separated thereby releasing a fluorescent signal. The second, *primer extension*, is based on the elongation of a primer hybridized to the oligonucleotide of interest, reaching just until the SNP site. If this elongation occurs with a labeled dideoxy-nucleotide, the incorporated nucleotide at the SNP side can be detected using fluorescence (minisequencing or Single Base Extension (140, 141). In *Oligonucleotide ligation* (OLA), Landegren et al (142) used the ability of DNA ligases to covalently join two oligonucleotides when they are hybridized to the sequence of interest. Three oligo's are designed: one common probe is complementary

Figure 3. Schematic overview of the possible combinations for genotyping techniques, based on the allelic discrimination reaction, assay environment and detection methods. Some examples are given and explained in the text: 1. Molecular Beacons, 2. Arrayed Primer Extension (APEX), 3. SNaPSHOT, 4. Oligonucleotide Ligation (OLA), 5. Invader assay, 6. Taqman Assay and 7. Restriction Fragment Length Polymorphism (RFLP).

just until the 3' of SNP position, the other 2 overlap the SNP site at their 3' side, with each probe containing one allele. The two probes compete to anneal to the target sequence just adjacent to the common probe. Only if the probe fully hybridizes, i.e. contains the appropriate nucleotide at the allelic side, the two parts are enzymatically ligated to each other by a DNA ligase. An example of signal detection is by using FRET, with the common probe containing a donor dye at the 5' site and the allele specific probes containing a specific acceptor dye at the 3' site. *Nuclease cleavage* methods are in some cases also based on hybridization, like the Invader[®] (143) or the Taqman[®] (144, 145) assay, however the eventual allelic discrimination is different, i.e. based on the degradation of nucleotide sequences. The invader assay[®] uses 2 oligonucleotides, namely the "invader oligonucleotide", complementary to the 3' side of the SNP, and the "probe", which overlaps the SNP and contains 2 regions: one contains a 3' fragment including the SNP side complementary to the sequence at the 5' side of the SNP, the other is a non binding 5' tail opposite to the 3' side of the SNP. This non-binding tail overlaps the invader in case of a successful hybridization of both oligonucleotides, that is if the SNP is complementary to the allelic base in the probe. This three-dimensional structure is recognized by a so called Flap endonuclease, which releases the 3' arm probe. This probe can subsequently serve as an invader oligonucleotide in a secondary cleavage reaction, increasing the signal. Detection of this signal can be linked to a change in fluorescence, e.g. FRET, or by analysis with Mass Spectrometry. Because of this amplification of the signal, no PCR is required for this assay. In contrast, the Taqman[®] assay (sometimes called 5' exonuclease assay) is conducted during PCR. In this case, two oligonucleotides which differ at the polymorphic sites (Taqman[®] probes) are labeled with a different fluorescent dyes at the 3' side and a quencher molecule at the 5' side of the sequence, quenching the fluorescence by FRET. If the oligonucleotide is complementary to the target sequence containing the SNP, hybridization occurs during the PCR annealing step. In the subsequent extension step, degradation from the 5' site is performed by the Taq polymerase and replaced with a new strand, thereby separating the dye from the quencher which produces a fluorescent signal. In case of unsuccessful hybridization, the new strand is synthesized without fragmentation of the probe.

The assay can be conducted in two different environments, namely *in solution* or on *solid support*. Several formats are developed for the latter, including glass slides, chips or beads. For example, primer extension can be conducted in both environments (example 3, Figure 3). In case of a solution, the oligonucleotide probes, or targets, are free present in a reaction solution containing also the fluorescent labeled di-deoxynucleotides and the PCR products. The primer extension reaction is conducted in this solution. In contrast, if a solid support is used, the targets are spotted on, e.g. a glass slide, and a solution containing the PCR products and labeled nucleotides is added. In this case, hybridization and incorporation of the di-deoxynucleotides takes place on the glass slide.

Finally, numerous detection methods are available for detection of the products obtained from the different allele discrimination reactions. The major ones are methods based on gel separation, microarray based analysis, mass spectrometry or fluorescent plate readers, some examples of these were described above and shown in Figure 3. Analyses based on *gel electrophoresis* are generally considered as the “golden standard”. In this case, automated high throughput can be achieved with the use of multi-capillary electrophoresis, which are able to operate unsupervised 24 hours.

Still, development of genotyping analyses, detection methods and platforms together with the optimization of the current techniques is continuously going on, eventually leading to whole genome SNP typing. One major tool in working towards the latter is provided by the publication of the human haplotype map (HapMap, www.hapmap.org) (146). The HapMap project aims to reduce the number of common SNPs required for the examination of the entire genome. This is achieved by using the fact that genetic variants that are located near to each other tend to be inherited together. For example, individuals carrying a certain SNP at one particular location of a chromosome, will also carry identical SNPs at other locations of that chromosome. These regions of SNPs that are linked to each other are called “haplotypes”. The HapMap project has identified approximately 500.000 “tag SNPs” out of the 10 million SNPs that exist in the human genome, which are able to identify these haplotypes (147). Thus, haplotype blocks may provide useful information to reduce the number of SNPs that need to be genotyped for whole genome scans. However, because the sizes of these haplotype blocks are relatively small, still a large number of SNPs needs to be assessed and appropriate strategies are still under development.

1.3.3 Genotyping techniques: Selection issues

The majority of the current high-throughput genotyping methods combines one of the above described available methods for allelic discrimination, with a certain detection method using a suitable assay format, as is shown in Figure 3. The currently available high throughput methods have been reviewed in the literature (148), discussing issues as the level of throughput, capability of multiplexing, sensitivity, reproducibility and accuracy of the techniques. Other important aspects considering the available assays are the flexibility of the technology, the time consumption of the analysis and the cost regarding both the required equipment and the cost per genotype determination. Comparing the advantages and disadvantages of all these aspects between the available methodologies, to define the most suitable technique for specific applications is difficult, as there is no single general protocol that fits all research needs. Often, crucial aspects are the cost, the level of throughput and the possibility for multiplexing. Multiplexing can reduce the cost per genotype, by the simultaneous analysis of multiple genotypes per conducted reaction. Moreover, multiplexing can increase the throughput of the reaction for the same reason. Multiplex capability and throughput largely depend on the assay

environment. High levels of multiplexing, implying the number of SNPs per analyzed sample, can be achieved by reactions on solid supports, for example by using microarrays. However, the throughput of these assays, in terms of the number of analyzed samples, is often low. A much higher level of throughput is feasible with assays conducted in solution. For example, primer extension reactions in solution can dramatically increase the level of samples analyzed per experiment, although the multiplex capability of these assays is often limited.

Crucial aspects for biomonitoring studies are in general the level of throughput and multiplexing capability, but also the flexibility of the assay. In this case, the latter refers to the possibility of selecting the SNPs to be analyzed which are often involved in particular pathways of interest, for example pathways related to environmental carcinogenesis. Because several pathways are often involved, multiple SNPs need to be assessed and at least a medium level of multiplexing is required. However, the major drawback in biomonitoring studies is often the number of samples that have to be analyzed, which requires a high throughput level.

1.4 Aim and outline of the thesis

Genetic susceptibility attracts major attention in molecular epidemiology studies, as it may influence every intermediate biomarker within the continuum of external exposure to a hazardous environmental compound, to development of disease. Where initially emphasis lied on the investigation of one single polymorphism in relation to a specific exposure, it has now become evident that gene-environment interactions are far more complex, involving genetic variations in multiple genes and pathways, which may furthermore interact with other susceptibility factors such as age and the extent of carcinogen exposure (18). Therefore, in order to assess the impact of the environment on human health, the need for new biomarkers of susceptibility, which enable investigating the role of multiple SNPs of interest in large populations, is urgent. Furthermore, it is important to incorporate this knowledge into the risk assessment process to sustain the policy debate about inter-individual differences in susceptibility regarding hazardous environmental exposures.

It is hypothesized that, instead of looking at single SNPs in relation to a single marker, the interaction between multiple genes encoding for susceptibility to environmental exposure, with biomarkers of exposure and early effect, explains inter-individual differences in responses to environmental exposure and may thus provide new insights on environmental health risk assessment.

This thesis therefore focuses on the development, validation and application of a novel multiplex genotyping technique, applicable as a biomarker of genetic susceptibility in biomonitoring studies. This technique provides the opportunity to simultaneously assess multiple polymorphisms in genes of interest, i.e. the ability to select functional polymorphisms in pathways relevant for environmental carcinogenesis, in large

populations. Development and validation of this genotyping technique is described in **Chapter 2**.

Cigarette smoking is generally considered a model for exposure to environmental carcinogens. To assess the susceptibility of smokers for developing cigarette smoke-related health effects by application of the multiplex genotyping technique as developed in Chapter 2, two main smoking-related biomarkers for internal dose/ early effect were evaluated, namely bulky DNA adducts (**Chapter 3**) and 4-amino-biphenyl-hemoglobin adducts (**Chapter 4**). In both studies, it was hypothesized that genetic polymorphisms in genes involved in carcinogenesis related-pathways might underly, in combination with other susceptibility factors, the inter-individual variations in the exposure-effect relationship between cigarette smoking and the specific biomarker. Moreover, in Chapter 3, a statistical analysis is developed and validated which is applied in the subsequent studies exploring similar research questions.

Application of the developed methodology in a large biomonitoring study, which has been conducted in Belgium in assignment of the Flemish Government, is presented in **Chapter 5**. The aim of this so-called Flemish Environment and Health Study (FLEHS) was to obtain insights in the relationship between exposures to environmental pollutants and human health risk in Flanders. Inter-individual differences in response of certain biomarkers of effect to these environmental exposures were investigated by determining the impact of SNPs, selected in genes involved in hypothetically relevant pathways, on the relationship between the measured exposure- and effect markers of the FLEHS.

Finally, to complete the circle to Darwin's theory, two subpopulations of the FLEHS biomonitoring study, namely adolescents and adults, were analyzed to test the hypothesis that, if an unfavorable combination of an at-risk genotype and environmental exposures is indeed associated with disease development, the prevalence of an at-risk genotype decreases in older- compared to younger populations (**Chapter 6**). Conducted studies are summarized and discussed in **Chapter 7**.

References

1. Venter, J.C., Adams, M.D., Myers, E.W., et al. (2001) The sequence of the human genome. *Science*, 291, 1304-51.
2. Sachidanandam, R., Weissman, D., Schmidt, S.C., et al. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 409, 928-33.
3. Watson, J.D. and Crick, F.H. (1953) The structure of DNA. *Cold Spring Harb Symp Quant Biol*, 18, 123-31.
4. Lander, E.S., Linton, L.M., Birren, B., et al. (2001) Initial sequencing and analysis of the human genome. *Nature*, 409, 860-921.
5. Venitt, S. (1994) Mechanisms of carcinogenesis and individual susceptibility to cancer. *Clin Chem*, 40, 1421-5.
6. Weinstein, I.B. (1995) The contributions of molecular biology to cancer epidemiology. *Ann N Y Acad Sci*, 768, 30-40.
7. Boyle, P. (2006) The globalisation of cancer. *Lancet*, 368, 629-30.
8. Perera, F.P. (1996) Molecular epidemiology: insights into cancer susceptibility, risk assessment, and prevention. *J Natl Cancer Inst*, 88, 496-509.
9. Harris, C.C. (1989) Interindividual variation among humans in carcinogen metabolism, DNA adduct formation and DNA repair. *Carcinogenesis*, 10, 1563-6.
10. King, R.A., Rotter, J. I., and Motulsky, A. G. (2002) *The Genetic Basis of Common Diseases*. Oxford University Press Oxford, UK.
11. Alberts, B., Johnson, A., Lewis, J., Raff, R., Roberts, K., and Walker, P. (2002) *Molecular Biology of the Cell*. Garland Publishing, New York.
12. Venkitaraman, A.R. (2002) Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*, 108, 171-82.
13. Miki, Y., Swensen, J., Shattuck-Eidens, D., et al. (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, 266, 66-71.
14. Easton, D., Ford, D. and Peto, J. (1993) Inherited susceptibility to breast cancer. *Cancer Surv*, 18, 95-113.
15. Orphanides, G. and Kimber, I. (2003) Toxicogenetics: applications and opportunities. *Toxicol Sci*, 75, 1-6.
16. Olden, K. and Guthrie, J. (2001) Genomics: implications for toxicology. *Mutat Res*, 473, 3-10.
17. Vineis, P. (1997) Molecular epidemiology: low-dose carcinogens and genetic susceptibility. *Int J Cancer*, 71, 1-3.
18. Perera, F.P. (1997) Environment and cancer: who are susceptible? *Science*, 278, 1068-73.
19. Kelada, S.N., Eaton, D.L., Wang, S.S., et al. (2003) The role of genetic polymorphisms in environmental health. *Environ Health Perspect*, 111, 1055-64.
20. Xu, X., Kelsey, K.T., Wiencke, J.K., et al. (1996) Cytochrome P450 CYP1A1 MspI polymorphism and lung cancer susceptibility. *Cancer Epidemiol Biomarkers Prev*, 5, 687-92.
21. Nakachi, K., Imai, K., Hayashi, S., et al. (1991) Genetic susceptibility to squamous cell carcinoma of the lung in relation to cigarette smoking dose. *Cancer Res*, 51, 5177-80.

22. Kawajiri, K., Nakachi, K., Imai, K., et al. (1990) Identification of genetically high risk individuals to lung cancer by DNA polymorphisms of the cytochrome P450IA1 gene. *FEBS Lett*, 263, 131-3.
23. Kawajiri, K., Eguchi, H., Nakachi, K., et al. (1996) Association of CYP1A1 germ line polymorphisms with mutations of the p53 gene in lung cancer. *Cancer Res*, 56, 72-6.
24. Watanabe, J., Shimada, T., Gillam, E.M., et al. (2000) Association of CYP1B1 genetic polymorphism with incidence to breast and lung cancer. *Pharmacogenetics*, 10, 25-33.
25. Taningher, M., Malacarne, D., Izzotti, A., et al. (1999) Drug metabolism polymorphisms as modulators of cancer susceptibility. *Mutat Res*, 436, 227-61.
26. Spivack, S.D., Fasco, M.J., Walker, V.E., et al. (1997) The molecular epidemiology of lung cancer. *Crit Rev Toxicol*, 27, 319-65.
27. Smith, G.B., Harper, P.A., Wong, J.M., et al. (2001) Human lung microsomal cytochrome P450IA1 (CYP1A1) activities: impact of smoking status and CYP1A1, aryl hydrocarbon receptor, and glutathione S-transferase M1 genetic polymorphisms. *Cancer Epidemiol Biomarkers Prev*, 10, 839-53.
28. Rojas, M., Cascorbi, I., Alexandrov, K., et al. (2000) Modulation of benzo[a]pyrene diol-epoxide-DNA adduct levels in human white blood cells by CYP1A1, GSTM1 and GSTT1 polymorphism. *Carcinogenesis*, 21, 35-41.
29. Crofts, F., Taioli, E., Trachman, J., et al. (1994) Functional significance of different human CYP1A1 genotypes. *Carcinogenesis*, 15, 2961-3.
30. Roos, P.H. and Bolt, H.M. (2005) Cytochrome P450 interactions in human cancers: new aspects considering CYP1B1. *Expert Opin Drug Metab Toxicol*, 1, 187-202.
31. Hecht, S.S. (1999) Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst*, 91, 1194-210.
32. Bartsch, H., Nair, U., Risch, A., et al. (2000) Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer Epidemiol Biomarkers Prev*, 9, 3-28.
33. Nebert, D.W., Petersen, D.D. and Puga, A. (1991) Human AH locus polymorphism and cancer: inducibility of CYP1A1 and other genes by combustion products and dioxin. *Pharmacogenetics*, 1, 68-78.
34. Garte, S.J., Trachman, J., Crofts, F., et al. (1996) Distribution of composite CYP1A1 genotypes in Africans, African-Americans and Caucasians. *Hum Hered*, 46, 121-7.
35. Garte, S. (1998) The role of ethnicity in cancer susceptibility gene polymorphisms: the example of CYP1A1. *Carcinogenesis*, 19, 1329-32.
36. Hayashi, S., Watanabe, J., Nakachi, K., et al. (1991) Genetic linkage of lung cancer-associated MspI polymorphisms with amino acid replacement in the heme binding region of the human cytochrome P450IA1 gene. *J Biochem (Tokyo)*, 110, 407-11.
37. Kawajiri, K., Nakachi, K., Imai, K., et al. (1993) The CYP1A1 gene and cancer susceptibility. *Crit Rev Oncol Hematol*, 14, 77-87.
38. Kawajiri, K., Nakachi, K., Imai, K., et al. (1990) Individual differences in lung cancer susceptibility in relation to polymorphisms of P-450IA1 gene and cigarette dose. *Princess Takamatsu Symp*, 21, 55-61.
39. Deakin, M., Elder, J., Hendrickse, C., et al. (1996) Glutathione S-transferase GSTT1 genotypes and susceptibility to cancer: studies of interactions with GSTM1 in lung, oral, gastric and colorectal cancers. *Carcinogenesis*, 17, 881-4.
40. Strange, R.C. and Fryer, A.A. (1999) The glutathione S-transferases: influence of polymorphism on cancer susceptibility. *IARC Sci Publ*, 231-49.

Chapter 1

41. Seidegard, J., Pero, R.W., Markowitz, M.M., et al. (1990) Isoenzyme(s) of glutathione transferase (class Mu) as a marker for the susceptibility to lung cancer: a follow up study. *Carcinogenesis*, 11, 33-6.
42. Nazar-Stewart, V., Motulsky, A.G., Eaton, D.L., et al. (1993) The glutathione S-transferase mu polymorphism as a marker for susceptibility to lung carcinoma. *Cancer Res*, 53, 2313-8.
43. McWilliams, J.E., Sanderson, B.J., Harris, E.L., et al. (1995) Glutathione S-transferase M1 (GSTM1) deficiency and lung cancer risk. *Cancer Epidemiol Biomarkers Prev*, 4, 589-94.
44. Bell, D.A., Taylor, J.A., Paulson, D.F., et al. (1993) Genetic risk and carcinogen exposure: a common inherited defect of the carcinogen-metabolism gene glutathione S-transferase M1 (GSTM1) that increases susceptibility to bladder cancer. *J Natl Cancer Inst*, 85, 1159-64.
45. Evans, D.A. (1989) N-acetyltransferase. *Pharmacol Ther*, 42, 157-234.
46. Yu, M.C., Skipper, P.L., Taghizadeh, K., et al. (1994) Acetylator phenotype, aminobiphenyl-hemoglobin adduct levels, and bladder cancer risk in white, black, and Asian men in Los Angeles, California. *J Natl Cancer Inst*, 86, 712-6.
47. Bell, D.A., Taylor, J.A., Butler, M.A., et al. (1993) Genotype/phenotype discordance for human arylamine N-acetyltransferase (NAT2) reveals a new slow-acetylator allele common in African-Americans. *Carcinogenesis*, 14, 1689-92.
48. Cartwright, R.A., Glashan, R.W., Rogers, H.J., et al. (1982) Role of N-acetyltransferase phenotypes in bladder carcinogenesis: a pharmacogenetic epidemiological approach to bladder cancer. *Lancet*, 2, 842-5.
49. Vineis, P., Bartsch, H., Caporaso, N., et al. (1994) Genetically based N-acetyltransferase metabolic polymorphism and low-level environmental exposure to carcinogens. *Nature*, 369, 154-6.
50. Landi, M.T., Zocchetti, C., Bernucci, I., et al. (1996) Cytochrome P4501A2: enzyme induction and genetic control in determining 4-aminobiphenyl-hemoglobin adduct levels. *Cancer Epidemiol Biomarkers Prev*, 5, 693-8.
51. Lilla, C., Verla-Tebit, E., Risch, A., et al. (2006) Effect of NAT1 and NAT2 genetic polymorphisms on colorectal cancer risk associated with exposure to tobacco smoke and meat consumption. *Cancer Epidemiol Biomarkers Prev*, 15, 99-107.
52. Bartsch, H. and Hietanen, E. (1996) The role of individual susceptibility in cancer burden related to environmental exposure. *Environ Health Perspect*, 104 Suppl 3, 569-77.
53. Pavanello, S. and Clonfero, E. (2000) Biological indicators of genotoxic risk and metabolic polymorphisms. *Mutat Res*, 463, 285-308.
54. Nebert, D.W., McKinnon, R.A. and Puga, A. (1996) Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer. *DNA Cell Biol*, 15, 273-80.
55. Hulla, J.E., Miller, M.S., Taylor, J.A., et al. (1999) Symposium overview: the role of genetic polymorphism and repair deficiencies in environmental disease. *Toxicol Sci*, 47, 135-43.
56. Nebert, D.W., Ingelman-Sundberg, M. and Daly, A.K. (1999) Genetic epidemiology of environmental toxicity and cancer susceptibility: human allelic polymorphisms in drug-metabolizing enzyme genes, their functional importance, and nomenclature issues. *Drug Metab Rev*, 31, 467-87.
57. Ames, B.N. (1989) Endogenous DNA damage as related to cancer and aging. *Mutat Res*, 214, 41-6.
58. Wood, R.D., Mitchell, M., Sgouros, J., et al. (2001) Human DNA repair genes. *Science*, 291, 1284-9.
59. Bohr, V.A. (1995) DNA repair fine structure and its relations to genomic instability. *Carcinogenesis*, 16, 2885-92.

60. Goode, E.L., Ulrich, C.M. and Potter, J.D. (2002) Polymorphisms in DNA repair genes and associations with cancer risk. *Cancer Epidemiol Biomarkers Prev*, 11, 1513-30.
61. Friedberg, E.C. (1995) Out of the shadows and into the light: the emergence of DNA repair. *Trends Biochem Sci*, 20, 381.
62. Berwick, M. and Vineis, P. (2000) Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. *J Natl Cancer Inst*, 92, 874-97.
63. Moses, R.E. (2001) DNA damage processing defects and disease. *Annu Rev Genomics Hum Genet*, 2, 41-68.
64. Wooster, R., Bignell, G., Lancaster, J., et al. (1995) Identification of the breast cancer susceptibility gene BRCA2. *Nature*, 378, 789-92.
65. Cleaver, J.E. (1968) Defective repair replication of DNA in xeroderma pigmentosum. *Nature*, 218, 652-6.
66. David-Beabes, G.L. and London, S.J. (2001) Genetic polymorphism of XRCC1 and lung cancer risk among African-Americans and Caucasians. *Lung Cancer*, 34, 333-9.
67. Ratnasinghe, D., Yao, S.X., Tangrea, J.A., et al. (2001) Polymorphisms of the DNA repair gene XRCC1 and lung cancer risk. *Cancer Epidemiol Biomarkers Prev*, 10, 119-23.
68. Duell, E.J., Millikan, R.C., Pittman, G.S., et al. (2001) Polymorphisms in the DNA repair gene XRCC1 and breast cancer. *Cancer Epidemiol Biomarkers Prev*, 10, 217-22.
69. Stern, M.C., Umbach, D.M., van Gils, C.H., et al. (2001) DNA repair gene XRCC1 polymorphisms, smoking, and bladder cancer risk. *Cancer Epidemiol Biomarkers Prev*, 10, 125-31.
70. Hou, S.M., Falt, S., Angelini, S., et al. (2002) The XPD variant alleles are associated with increased aromatic DNA adduct level and lung cancer risk. *Carcinogenesis*, 23, 599-603.
71. Zhou, W., Liu, G., Miller, D.P., et al. (2002) Gene-environment interaction for the ERCC2 polymorphisms and cumulative cigarette smoking exposure in lung cancer. *Cancer Res*, 62, 1377-81.
72. Winsey, S.L., Haldar, N.A., Marsh, H.P., et al. (2000) A variant within the DNA repair gene XRCC3 is associated with the development of melanoma skin cancer. *Cancer Res*, 60, 5612-6.
73. Matullo, G., Guarrera, S., Carturan, S., et al. (2001) DNA repair gene polymorphisms, bulky DNA adducts in white blood cells and bladder cancer in a case-control study. *Int J Cancer*, 92, 562-7.
74. Halliwell, B. (1996) Mechanisms involved in the generation of free radicals. *Pathol Biol (Paris)*, 44, 6-13.
75. Collins, A.R. (1999) Oxidative DNA damage, antioxidants, and cancer. *Bioessays*, 21, 238-46.
76. Lu, A.L., Li, X., Gu, Y., et al. (2001) Repair of oxidative DNA damage: mechanisms and functions. *Cell Biochem Biophys*, 35, 141-70.
77. Forsberg, L., de Faire, U. and Morgenstern, R. (2001) Oxidative stress, human genetic variation, and disease. *Arch Biochem Biophys*, 389, 84-93.
78. Caporaso, N. (2003) The molecular epidemiology of oxidative damage to DNA and cancer. *J Natl Cancer Inst*, 95, 1263-5.
79. Loft, S. and Poulsen, H.E. (1996) Cancer risk and oxidative DNA damage in man. *J Mol Med*, 74, 297-312.
80. Ratnasinghe, D., Tangrea, J.A., Andersen, M.R., et al. (2000) Glutathione peroxidase codon 198 polymorphism variant increases lung cancer risk. *Cancer Res*, 60, 6381-3.

81. Wang, L.I., Miller, D.P., Sai, Y., et al. (2001) Manganese superoxide dismutase alanine-to-valine polymorphism at codon 16 and lung cancer risk. *J Natl Cancer Inst*, 93, 1818-21.
82. Ahn, J., Gammon, M.D., Santella, R.M., et al. (2005) Associations between breast cancer risk and the catalase genotype, fruit and vegetable consumption, and supplement use. *Am J Epidemiol*, 162, 943-52.
83. Le Marchand, L., Donlon, T., Lum-Jones, A., et al. (2002) Association of the hOGG1 Ser326Cys polymorphism with lung cancer risk. *Cancer Epidemiol Biomarkers Prev*, 11, 409-12.
84. Xu, J., Zheng, S.L., Turner, A., et al. (2002) Associations between hOGG1 sequence variants and prostate cancer susceptibility. *Cancer Res*, 62, 2253-7.
85. Xing, D.Y., Tan, W., Song, N., et al. (2001) Ser326Cys polymorphism in hOGG1 gene and risk of esophageal cancer in a Chinese population. *Int J Cancer*, 95, 140-3.
86. Perera, F.P. and Weinstein, I.B. (1982) Molecular epidemiology and carcinogen-DNA adduct detection: new approaches to studies of human cancer causation. *J Chronic Dis*, 35, 581-600.
87. Boffetta, P. (2000) Molecular epidemiology. *J Intern Med*, 248, 447-54.
88. Merlo, D.F., Sormani, M.P. and Bruzzi, P. (2006) Molecular epidemiology: New rules for new tools? *Mutat Res*.
89. Hecht, S.S. (2002) Human urinary carcinogen metabolites: biomarkers for investigating tobacco and cancer. *Carcinogenesis*, 23, 907-22.
90. Saieva, C., Tumino, R., Masala, G., et al. (2003) Urinary 1-hydroxypyrene and t,t-muconic acid as biomarkers of exposure to environmental pollutants in two areas in Italy (EPIC-Florence and Ragusa). *Tumori*, 89, 679-86.
91. Jongeneelen, F.J., Anzion, R.B., Leijdekkers, C.M., et al. (1985) 1-hydroxypyrene in human urine after exposure to coal tar and a coal tar derived product. *Int Arch Occup Environ Health*, 57, 47-55.
92. Van Rooij, J.G., Veege, M.M., Bodelier-Bade, M.M., et al. (1994) Smoking and dietary intake of polycyclic aromatic hydrocarbons as sources of interindividual variability in the baseline excretion of 1-hydroxypyrene in urine. *Int Arch Occup Environ Health*, 66, 55-65.
93. Cocco, P., Moore, P.S., Ennas, M.G., et al. (2006) Effect of Urban Traffic, Individual Habits, and Genetic Polymorphisms on Background Urinary 1-Hydroxypyrene Excretion. *Ann Epidemiol*.
94. Baird, W.M., Hooven, L.A. and Mahadevan, B. (2005) Carcinogenic polycyclic aromatic hydrocarbon-DNA adducts and mechanism of action. *Environ Mol Mutagen*, 45, 106-14.
95. Hayes, J.D. and Pulford, D.J. (1995) The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit Rev Biochem Mol Biol*, 30, 445-600.
96. Alexandrie, A.K., Warholm, M., Carstensen, U., et al. (2000) CYP1A1 and GSTM1 polymorphisms affect urinary 1-hydroxypyrene levels after PAH exposure. *Carcinogenesis*, 21, 669-76.
97. Abnet, C.C., Fagundes, R.B., Strickland, P.T., et al. (2006) The influence of genetic polymorphisms in Ahr, CYP1A1, CYP1A2, CYP1B1, GST M1, GST T1, and UGT1A1 on urine 1-hydroxypyrene glucuronide concentrations in healthy subjects from Rio Grande do Sul, Brazil. *Carcinogenesis*.
98. IARC (1982) Some industrial chemicals and dyestuffs. IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans. IARC, Lyon, 29, 93-148.

99. Smith, C.J., Livingston, S.D. and Doolittle, D.J. (1997) An international literature survey of "IARC Group I carcinogens" reported in mainstream cigarette smoke. *Food Chem Toxicol*, 35, 1107-30.
100. Wallace, L. (1996) Environmental exposure to benzene: an update. *Environ Health Perspect*, 104 Suppl 6, 1129-36.
101. Cocco, P., Tocco, M.G., Ibba, A., et al. (2003) trans,trans-Muconic acid excretion in relation to environmental exposure to benzene. *Int Arch Occup Environ Health*, 76, 456-60.
102. Qu, Q., Melikian, A.A., Li, G., et al. (2000) Validation of biomarkers in humans exposed to benzene: urine metabolites. *Am J Ind Med*, 37, 522-31.
103. Scherer, G., Renner, T. and Meger, M. (1998) Analysis and evaluation of trans,trans-muconic acid as a biomarker for benzene exposure. *J Chromatogr B Biomed Sci Appl*, 717, 179-99.
104. Pezzagno, G., Maestri, L. and Fiorentino, M.L. (1999) Trans,trans-muconic acid, a biological indicator to low levels of environmental benzene: some aspects of its specificity. *Am J Ind Med*, 35, 511-8.
105. Rossi, A.M., Guarnieri, C., Rovesti, S., et al. (1999) Genetic polymorphisms influence variability in benzene metabolism in humans. *Pharmacogenetics*, 9, 445-51.
106. Ruppert, T., Scherer, G., Tricker, A.R., et al. (1997) trans,trans-muconic acid as a biomarker of non-occupational environmental exposure to benzene. *Int Arch Occup Environ Health*, 69, 247-51.
107. Bergamaschi, E., Brustolin, A., De Palma, G., et al. (1999) Biomarkers of dose and susceptibility in cyclists exposed to monoaromatic hydrocarbons. *Toxicol Lett*, 108, 241-7.
108. Chanvaivit, S., Navasumrit, P., Hunsonti, P., et al. (2006) Exposure assessment of benzene in Thai workers, DNA-repair capacity and influence of genetic polymorphisms. *Mutat Res*.
109. Dallinga, J.W., Pachen, D.M., Wijnhoven, S.W., et al. (1998) The use of 4-aminobiphenyl hemoglobin adducts and aromatic DNA adducts in lymphocytes of smokers as biomarkers of exposure. *Cancer Epidemiol Biomarkers Prev*, 7, 571-7.
110. Phillips, D.H. (1996) DNA adducts in human tissues: biomarkers of exposure to carcinogens in tobacco smoke. *Environ Health Perspect*, 104 Suppl 3, 453-8.
111. Hemminki, K., Dickey, C., Karlsson, S., et al. (1997) Aromatic DNA adducts in foundry workers in relation to exposure, life style and CYP1A1 and glutathione transferase M1 genotype. *Carcinogenesis*, 18, 345-50.
112. Rojas, M., Alexandrov, K., Cascorbi, I., et al. (1998) High benzo[a]pyrene diol-epoxide DNA adduct levels in lung and blood cells from individuals with combined CYP1A1 MspI/Msp-GSTM1*0/*0 genotypes. *Pharmacogenetics*, 8, 109-18.
113. Godschalk, R.W., Dallinga, J.W., Wikman, H., et al. (2001) Modulation of DNA and protein adducts in smokers by genetic polymorphisms in GSTM1, GSTT1, NAT1 and NAT2. *Pharmacogenetics*, 11, 389-98.
114. Ketelslegers, H.B., Gottschalk, R.W., Godschalk, R.W., et al. (2006) Interindividual variations in DNA adduct levels assessed by analysis of multiple genetic polymorphisms in smokers. *Cancer Epidemiol Biomarkers Prev*, 15, 624-9.
115. Duell, E.J., Wiencke, J.K., Cheng, T.J., et al. (2000) Polymorphisms in the DNA repair genes XRCC1 and ERCC2 and biomarkers of DNA damage in human blood mononuclear cells. *Carcinogenesis*, 21, 965-71.
116. Lunn, R.M., Langlois, R.G., Hsieh, L.L., et al. (1999) XRCC1 polymorphisms: effects on aflatoxin B1-DNA adducts and glycophorin A variant frequency. *Cancer Res*, 59, 2557-61.

Chapter 1

117. Matullo, G., Palli, D., Peluso, M., et al. (2001) XRCC1, XRCC3, XPD gene polymorphisms, smoking and (32)P-DNA adducts in a sample of healthy subjects. *Carcinogenesis*, 22, 1437-45.
118. Vineis, P., Caporaso, N., Tannenbaum, S.R., et al. (1990) Acetylation phenotype, carcinogen-hemoglobin adducts, and cigarette smoking. *Cancer Res*, 50, 3002-4.
119. Kimura, S., Kawabe, M., Ward, J.M., et al. (1999) CYP1A2 is not the primary enzyme responsible for 4-aminobiphenyl-induced hepatocarcinogenesis in mice. *Carcinogenesis*, 20, 1825-30.
120. Tsuneoka, Y., Dalton, T.P., Miller, M.L., et al. (2003) 4-aminobiphenyl-induced liver and urinary bladder DNA adduct formation in Cyp1a2(-/-) and Cyp1a2(+/+) mice. *J Natl Cancer Inst*, 95, 1227-37.
121. Wogan, G.N. (1989) Markers of exposure to carcinogens. *Environ Health Perspect*, 81, 9-17.
122. Heddle, J.A., Hite, M., Kirkhart, B., et al. (1983) The induction of micronuclei as a measure of genotoxicity. A report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutat Res*, 123, 61-118.
123. Forni, A. (1984) Chromosomal aberrations in monitoring exposure to mutagens-carcinogens. *IARC Sci Publ*, 325-37.
124. Carrano, A.V. and Thompson, L.H. (1982) Sister chromatid exchange and gene mutation. *Cytogenet Cell Genet*, 33, 57-61.
125. Albertini, R.J., Nicklas, J.A., Fuscoe, J.C., et al. (1993) In vivo mutations in human blood cells: biomarkers for molecular epidemiology. *Environ Health Perspect*, 99, 135-41.
126. van Leeuwen, D.M., Gottschalk, R.W., van Herwijnen, M.H., et al. (2005) Differential gene expression in human peripheral blood mononuclear cells induced by cigarette smoke and its constituents. *Toxicol Sci*, 86, 200-10.
127. Park, S.Y., Lee, K.H., Kang, D., et al. (2006) Effect of genetic polymorphisms of MnSOD and MPO on the relationship between PAH exposure and oxidative DNA damage. *Mutat Res*, 593, 108-15.
128. Weinberg, R.A. (1989) Oncogenes, antioncogenes, and the molecular bases of multistep carcinogenesis. *Cancer Res*, 49, 3713-21.
129. Weber, W.W. and Brenner, W. (1974) A filter paper method for determining isoniazid acetylase phenotype. *Am J Hum Genet*, 26, 467-73.
130. Mullis, K., Faloona, F., Scharf, S., et al. (1992) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986. *Biotechnology*, 24, 17-27.
131. Case, R.A., Hosker, M.E., Mc, D.D., et al. (1954) Tumours of the urinary bladder in workmen engaged in the manufacture and use of certain dyestuff intermediates in the British chemical industry. I. The role of aniline, benzidine, alpha-naphthylamine, and beta-naphthylamine. *Br J Ind Med*, 11, 75-104.
132. Case, R.A. and Hosker, M.E. (1954) Tumour of the urinary bladder as an occupational disease in the rubber industry in England and Wales. *Br J Prev Soc Med*, 8, 39-50.
133. Miller, J.A. and Miller, E.C. (1969) The metabolic activation of carcinogenic aromatic amines and amides. *Prog Exp Tumor Res*, 11, 273-301.
134. Weisburger, J.H. and Weisburger, E.K. (1973) Biochemical formation and pharmacological, toxicological, and pathological properties of hydroxylamines and hydroxamic acids. *Pharmacol Rev*, 25, 1-66.
135. Lower, G.M., Jr., Nilsson, T., Nelson, C.E., et al. (1979) N-acetyltransferase phenotype and risk in urinary bladder cancer: approaches in molecular epidemiology. Preliminary results in Sweden and Denmark. *Environ Health Perspect*, 29, 71-9.

136. Sanger, F., Nicklen, S. and Coulson, A.R. (1992) DNA sequencing with chain-terminating inhibitors. 1977. *Biotechnology*, 24, 104-8.
137. D'Surney, S.J., Shugart, L.R. and Theodorakis, C.W. (2001) Genetic markers and genotyping methodologies: an overview. *Ecotoxicology*, 10, 201-4.
138. Tyagi, S. and Kramer, F.R. (1996) Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol*, 14, 303-8.
139. Tyagi, S., Bratu, D.P. and Kramer, F.R. (1998) Multicolor molecular beacons for allele discrimination. *Nat Biotechnol*, 16, 49-53.
140. Syvanen, A.C., Aalto-Setälä, K., Harju, L., et al. (1990) A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics*, 8, 684-92.
141. Sokolov, B.P. (1990) Primer extension technique for the detection of single nucleotide in genomic DNA. *Nucleic Acids Res*, 18, 3671.
142. Landegren, U., Kaiser, R., Sanders, J., et al. (1988) A ligase-mediated gene detection technique. *Science*, 241, 1077-80.
143. Mein, C.A., Barratt, B.J., Dunn, M.G., et al. (2000) Evaluation of single nucleotide polymorphism typing with invader on PCR amplicons and its automation. *Genome Res*, 10, 330-43.
144. Holland, P.M., Abramson, R.D., Watson, R., et al. (1991) Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci U S A*, 88, 7276-80.
145. Kalinina, O., Lebedeva, I., Brown, J., et al. (1997) Nanoliter scale PCR with TaqMan detection. *Nucleic Acids Res*, 25, 1999-2004.
146. (2003) The International HapMap Project. *Nature*, 426, 789-96.
147. Gabriel, S.B., Schaffner, S.F., Nguyen, H., et al. (2002) The structure of haplotype blocks in the human genome. *Science*, 296, 2225-9.
148. Tsuchihashi, Z. and Dracopoli, N.C. (2002) Progress in high throughput SNP genotyping methods. *Pharmacogenomics J*, 2, 103-10.

Chapter 2

Simultaneous genotyping of nine polymorphisms in xenobiotic metabolizing enzymes using multiplex PCR amplification and Single Base Extension

Ad M. Knaapen¹, Hans B. Ketelslegers¹, Ralph W.H. Gottschalk¹, Rob G.J.H. Janssen², Aimee D.C. Paulussen², Hubert J.M. Smeets², Roger W.L. Godschalk¹, Frederik J. van Schooten¹, Jos C.S. Kleinjans and Joost H.M. van Delft

¹Department of Health Risk Analysis and Toxicology, and ²Department of Population Genetics, Genomics and Bioinformatics, Maastricht University, Maastricht, the Netherlands

Clinical Chemistry 2004;50(9):1664-8

Abstract

There is a body of evidence suggesting that interindividual differences in susceptibility towards xenobiotic exposures might at least partly be explained by variations (polymorphisms) in genes that code for xenobiotic-metabolizing enzymes. In order to assess the impact of genetic variation on susceptibility towards xenobiotics, methods for high throughput multiplexed genotyping are urgently needed.

In the present study we developed an assay for simultaneous genotyping of nine polymorphisms in five different genes involved in metabolism of xenobiotic compounds: *CYP1A2* (*1C, *1F), *GSTM1* (null), *GSTP1* (*2, *3), *GSTT1* (null), and *NAT2* (*5, *6, *7). The assay involves pooling of two multiplex PCR products, followed by simultaneous genotyping of nine polymorphisms in one single tube using primer extension and automated capillary gel electrophoresis.

The method is able to genotype single nucleotide polymorphisms, as well as gene-deletions (*GSTM1*-null, *GSTT1*-null). In order to test the validity and applicability of the method, genomic DNA from 67 healthy Caucasians was assayed. Accurate and complete genotypes were obtained from all samples, also upon repeated measurements. The assay was validated by comparison of the obtained genotypes with conventional methods such as restriction fragment length polymorphism-analysis and direct sequencing, which showed 100% accordance.

We have developed a thoroughly validated, robust and rapid assay for simultaneous genotyping of nine polymorphisms in xenobiotic metabolizing enzymes. The method is well suited for large-scale molecular epidemiological studies that aim to identify health risk by studying the relation between xenobiotic dose, multiple interacting susceptibility genes and phenotypical parameters like protein/DNA-adduct levels.

Introduction

In daily life humans are constantly exposed to xenobiotic compounds such as drugs or chemical carcinogens which are abundantly present in for instance polluted air, food and tobacco smoke. It is generally accepted that many of these compounds exert their toxic activity only after bioactivation into reactive metabolites that for instance covalently bind to the DNA (1-3). In the past 20 years, it has become clear that this toxic effect is the integrated result of both dose and the combined action of an array of enzymes that is involved in metabolic activation and detoxification routes (4). These enzymatic processes play a crucial role in for instance chemical carcinogenesis, which is illustrated by the observed large interindividual variation in response to a given carcinogenic exposure (5).

Currently, there is a body of evidence showing that this variation in susceptibility of individuals towards health effects of xenobiotics can at least partly be explained by polymorphisms in genes that code for metabolic enzymes. For instance, till now the relationship between cancer risk and genetic polymorphisms in specific xenobiotic-metabolizing enzymes, such as cytochrome P450 enzymes (CYP450), *N*-acetyltransferases (e.g. NAT2) and glutathione *S*-transferases (e.g. GSTM, GSTT, GSTP) has received widespread attention (e.g. (6-9)). However, conflicting results have been reported, which is probably caused by low number of cases, and hence an insufficient power of many of these studies (10). Moreover, the majority of molecular epidemiology studies investigating the implication of genetic variation on health risk modification has merely focussed on single polymorphisms. Regarding the amount of genes that are implicated in metabolism of xenobiotics, and the large number of polymorphisms present in the human genome (11), these approaches will fail to fully determine the role of genetic variation in the individual susceptibility towards xenobiotic exposures. As such, these observations underline the need for studies that investigate multiple susceptibility genes as well as their mutual interactions in larger populations. It is however obvious, that conventional methods such as restriction fragment length polymorphism (RFLP) analysis and DNA sequencing, are either time-consuming and/or too expensive for large-scale analysis. Therefore, new procedures that allow for high-throughput, low cost and accurate genotyping of polymorphisms are urgently needed (12,13).

In the present study we aimed to develop a cost-effective and rapid method for simultaneous genotyping of 9 polymorphisms (either single nucleotide polymorphisms, SNP's, or gene-deletions) in 5 key enzyme genes involved in metabolism of xenobiotics: CYP1A2, GSTM1, GSTP1, GSTT1, and NAT2. These polymorphisms were specifically selected based on their functional effect on enzyme activity and association with health effects. The method involves multiplex PCR-based DNA template amplification followed by single base primer extension (SBE) and subsequent analysis using capillary gel electrophoresis (14-16). Here we report on development, validation and

application of the assay. Moreover, the potential for its use in future large scale molecular epidemiology studies on the relation between xenobiotic exposures, genetic susceptibility and health risk identification is discussed.

Materials and Methods

Primer design and multiplex PCR amplification

All polymorphisms included in the assay were selected based on known effects on enzyme activity and association with cancer development (Table 1).

Table 1. Overview of the polymorphisms included in the study. Inclusion criteria for selection of the polymorphisms were: a) known effects on enzyme activity, and b) association with cancer.

Gene	Polymorphism	Genotype	Phenotype*	Association with cancer*
CYP1A2	*1C	-3858G>A	Decreased activity (17)	Colorectal cancer (18)
	*1F	-164A>C	Decreased activity (19)	
GSTM1	Deletion	Gene deletion	Loss of activity (20)	Lung, colorectal, bladder cancer(7)
GSTP1	*2	1404A>G	Reduced activity (21)	Bladder, testicular, prostate (6) Lung (22)
	*3	2294C>T	Reduced activity (21)	
GSTT1	Deletion	Gene deletion	Loss of activity (20)	Gastro-intestinal, lung, kidney, liver (9)
NAT2	*5	341T>C	Slow acetylation (20)	NAT2-related acetylation differences associated with: bladder, breast, liver, colon and lung cancers. (8)
	*6	590G>A	Slow acetylation, less stable protein (20)	
	*7	857G>A	Slow acetylation (20)	

*Only a selection of references is used. The table should not be considered as a thorough review of the existing literature.

PCR primers were designed using the internet Primer 3 software (www.broad.mit.edu/cgi-bin/primer/primer3_www.cgi) and the NetPrimer software (www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html). Primers were all designed to have a melting temperature of 60±2°C, with a primer length of 19-23 bp. Primers (desalted) were obtained from Qiagen Operon (Hilden, Germany) and were dissolved in PCR water (Acros, Den Bosch, The Netherlands) at a concentration of 140-174 µM. Initially, all fragments were amplified using simplex PCR reactions to evaluate the specificity of the primer pairs. Moreover, each simplex PCR was run at various

Multiplex genotyping of carcinogen metabolizing enzymes

annealing temperatures to optimise and synchronise the thermal cycling reactions for multiplexing PCR purposes.

For the 7-plex PCR (Table 2) a 50 µl reaction mixture was prepared containing 1x PCR buffer (Invitrogen, Breda, The Netherlands), 0.2 mM dNTP's (Invitrogen), 0.5 mM MgCl₂ (Invitrogen), 1.25 U Platinum[®] Taq-Polymerase (Invitrogen), and 200 ng template DNA. The final concentrations of the primer pairs in this reaction mixture were 0.22 µM for respectively *GSTP1**3, *GSTT1* (null), *NAT2**6, *NAT2**7; 0.45 µM for respectively *CYP1A2**1F and *NAT2**5; and 0.16 µM for *GSTP1**2. PCR was conducted on a Tgradient 96 well Thermal cycler (Biometra) using the following program: denaturation at 94°C (3 min), 30 cycles of 94°C (30 sec), 56°C (30 sec) and 72°C (30 sec), and a final extension at 72°C (5 min).

For the 2-plex PCR (Table 2) a 50 µl reaction mixture was prepared as described above. The final concentration of all primers in this reaction mixture was 0.1 µM. The following PCR program was used: denaturation at 95°C for (5 min), 40 cycles of 95°C (30 sec), 59°C (60 sec), and 72°C (1min 30 sec), followed by a final extension at 72°C (5 min). Negative controls in all PCR reactions consisted of the same reaction mixtures without template DNA. In later experiments the 2- and 7-plex PCR reaction volumes were down-scaled to a final volume of 10 µl. Using this approach a total amount of only 80 ng DNA (2x 40) is needed for amplification of the 9 different fragments.

Table 2. Primer sequences for initial PCR amplification.

Allele	Forward primer	Product length (bp)
<i>7-plex PCR</i>		
<i>CYP1A2</i> *1F	Forward 5'-GAGGCTCCTTTCCAGCTCTC-3' Reverse 5'-CTCCCAGCTGGATACCAGA-3'	106
<i>GSTP1</i> *2	Forward 5'-TGGTGGACATGGTGAATGAC-3' Reverse 5'-AGCCCCTTTCTTTGTTTCAGC-3'	123
<i>GSTP1</i> *3	Forward 5'-TGGGAGGGATGAGAGTAGGA-3' Reverse 5'-CAGGGTCTCAAAAGGCTTCA-3'	106
<i>GSTT1</i> *0	Forward 5'-GTAGCCATCACGGAGCTGAT-3' Reverse 5'-GGCAGCATAAGCAGGACTTC-3'	97
<i>NAT2</i> *5	Forward 5'-CAAATACAGCACTGGCATGG-3' Reverse 5'-GGCTGATCCTTCCCAGAAAT-3'	135
<i>NAT2</i> *6	Forward 5'-CCTGCCAAAGAAGAAACACC-3' Reverse 5'-GGGTCTGCAAGGAACAAAAT-3'	143
<i>NAT2</i> *7	Forward 5'-TCCTTGGGGAGAAATCTCGT-3' Reverse 5'-GGGTGATACATACACAAGGGTTT-3'	92
<i>2-plex PCR</i>		
<i>CYP1A2</i> *1C	Forward 5'-AACACATGATCGAGCTATAC-3' Reverse 5'-GTGGTCTCTTCACTGTAAAGTTA-3'	596
<i>GSTM1</i> *0	Forward 5'-CTCCTGATTATGACAGAAGCC-3' Reverse 5'-CTGGATTGTAGCAGATCATGC-3'	648

Chapter 2

Multiplex PCR product purification

After PCR-amplification, contaminating dNTP's and primers were removed using a single step enzymatic digestion with ExoSAP-IT (containing exonuclease I and shrimp alkaline phosphatase, Amersham, Roosendaal, The Netherlands), to avoid participation in the subsequent primer-extension reaction. Before purification, the two multiplexed PCR reactions were pooled using 5 µl of the 7-plex PCR product and 4 µl of the 2-plex PCR product. This pool of 9 amplified alleles was mixed with 4 µl of ExoSAP-IT and incubated at 37°C for 45 min. Enzymes were subsequently inactivated by incubation at 75°C for 15 min. Purified and pooled PCR products were stored at -20°C until genotyping.

Multiplex genotyping using single base extension

Single base extension (SBE) primers were also designed using Primer 3 and Netprimer software. Primers were all designed to have a template-specific part of 20-33 bp, with a T_m of 66-69°C (Table 3).

Table 3. Extension primer sequences.

Polymorphism	Sequence*	Length
CYP1A2*1C_3858	5'-TGGGGCATGACAATTGCTTGAATC-3'	24
GSTP1*2_313	5'-AACTCTGGAGGACCTCCGCTGCAAATAC-3'	23+ 5=28
NAT2*6_590	5'-CCTACCAAAAAATATACTTATTTACGCTTGAACCTC-3'	33+ 3=36
CYP1A2*1F_164	5'-AACTGACTAAACTAGGTGCCACTCAAAGGGTGAGCTCTG TGGGC-3'	23+21=44
GSTM1*0	5'-AACTGACTAAACTAGGTGCCACGCTGGAGAACCAGACCAT GGACAACC-3'	24+24=48
GSTP1*3_341	5'-AACTGACTAAACTAGGTGCCACGTCGTGACCATGGTGGTGT CTGGCAGGAGG-3'	22+30=52
GSTT1*0	5'-AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGGGCAG GTGAACCCACTAGGC-3'	22+34=56
NAT2*5_341	5'-AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTATTTCAC CTTCTCCTGCAGGTGACCA-3'	24+36=60
NAT2*7_857	5'-AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAGCC CTCGTGCCCAAACCTGGTGATG-3'	22+42=64

*Neutral non-binding tails are represented in *italic*. Tails were taken from the 5' site of either 5'-AAC TGA CTA AAC TAG GTG CCA CGT CGT GAA AGT CTG ACA A-3' or 5'-ATG CTC AGA CAC AAT TAG CGC GAC CCT TAA TCC TTA GGT A-3' as described by Lindblad-Toh et al., 2000 (16). In some cases part of these sequences were complementary to the templates, causing elongation of the template-specific part of the extension-primers. These complementary bases were substituted by neutral non-binding bases (underlined) to avoid changes in T_m of the extension-primers.

Multiplex genotyping of carcinogen metabolizing enzymes

For the SNP's included in the gene-set, the extension primers were designed to bind immediately adjacent 5' to the specific SNP. During the cycling reaction (see below) the primers will be extended at their 3'-end with one single dideoxynucleotide triphosphate (ddNTP) that is labelled with a spectral distinct fluorophore, revealing the genotype of the SNP. For *GSTT1* and *GSTM1* (gene deletions), the extension primers were designed to hybridise in the middle of the amplified allele. All primers were designed to anneal to the antisense strand (forward primers), except for the *CYP1A2*1C* SBE-primer which was designed to anneal to the sense strand (reverse primer). To facilitate detection of all 9 polymorphisms in one single run, the total length of the extension primers was adjusted to a distinct size (resp. 24, 28, 36, 44, 48, 52, 56, 60, 64 bp) by addition of a neutral, non-binding sequence to their 5'-site. These non-binding 'tails' were taken from the 5' site of either 5'-AAC TGA CTA AAC TAG GTG CCA CGT CGT GAA AGT CTG ACA A-3' or 5'-ATG CTC AGA CAC AAT TAG CGC GAC CCT TAA TCC TTA GGT A-3' as described by Lindblad-Toh et al., 2000 (16) (Table 3). Primers in which four or more bases at the 3'-site were complementary to another site of the primer were redesigned to avoid artefacts due to fold-back of the primers. Primers (desalted) were purchased at Qiagen Operon. All SBE primers were dissolved in PCR water (Acros) at 100 µM.

Prior to single base extension a pooled SBE-primer premix was made up in water. SNaPshot™ reactions were performed as described by the manufacturer (Applied Biosystems, Nieuwerkerk a/d IJssel, The Netherlands). In brief, 4 µl of the ExoSAP-IT purified PCR product (containing the 9 fragments) was mixed with 3.5 µl SNaPshot multiplex ready reaction mix (containing AmpliTaq® polymerase, PCR buffer and fluorescently-labelled ddNTP's), 1 µl of the pooled SBE primers and 1.5 µl PCR water (perform in dimmed light). The final concentration of the SBE primers was 0.2 µM for *CYP1A2*1C/1F*, *GSTM1*0*, *GSTT1*0*, *GSTP1*2*, and *NAT2*7*; 1 µM for *NAT2*5* and *GSTP1*3*; and 0.04 µM for *NAT2*6*. As a control, each run contained a SBE reaction on a water sample that was included in the foregoing multiplex PCR reactions. Moreover, in each measurement a negative and a positive control (containing 6 distinct primers and control DNA, as delivered by the manufacturer) were run in parallel. The samples were placed in a thermal cycler (T gradient, Biometra), and single base extension was performed using 25 cycles of 96°C (10 sec) and 60°C (30 sec). In order to avoid interference of the unincorporated ddNTPs in subsequent genotype analysis, the samples were incubated at 37°C (1hr) with 1U shrimp alkaline phosphatase (SAP, Amersham) which removes the 5' phosphoryl groups from the ddNTP. The enzyme was deactivated by incubation at 75°C for 15'. Samples were kept at 4°C for up to 24 hrs prior to electrophoresis. For longer storage, samples were frozen at -20°C.

Genescan analysis

Samples (except for positive and negative control) were diluted 5x in water. Then 1 µl of (diluted) samples was mixed with 13 µl deionised formamide and 0.4 µl of Genescan™ 120 LIZ™ size standard. Samples were denatured at 95°C (5 min) and run on an

Chapter 2

ABI Prism® 3100 genetic analyser using a 36-cm capillary array and POP-6 polymer. The standard run module for SNaPshot samples was adapted for the use of POP-6 polymer by increasing the collection time to 1400 seconds. Samples were analysed using Genescan™ Analysis software (Version 3.7). Since each SBE primer was designed to have a distinct size, the relative position of each primer peak as determined by sizing using the internal LIZ size standard indicates the specific polymorphism, whereas the peak colour specifies the genotype.

Validation of the assay

For validation purposes a subset of 67 lymphocyte DNA samples, obtained from healthy Caucasian volunteers (26 males, 41 females, average age of 41 ± 8 years) was used. Informed consent was obtained from all individuals (23). All 67 samples were initially genotyped using the multiplex SBE assay as described above. Validation of the SBE-obtained genotypes was performed using conventional methods in at least 10 samples. *GSTP1**3 genotypes determined by the SBE primer that binds to the antisense strand were compared with genotypes as found upon extension of the sense SBE primer (5'-TGC CTT CAC ATA GTC ATC CTT GCC C-3'). Other genotypes were validated using either gel-electrophoresis, RFLP-analysis or sequencing (see below).

GEL ELECTROPHORESIS.

GSTM1 and *GSTT1* genotypes (n=67) as determined using the SBE assay, were validated using gel electrophoresis (agarose, 2%) as described previously (24).

RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP).

For validation of SBE genotyping of the *NAT2*, *GSTP1**2 and *CYP1A2**1F polymorphisms, restriction fragment length polymorphism (RFLP) analysis was used. All *NAT2**5 subtypes include the functional *341T>C* (*1114T*) missence variation which is the SNP evaluated in the SBE assay. However, since this SNP does not contain a restriction site, an RFLP assay for detection of the silent *481C>T* variation was used. This SNP is in linkage disequilibrium (LD) with the *341T>C* variant for the majority of the existing *NAT2**5 Caucasian subtypes. Further validation of *NAT2**5 *341T>C* was performed using sequencing (see below).

The RFLP assays for the *NAT2**5 (*481C>T*) and *NAT2**7 (*857G>A*) SNPs were performed as previously described (25). In short, for amplification of a 1093 bp fragment the following primer sequences were used: 5'-GGA ACA AAT TGG ACT TGG-3' (Forward) and 5'-TCT AGC ATG AAT CAC TCT GC-3' (Reverse). The PCR product was digested using either 5U of *KpnI* (Roche, Almere, The Netherlands) for *NAT2**5 (*481C>T*), or 5U *BamHI* (Roche) for *NAT2**7 (*857G>A*). Both digestions were performed at 37°C for 3h. Digested products were evaluated using 3% agarose gel electrophoresis.

Multiplex genotyping of carcinogen metabolizing enzymes

*NAT2*6* (590G>A) was determined as described by Gonzalez et al., 1998 (26), although with different primers. Briefly: a 327 bp product was produced using the following primers 5'-GCC TCT AGA ATT AAT TTC TGG G-3' (Forward) and 5'-CTA TAG GTG AGG ATG AAG CC-3' (Reverse). The PCR product was digested using 5U *TaqI* (Roche) (65°C, 3h). Genotypes were evaluated using 3% agarose gel electrophoresis.

PCR amplification for RFLP analysis of *GSTP1*2* and *CYP1A2*1F* genotypes was performed using the same primers and cycling reaction as described for the multiplex PCR amplification (Table 2). For subsequent enzymatic digestion, the PCR products were digested using either 5U *BsmAI* (55°C, 2.5h) (New England Biolabs) or 10U *ApaI* (30°C, 2.5h) (Invitrogen) for *GSTP1*2* and *CYP1A2*1F* respectively. The digested *GSTP1*2* product shows bands at 123 bp for A/A genotypes (wild type), whereas the presence of bands at 53 and 70 bp indicates G/G. A/G genotypes will show all 3 bands. Digestion of *CYP1A2*1F* products yields fragments of 106 bp for A/A, or 49 and 57 bp for C/C. A/C genotypes will resolve bands at 106, 49 and 57 bp.

SEQUENCING.

To evaluate the accuracy of the SBE genotyping of *CYP1A2*1C*, we determined the sequence of this SNP in 10 subjects of our validation population. This was performed using the BigDye® Terminator v3.1 cycle sequencing kit from Applied Biosystems. PCR-amplified template DNA was purified using QIAquick PCR purification kit (Qiagen). Both forward and reverse PCR primers for *CYP1A2* (Table 2) were used as sequencing primers. Sequencing was performed according to the manufacturer's recommendations using the ABI3100 genetic analyser. Sequencing of *NAT2*5* (341T>C) was performed according to the same protocol, using the primer sequences as described above under RFLP methods.

Statistics

The χ^2 -test was used to evaluate whether the observed genotype frequencies in the validation subset were in Hardy-Weinberg equilibrium. The significance level was set at $P=0.05$.

Results

Multiplex PCR

Although several distinct primer pairs for *CYP1A2*1C* and *GSTM1* were tested, we were unable to amplify these 2 specific alleles without causing interference in the multiplex PCR (data not shown). Therefore, we decided to design primer pairs that produced longer PCR fragments for both *GSTM1* and *CYP1A2*1C* (respectively 648 and 596 bp). As a consequence, amplification of these products could not be performed in

the same cycling protocol as used for the other 7 variants (Table 2). Gel electrophoresis of the duplex PCR reaction product yields 2 distinguishable fragments corresponding with 648 and 596 bp, indicating successful simultaneous amplification of DNA templates containing both *GSTM1* and *CYP1A2*1C* (Figure 1a).

Figure 1. Validation of 2-plex and 7-plex PCR reactions. Figure 1a: Gel electrophoresis of 2-plex PCR (*GSTM1**0, *CYP1A2**1C). Figure 1b (left panel) shows a representative picture of the gel electrophoresis of the 7-plex PCR products. Figure 1b (right panel): Gel electrophoresis of simplex PCR reactions, using the 7 respective primer pairs, performed on either the 1000x diluted 7-plex (lanes 1a-7a), or on the 10,000x dilution of the original template concentration (lanes 1b-7b). 1=*CYP1A2**1F (106 bp), 2=*GSTP1**2 (123 bp), 3=*GSTP1**3 (106 bp), 4=*GSTM1**0 (97 bp), 5=*NAT2**5 (135 bp), 6=*NAT2**6 (143 bp), 7=*NAT2**7 (92 bp).

Gel electrophoresis could not be used to assess whether all fragments were successfully amplified in the 7-plex PCR reaction, since these products are all from more or less the same size (92-143 bp). Gel electrophoresis yields a rather vague smear at about 100 bp (Figure 1b, left panel). To demonstrate that all single PCR products are amplified in the 7-plex PCR, a 1000x dilution of the 7-plex PCR product was used as a 'template' for subsequent simplex PCR reactions using the respective primer pairs. Using this approach, 7 sharp bands appeared on a gel, corresponding with the expected fragment lengths. No bands appeared using a 10,000x dilution of original template DNA (yielding the same amount of original template DNA as present in the 1000x dilution of the 7-plex PCR), demonstrating that all fragments were simultaneously amplified in the original 7-plex PCR (Figure 1b, right panel). After PCR amplification, products of the 2-plex and 7-plex reactions were mixed to obtain one single pool of 9 amplified templates for subsequent simultaneous genotyping.

SBE genotyping

In initial experiments, SBE reactions were performed on the respective 9 simplex PCR products to evaluate the specificity of the designed SBE primers (data not shown). Fig-

Multiplex genotyping of carcinogen metabolizing enzymes

Figure 2 shows representative signals obtained after multiplex SBE reactions, in which 9 alleles are simultaneously genotyped in one vial (= one subject). Using this assay we were able to genotype SNP's, as well as (homozygous) gene deletions (*GSTM1*-null, *GSTT1*-null). Upon performance of the SBE reaction on control multiplex PCR products (using water as 'template'), no signals were observed (not shown), which indicates the absence of possible artefacts due to single base-extendable SBE primer-dimers or hair-pins.

Figure 2. SBE genotyping results. Representative output of 2 different subjects is shown. The horizontal axis indicates the length of the SBE products (20-70 bp.), the vertical axis indicates the peak intensity (fluorescence units). The colour of the peak identifies the genotype (black=C; red=T; blue=G; green=A). The SBE primers will resolve at $n+1$ bp. (cf. Table 3), due to extension of the primers with one single base. The polymorphisms are: 1=CYP1A2*1C; 2=GSTP1*2; 3=NAT2*6; 4=CYP1A2*1F; 5=GSTM1; 6=GSTP1*3; 7=GSTT1; 8=NAT2*5; 9=NAT2*7. The lower panel represents a person having 2 Null-alleles ('0/0') for *GSTM1* (no. 5).

Validation outcomes

In order to test the validity and applicability of our genotyping assay we analysed the 9 polymorphisms in genomic DNA obtained from 67 healthy Caucasians. Accurate and complete genotypes were obtained from all samples (Table 4). Furthermore, to test for inter-assay variation, SBE genotyping of 20 different subjects was performed twice, and no differences in genotype-output were found. Although the test population used for validation of the assay is rather small, all variants of the different genotypes were represented. Moreover, the obtained frequencies are in accordance with previous observations in larger study populations (27,28), supporting the validity of the assay. Further calculations showed that distributions of all genotypes in the validation subset were in Hardy-Weinberg equilibrium ($p>0.05$, Table 4), except for *GSTP1**2.

Table 4. Frequencies of genotypes as determined by multiplex SBE in the subset of 67 DNA samples from Caucasians, which was used for validation purposes.

Gene	Polymorphism	Genotype	Frequency (%) according to SBE	Frequency (%) Literature*
<i>CYP1A2</i>	*1C	GG	98.5	96.0
		GA	1.5	3.9
		AA	0.0	0.04
	*1F	AA	47.1	33.0
		AC	44.1	45.1
		CC	8.8	10.9
<i>GSTM1</i>	Gene deletion	AA/A-	44.1	46.9
		-/-	55.9	53.1
<i>GSTP1</i>	*2 [#]	AA	36.8b	51.7
		AG	48.5	44.2
		GG	14.7	4.1
	*3	CC	83.8	86.8
		CT	14.7	12.6
		TT	1.5	0.6
<i>GSTT1</i>	Gene deletion	AA/A-	83.8	80.3
		-/-	16.2	19.7
<i>NAT2</i>	*5	TT	37.3	29.9
		TC	49.3	48.2
		CC	13.4	21.9
	*6	GG	55.9	50.0
		GA	38.2	43.0
		AA	5.9	7.0
	*7	GG	91.2	94.4
		GA	8.8	5.5
		AA	0.0	0.1

*Frequencies in Caucasians. *CYP1A2**1C, *CYP1A2**1F (27), and *GSTM1**0, *GSTP1**2, *GSTP1**3, *GSTT1**0, *NAT2**5, *NAT2**6, *NAT2**7 (28).

[#]Not in Hardy-Weinberg equilibrium ($\chi^2=32.2$, $P<0.05$). All other genotype frequencies of the study population were in Hardy-Weinberg equilibrium ($P>0.05$).

For more specific validation purposes several independent (conventional) methods were applied, including RFLP analysis and direct sequencing. All genotypes were validated in at least 10 different DNA samples from the test population. For each polymorphism, every variant was assayed in these validation studies. As shown in Table 5, there was 100% accordance between the various genotypes as determined by our multiplex SBE assay and the outcomes obtained by the conventional methods. For determination of the *NAT2**5 polymorphism we evaluated the SNP at position 341 since this SNP covers all possible *NAT2**5 subtypes. However, as the 341T>C variation does not contain a restriction site, *NAT2**5 was initially validated using PCR-RFLP determination of the 481C>T variation, which for the majority of the existing *NAT2**5 Caucasian subtypes is in LD with the 341T>C variant. Using this method we observed 96% (64 out of 67 samples) accordance between the SBE outcomes and the RFLP method.

Table 5. *Validation of the SBE method.*

Gene	Polymorphism	Validation method	N*	% similarity with SBE
<i>CYP1A2</i>	*1C	Sequencing	10	100
	*1F	RFLP	20	100
<i>GSTM1</i>	Gene deletion	Gel electrophoresis	67	100
<i>GSTP1</i>	*2	RFLP	20	100
	*3	Reverse SBE primer	20	100
<i>GSTT1</i>	Gene deletion	Gel electrophoresis	67	100
<i>NAT2</i>	*5	RFLP (481C>T)	67	96
		Sequencing	3	100
	*6	RFLP	10	100
	*7	RFLP	10	100

*All taken from the same subset of 67 Caucasian DNA samples.

Direct sequencing of the remaining 3 samples revealed discordant 341T>C and 481C>T genotypes (Table 6). Both the SBE (341T>C) and RFLP (481C>T) obtained genotypes were in 100% accordance with the direct sequencing results (see Table 5). However, since the 341T>C genotype is the responsible variation for the NAT2 dysfunction, this variation determines the final genotype (Table 6). In addition, haplotype analysis of defective NAT2 alleles has proven to be a good predictor of acetylation phenotype. By analyzing variation at position 341, 590 and 857, we expect the phenotype to be correctly predicted in 90-95% of the population (29,30).

Table 6. *Genotyping of NAT2*5. Comparison of 341T>C (used in the SBE-assay), and 481C>T (RFLP-method).*

Sample code	SBE genotype (341T>C)	RFLP genotype (481C>T)	Final NAT2 genotype*
C54	T/C (heterozygous)	C/C (wildtype)	*1/*5
C66	C/C (homozygous mutant)	C/T (heterozygous)	*5/*5
N8	C/C (homozygous mutant)	C/T (heterozygous)	*5/*5

*Final genotype is determined by the 341T>C variation.

Discussion

Methods for simultaneous genotyping of multiple polymorphisms in large populations are needed to elucidate the involvement of genetic variation in the variable susceptibility of subjects towards xenobiotic exposures. This paper presents a thoroughly validated assay for simultaneous genotyping of nine polymorphisms in five different genes cru-

cially involved in metabolism of xenobiotics such as chemical carcinogens. The assay is based on single base extension (SBE) with subsequent automated genotyping using capillary gel-electrophoresis. Parallel genotyping of a subset of DNA samples with several conventional methods (RFLP, direct sequencing, gel-electrophoresis) clearly demonstrated the validity and speed of the SBE assay. Besides, our data show that the SBE method, which involves analysis of the *34/T>C* variation to determine the *NAT2*5* polymorphism, is more accurate than the conventionally used RFLP (*48/C>T*) method.

This SBE assay is specifically developed with a view to forthcoming large-scale (molecular) epidemiological studies that aim to examine the relation between xenobiotic exposures, interacting genetic polymorphisms, and ultimate health effects such as cancer development. Such studies will help to develop and specify individual health risk profiles and susceptibility (23). All selected polymorphisms are known to affect enzyme activities and are for instance associated with chemical carcinogenesis caused by xenobiotics like polycyclic aromatic hydrocarbons (PAH's) or heterocyclic aromatic amines (Table 1). The selected polymorphisms comprise genes involved in both phase I (*CYP1A2*) and phase II (*GSTM1*, *P1*, *T1*, *NAT2*) metabolic pathways. As such, both activating and detoxifying enzymes are included. Although *CYP1A1* is critically involved in the activation of PAH's (1), it is not included in our assay. This is primarily because the effects of *CYP1A1* polymorphisms on enzyme activity, DNA adduct formation and/or cancer development in Caucasians is still controversial (31,32).

Although SBE technology was originally designed to genotype SNP's, we demonstrated that it can also be applied for analysis of gene-deletions, such as *GSTM1-null* and *GSTT1-null*. For SNP's, heterozygous genotypes can be distinguished due to incorporation of differentially labelled ddNTPs. No differences can be made between wild-type (no deletions) and heterozygous gene deletions, since peak height in the SBE output does not quantitatively reflect gene-copy number. However, as *GSTM1* and *GSTT1* enzyme activity is only completely eliminated through homozygous gene deletions, the distinction between heterozygous and wild-type carriers is of less importance. As shown in Figure 2, the two peaks of heterozygous SNP-genotypes do generally not reach the same height. This can possibly be explained by either unequal allelic amplification during PCR or differential efficiencies of incorporation of the ddNTPs for each allele-specific reaction (33).

Earlier studies that have used comparable SBE (SNaPshot) technology to genotype multiple polymorphisms either started from pools of simplex PCR products (34), pooled DNA samples from cases or controls (33), or single long PCR fragments (35). Others have used PCR multiplexing of 5 segments in one single gene (36). In contrast, in order to fully employ the high-throughput capabilities of the SNaPshot technology, we designed a multiplex PCR-based amplification of 9 alleles in 5 different genes, involving two separate multiplex PCR reactions. The PCR reactions can be performed in 10 µl reaction volumes, using a total amount of 80 ng DNA. This multiplex approach significantly reduces total costs, avoids the use of relative large amounts of template

Multiplex genotyping of carcinogen metabolizing enzymes

DNA, and increases the speed of the assay. In comparison, using conventional simplex PCR-based RFLP methods, genotyping of all polymorphisms in our rather small subset of 67 DNA samples would already involve 603 (67 x 9) different PCR and subsequent genotyping reactions. In the presented assay all PCR reactions and subsequent purification steps can be performed manually in 96 well plates, at a speed of about 2 plates per day. Considering a possible throughput of 4 plates per 24 h using the automated capillary sequencer, 384 subjects can be genotyped for 9 polymorphisms (total number of 3456 genotypes) within 3 days. This throughput can even further be increased by the implementation of fully automated, assay-specific genotyping software and by the use of robotic liquid handling stations.

In preliminary experiments, we demonstrated that the SBE assay was able to successfully genotype polymorphisms in DNA obtained from mouth swabs. Additionally, except for the longer fragments (>500 bp, *CYP1A2*1C*, *GSTM1*), the method could also be applied to perform accurate genotyping in highly fragmented DNA, such as isolated from paraffin-embedded tissue (data not shown). These observations demonstrate the applicability of the method for future large-scale surveys in which either low amounts or relatively poor quality DNA is available. Moreover, apart from the fact that the presented method is quick, accurate and cheap, its versatility and flexibility adds another major advantage: the reported protocol can be easily adapted for new sets or combinations of polymorphisms.

In conclusion, we developed a thoroughly validated method for simultaneous genotyping of 9 polymorphisms, based on multiplexing of PCR reactions and subsequent genotyping using SBE technology and capillary gel electrophoresis. Using this method 7 SNP's and 2 gene deletions in 5 different xenobiotic metabolizing enzymes can be accurately assayed in one single tube, starting from template DNA from various sources. In contrast to for instance microarray-based technologies (e.g. arrayed primer extension, APEX) that mainly focus on simultaneous analysis of large number of polymorphisms in small populations (37), our SBE assay is highly applicable for genotyping of a low to medium number of polymorphisms in large populations. Indeed, we showed that our assay is accurate and robust, and in combination with its semi-automated analysis it is able to rapidly genotype large numbers of subjects. Moreover, the assay specifically focuses on polymorphisms that all have known effects on enzyme activity and which are associated with health effects such as cancer development. Therefore, this method is well suited for large scale studies that aim to identify xenobiotic-related health effects by studying polymorphisms in multiple susceptibility genes, their mutual interactions, and their effect on phenotypical parameters like protein/DNA-adduct formation.

ACKNOWLEDGMENTS

This study was supported by the research program "Environment and Health" of the Flemish Government (Belgium), and by the Trans-national University Limburg.

References

1. Hall, M. and Grover, P. (1990) Polycyclic aromatic hydrocarbons: metabolism, activation and tumour initiation. In: Cooper CS, Grover PL, eds. *Chemical Carcinogenesis and Mutagenesis*, Vol. I. New York: Raven Press, 1990:327-372.
2. Pitot, H. (1990) Mechanisms of chemical carcinogenesis: theoretical and experimental bases. In: Cooper CS, Grover PL, eds. *Chemical Carcinogenesis and Mutagenesis*, Vol. I. New York: Raven Press, 1990:3-32.
3. Schut, H.A. and Snyderwine, E.G. (1999) DNA adducts of heterocyclic amine food mutagens: implications for mutagenesis and carcinogenesis. *Carcinogenesis*, 20, 353-68.
4. Guengerich, F.P. (2000) Metabolism of chemical carcinogens. *Carcinogenesis*, 21, 345-51.
5. Harris, C.C. (1989) Interindividual variation among humans in carcinogen metabolism, DNA adduct formation and DNA repair. *Carcinogenesis*, 10, 1563-6.
6. Harries, L.W., Stubbins, M.J., Forman, D., et al. (1997) Identification of genetic polymorphisms at the glutathione S-transferase Pi locus and association with susceptibility to bladder, testicular and prostate cancer. *Carcinogenesis*, 18, 641-4.
7. Hengstler, J.G., Arand, M., Herrero, M.E., et al. (1998) Polymorphisms of N-acetyltransferases, glutathione S-transferases, microsomal epoxide hydrolase and sulfotransferases: influence on cancer susceptibility. *Recent Results Cancer Res*, 154, 47-85.
8. Hirvonen, A. (1999) Polymorphic NATs and cancer predisposition. *IARC Sci Publ*, 251-70.
9. Landi, S. (2000) Mammalian class theta GST and differential susceptibility to carcinogens: a review. *Mutat Res*, 463, 247-83.
10. Brennan, P. (2002) Gene-environment interaction and aetiology of cancer: what does it mean and how can we measure it? *Carcinogenesis*, 23, 381-7.
11. Venter, J.C., Adams, M.D., Myers, E.W., et al. (2001) The sequence of the human genome. *Science*, 291, 1304-51.
12. Syvanen, A.C. (2001) Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat Rev Genet*, 2, 930-42.
13. Kristensen, V.N., Kelefiotis, D., Kristensen, T., et al. (2001) High-throughput methods for detection of genetic variation. *Biotechniques*, 30, 318-22, 324, 326 passim.
14. Sokolov, B.P. (1990) Primer extension technique for the detection of single nucleotide in genomic DNA. *Nucleic Acids Res*, 18, 3671.
15. Kuppuswamy, M.N., Hoffmann, J.W., Kasper, C.K., et al. (1991) Single nucleotide primer extension to detect genetic diseases: experimental application to hemophilia B (factor IX) and cystic fibrosis genes. *Proc Natl Acad Sci U S A*, 88, 1143-7.
16. Lindblad-Toh, K., Winchester, E., Daly, M.J., et al. (2000) Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nat Genet*, 24, 381-6.
17. Nakajima, M., Yokoi, T., Mizutani, M., et al. (1999) Genetic polymorphism in the 5'-flanking region of human CYP1A2 gene: effect on the CYP1A2 inducibility in humans. *J Biochem (Tokyo)*, 125, 803-8.
18. Lang, N.P., Butler, M.A., Massengill, J., et al. (1994) Rapid metabolic phenotypes for acetyltransferase and cytochrome P4501A2 and putative exposure to food-borne heterocyclic amines increase the risk for colorectal cancer or polyps. *Cancer Epidemiol Biomarkers Prev*, 3, 675-82.

Multiplex genotyping of carcinogen metabolizing enzymes

19. Sachse, C., Brockmoller, J., Bauer, S., et al. (1999) Functional significance of a C-->A polymorphism in intron 1 of the cytochrome P450 CYP1A2 gene tested with caffeine. *Br J Clin Pharmacol*, 47, 445-9.
20. Wormhoudt, L.W., Commandeur, J.N. and Vermeulen, N.P. (1999) Genetic polymorphisms of human N-acetyltransferase, cytochrome P450, glutathione-S-transferase, and epoxide hydrolase enzymes: relevance to xenobiotic metabolism and toxicity. *Crit Rev Toxicol*, 29, 59-124.
21. Ali-Osman, F., Akande, O., Antoun, G., et al. (1997) Molecular cloning, characterization, and expression in *Escherichia coli* of full-length cDNAs of three human glutathione S-transferase Pi gene variants. Evidence for differential catalytic activity of the encoded proteins. *J Biol Chem*, 272, 10004-12.
22. Wang, Y., Spitz, M.R., Schabath, M.B., et al. (2003) Association between glutathione S-transferase p1 polymorphisms and lung cancer risk in Caucasians: a case-control study. *Lung Cancer*, 40, 25-32.
23. Godschalk, R.W., Dallinga, J.W., Wikman, H., et al. (2001) Modulation of DNA and protein adducts in smokers by genetic polymorphisms in GSTM1, GSTT1, NAT1 and NAT2. *Pharmacogenetics*, 11, 389-98.
24. Hung, H.C., Chuang, J., Chien, Y.C., et al. (1997) Genetic polymorphisms of CYP2E1, GSTM1, and GSTT1; environmental factors and risk of oral cancer. *Cancer Epidemiol Biomarkers Prev*, 6, 901-5.
25. Bell, D.A., Taylor, J.A., Butler, M.A., et al. (1993) Genotype/phenotype discordance for human arylamine N-acetyltransferase (NAT2) reveals a new slow-acetylator allele common in African-Americans. *Carcinogenesis*, 14, 1689-92.
26. Gonzalez, M.V., Alvarez, V., Pello, M.F., et al. (1998) Genetic polymorphism of N-acetyltransferase-2, glutathione S-transferase-M1, and cytochromes P450IIE1 and P450IID6 in the susceptibility to head and neck cancer. *J Clin Pathol*, 51, 294-8.
27. Todesco, L., Torok, M., Krahenbuhl, S., et al. (2003) Determination of -3858G-->A and -164C-->A genetic polymorphisms of CYP1A2 in blood and saliva by rapid allelic discrimination: large difference in the prevalence of the -3858G-->A mutation between Caucasians and Asians. *Eur J Clin Pharmacol*, 59, 343-6.
28. Garte, S., Gaspari, L., Alexandrie, A.K., et al. (2001) Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev*, 10, 1239-48.
29. Gross, M., Kruisselbrink, T., Anderson, K., et al. (1999) Distribution and concordance of N-acetyltransferase genotype and phenotype in an American population. *Cancer Epidemiol Biomarkers Prev*, 8, 683-92.
30. Cascorbi, I., Drakoulis, N., Brockmoller, J., et al. (1995) Arylamine N-acetyltransferase (NAT2) mutations and their allelic linkage in unrelated Caucasian individuals: correlation with phenotypic activity. *Am J Hum Genet*, 57, 581-92.
31. Fryer, A.A. and Jones, P.W. (1999) Interactions between detoxifying enzyme polymorphisms and susceptibility to cancer. *IARC Sci Publ*, 303-22.
32. Bartsch, H., Nair, U., Risch, A., et al. (2000) Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer Epidemiol Biomarkers Prev*, 9, 3-28.
33. Norton, N., Williams, N.M., Williams, H.J., et al. (2002) Universal, robust, highly quantitative SNP allele frequency measurement in DNA pools. *Hum Genet*, 110, 471-8.
34. Turner, D., Choudhury, F., Reynard, M., et al. (2002) Typing of multiple single nucleotide polymorphisms in cytokine and receptor genes using SNaPshot. *Hum Immunol*, 63, 508-13.

Chapter 2

35. Makridakis, N.M. and Reichardt, J.K. (2001) Multiplex automated primer extension analysis: simultaneous genotyping of several polymorphisms. *Biotechniques*, 31, 1374-80.
36. Gwee, P.C., Tang, K., Chua, J.M., et al. (2003) Simultaneous genotyping of seven single-nucleotide polymorphisms in the MDR1 gene by single-tube multiplex minisequencing. *Clin Chem*, 49, 672-6.
37. Landi, S., Gemignani, F., Gioia-Patricola, L., et al. (2003) Evaluation of a microarray for genotyping polymorphisms related to xenobiotic metabolism and DNA repair. *Biotechniques*, 35, 816-20, 822, 824-7.

Chapter 3

Inter-individual variations in DNA adduct levels assessed by analysis of multiple polymorphisms in smokers

Hans B. Ketelslegers¹, Ralph W.H. Gottschalk¹, Roger W.L. Godschalk¹, Ad M. Knaapen¹, Frederik J. van Schooten¹, Robert F.M.H. Vlietinck², Jos C.S. Kleinjans¹
and Joost H.M. van Delft¹

¹ Department of Health Risk Analysis and Toxicology,
Maastricht University, Maastricht, the Netherlands

² Centre for Human Genetics, University of Leuven, Leuven, Belgium

Cancer Epidemiology Biomarkers and Prevention 2006;15(4):624-9

Abstract

Genetic polymorphisms in genes involved in pathways that lead to DNA damage and its repair may explain part of the large inter-individual variation in DNA adduct levels in smokers. To this end, we investigated 19 polymorphisms in 12 genes involved in carcinogen metabolism, DNA repair and oxidative stress defense on DNA adduct levels (determined by ^{32}P -postlabelling) in lymphocytes of 63 healthy Caucasian smokers. The total number of alleles that were categorized as putatively high risk alleles, appeared positively related with bulky DNA adduct levels ($P=0.001$). Subsequently, to investigate which polymorphisms may have the highest contribution to DNA adduct levels observed in these smokers, discriminant analysis was performed. In the investigated set of polymorphisms, *GSTM1**0 ($P<0.001$), *mEH**2 ($P=0.001$) and *GPX1* ($P<0.001$) in combination with the level of exposure ($P<0.001$) were found to be the key effectors. This is the first demonstration of the involvement of GPX in DNA adduct formation. DNA-adduct levels in subjects with a relatively high number of risk alleles for these three genes were more than 2-fold higher than in individuals having no risk alleles (1.97 ± 1.026 adducts per 108 nucleotides versus 0.79 ± 0.49). This study demonstrates that analysis of multiple genetic polymorphisms may predict the inter-individual variations in DNA adduct levels caused by exposure to a complex mixture of compounds such as cigarette smoke. Noteworthy, all three genes are involved in processes on deactivation of reactive carcinogenic metabolites. It is concluded that discriminant analysis presents an important statistical tool studying the impact of multiple genotypes on molecular biomarkers.

Introduction

Cigarette smoke is a complex mixture of chemicals including many genotoxic carcinogens (1). Most of these compounds induce their carcinogenic effect by covalently binding to the DNA to form so-called DNA adducts. It has repeatedly been shown that a dose-response relationship exists between the amount of cigarettes smoked per day, bulky DNA adduct levels and (lung-)cancer risk (2-5). However, within these relationships, large inter-individual variations are observed: individuals with approximately similar exposures may have quite different DNA adduct levels.

It was postulated that genetic polymorphisms in genes involved in the process of DNA adduct formation and repair explain at least part of this variation (6). However, DNA adduct formation after exposure to mixtures of compounds, such as cigarette smoke, and the endogenous response to it, implicates a complex network of pathways in which many different proteins are involved. For instance, some chemical carcinogens in cigarette smoke require metabolic activation before they can react with DNA and exert their genotoxic effect. These reactions are catalyzed by phase I enzymes such as cytochrome P450. Other genotoxic compounds, including the reactive intermediates from the above mentioned reactions, are detoxified by phase II reactions which are catalyzed by enzymes such as glutathione-S-transferases and the *N*-acetyltransferases. The balance between both bioactivation and -deactivation may differ between individuals due to polymorphisms in genes involved in these biotransformation processes. Furthermore, impaired DNA repair capacity as caused by polymorphisms in DNA repair genes may lead to relatively high DNA adduct levels (6,7). There are also indications that oxidative stress may contribute to the formation of bulky adducts in cigarette smokers (8), implying a hypothetical role in DNA adduct formation for polymorphisms in genes modulating the impact of oxidative stress (1). Therefore, some polymorphisms in genes involved in protection against oxidative stress (e.g. *GPX1*) have been included in the present analysis.

Because of the complex exposure of cigarette smokers and the array of processes that are involved in modulation of DNA adduct levels, it is thus concluded that it is very unlikely that there is one single mechanism or gene that can explain the inter-individual differences in DNA adduct levels in smokers upon adjusting for differences in smoking behavior. As a result, studies investigating single polymorphisms in relation to DNA adduct levels might either over- or under-estimate the involvement of these polymorphisms. In order to identify the most relevant genes and polymorphisms and to quantitate possible interactions between them, studies are required that analyze many genes and polymorphisms simultaneously in each individual of an exposed population.

In the present study, 19 polymorphisms in metabolic, DNA repair and oxidative stress related genes were investigated. These polymorphisms were selected on the basis of their known association with cancer development or their hypothetical effect on bulky DNA adduct levels. Multiplex genotyping was performed using a single base

extension-based method (9). We hypothesize that the simultaneous assessment of multiple genotypes yields a better prediction of DNA adduct levels in peripheral lymphocytes of smokers as compared to the analysis of single gene polymorphisms. To our knowledge, such a large number of polymorphisms has never been assessed in a single population in relation to a biomarker like carcinogen-DNA adducts. In the current report, we present a reliable classification of subgroups of smokers (in terms of low, medium and high responders regarding DNA adduct levels) using discriminant analysis as a tool for identifying the most relevant genes and polymorphisms.

Materials and Methods

Study population

Genotyping was performed using lymphocytic DNA from 63 healthy smoking Caucasians (29 males, 34 females) with an average $\pm$ SD age of 43 ± 9 years. These individuals reported smoking between 5 and 50 cigarettes per day (overall mean $\pm$ SD, 26 ± 9 cigarettes per day) for at least 10 years. Because the half life of lymphocytic DNA adducts is short (11 weeks(10)), the amount of cigarettes smoked per day instead of pack-years was used as parameter for exposure. Informed consent was obtained from all individuals.

DNA isolation and ^{32}P -postlabelling

Peripheral blood lymphocytes were isolated by gradient centrifugation on lymphoprep according to Bøyum (11). Standard phenol-extraction procedures were used for the isolation of DNA from peripheral blood lymphocytes. Subsequently, the nuclease P1-enriched ^{32}P -postlabeling assay was performed as described previously (12,13). Quantitation was performed by using phosphor imaging technology (Molecular Dynamics™, Sunnyvale) with detection limit of <0.01 adducts per 108 nucleotides. DNA adduct analysis was performed at least in duplicate, and inter-assay variation was less than 20%.

Selection of polymorphisms

All Single Nucleotide Polymorphisms (SNPs) included in this study (Table 1) were selected on basis of i) their association with cancer development and/or known effects on enzyme activity, ii) their expected influence on DNA adduct levels as measured by ^{32}P -postlabelling, and iii) a frequency of occurrence in the population of more than 5%. DNA sequences and allele frequencies were obtained from the Cancer SNP 500 database (<http://snp500cancer.nci.nih.gov>). The selected polymorphisms can be divided into

Table 1. Overview of polymorphisms included in the study, together with their PCR-primers (Cont'd on next page).

Polymorphism	Frequency (%) [*]	DbSNP ID	PCR Primers	Product (bp) ^{**}	Length (bp)
<i>CYP1A2</i> *1F (-164A>C)	6/43/51	rs762551	F5'- GAGGCTCCCTTCCAGCTCTC- 3' R5'- CTCCCAGCTGGATACCAGA- 3'	106 (1)	44
<i>GSTM1</i> Deletion	40/60		F5'- CTCCTGATTATGACAGAAGCC- 3' R5'- CTGGATTGTAGCAGATCATGC- 3'	648 (1)	48
<i>GSTP1</i> *2 (1404A>G)	46/39/15	rs947894	F5'- TGGTGGACATGGTGAATGAC- 3' R5'- AGCCCTTTCTTTGTTCAAGC- 3'	123 (1)	28
*3 (2294C>T)	89/11/0	rs1799811	F5'- TGGGAGGGATGAGAGTAGGA- 3' R5'- CAGGGTCTCAAAAGGCTTCA- 3'	106 (1)	52
<i>GSTT1</i> Deletion	85/15		F5'- GTAGCCATCACGGAGCTGAT- 3' R5'- GGCAGCATAAGCAGGACTTC- 3'	97 (1)	56
<i>NAT2</i> *5 (341T>C)	39/49/12	rs1801280	F5'- CAAATACAGCACTGGCATGG- 3' R5'- GGCTGATCCTTCCCAGAAAT- 3'	13 (1)	60
*6 (590G>A)	49/42/9	rs1799930	F5'- CCTGCCAAAGAAAGAAACACC- 3' R5'- GGGTCTGCAAGGAACAAAAT- 3'	143 (1)	36
*7 (857G>A)	92/8/0	rs1799931	F5'- TCCTTGGGGAGAAATCTCGT- 3' R5'- GGGTGATACATACACAAGGGTTT- 3'	92 (1)	64
<i>mEH</i> *2 (C>T)	48/47/5	rs1051740	F5'-CTCTCAACTTGGGGTCTCTGA- 3' R5'-GGCGTTTTCGAAACATACCT- 3'	231 (3)	46
*3 (A>G)	59/38/3	rs2234922	F5'-CGTGCAGGGTCTCTCTCTC- 3' R5'- GTTCTTGGGGTCAGTCAGGA- 3'	194 (3)	50
<i>XRCC1</i> *2 (26304C>T)	89/11/0	rs1799782	F5'-TGAAGGAGGAGGATGAGAGC-3' R5'-CTCTACCCTCAGACCCACGA-3'	147 (2)	21
*3 (27466G>A)	94/6/0	rs25489	F5'-CCCCAGTGGTGCTAACCTAA-3' R5'-GGGGTTTGCCTGTCACTG-3'	116 (2)	25
*4 (28152G>A)	45/40/15	rs25487	F5'-TAAGGAGTGGGTGCTGGACT-3' R5'-ATTGCCAGCACAGGATAAG-3'	101 (2)	37
<i>XRCC3</i> *1 (18067C>T)	39/40/21	rs861539	F5'-GCCTGGTGGTCATCGACTC-3' R5'-ACAGGGCTCTGGAAGGCA-3'	136 (2)	63
<i>XPD</i> *5 (35931A>C)	35/49/15	rs1052559	F5'-TTCTCTGCAGGAGGATCAGC-3' R5'-CTCAGGAGTCACCAGGAACC-3'	146 (2)	45
<i>BRCA2</i> *1 (-26G>A)	60/34/6	rs1799943	F5'-AAATTTTCCAGCGCTTCTGA-3' R5'-AATGTTGGCCTCTCTTTGGA-3'	159 (2)	57
*3 (1342A>C)	48/44/8	rs144848	F5'-AGCAAACGCTGATGAATGTG-3' R5'-TTGGAGATTITGTCACCTCCAC-3'	150 (2)	53
<i>NQO1</i> *2 (609C>T)	91/9/0	rs1800566	F5'-TGAAGTCAAGGAGGTGGAGGT- 3' R5'- CTGGTTTGAGCGAGTGTTC- 3'	240 (3)	26
<i>GPX1</i> *1 (593C>T)	57/34/9	rs1050450	F5'-ACTGGGATCAACAGGACCAG- 3' R5'- TTGACATCGAGCCTGACATC- 3'	213 (3)	38

^{*}Frequencies in the currently investigated population are shown as fully wild types/heterozygous/fully mutants. In case of *GSTM1* and *GSTT1*, no differences can be made between wild types (no deletions) and heterozygous gene deletions

^{**}(1) 8-plex PCR, (2) 7-plex PCR, (3) 4-plex PCR (see material and methods)

Table 1 (Cont'd). Overview of polymorphisms included in the study, together with their SBE primers.

Polymorphism		SBE Primers ^a	Length (bp)
<i>CYP1A2</i>	*1F	5'-AACTGACTAACTAGGTGCCACTCAAAGGGTGAGCTCTGTGGGC- 3'	44
<i>GSTM1</i>	Del	5'-AACTGACTAACTAGGTGCCACGCTGGAGAACCAGACCATGGACAACC- 3'	48
<i>GSTP1</i>	*2	5'-AACTC TGGAGGACCTCCGCTGCAAAATAC- 3'	28
	*3	5'-AACTGACTAACTAGGTGCCACGTCGTGAC CATGGTGGTGTCTGGCAGGAGG- 3'	52
<i>GSTT1</i>	Del	5'-AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGGGCAGGTGAACCCACTAG-GC- 3'	56
<i>NAT2</i>	*5	5'-AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTATTCACCTTCTCCTGCAGGTG-ACCA- 3'	60
	*6	5'- CCTACCAAAAAATATACTTATTTACGCTTGAACCTC- 3'	36
	*7	5'-AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACAGCCCTCGTGCCCAAAC-CTGGTGATG-3'	64
<i>mEH</i>	*2	5'-AACTGACTAACTAGGTGGAAGAAGCAGGTGGAG-ATTCTCAACAGA- 3'	46
	*3	5'-AACTGACTAACTAGGTGCCACGTCGTGAAAGTC-CAGCTGCCCGCAGGCC- 3'	50
<i>XRCC1</i>	*2	5'-CGGGGGCTCTCTTCTTCAGC- 3'	21
	*3	5'-TCTTCTCCAGTGCCAGCTCCAACTC- 3'	25
	*4	5'-AACTGACTAACTAGTTGGCGTGTGAGGCCCTTACCTC- 3'	37
<i>XRCC3</i>	*1	5'-AACTGACTAACTAGGTGCCACGTCGTGAAAGTCTGACATGCGCACTGCTCAGCTC-ACGCAGC- 3'	63
<i>XPB</i>	*5	5'-AACTGACTAACTAGGTGGCTGCTGAGCAATCTGCTCTA-TCCTCT- 3'	45
<i>BRCA2</i>	*1	5'-AACTGACTAACTAGGTGCCACGTCGAGGTCTTCTGTTTTCAGACTTATTTACC-AA- 3'	57
	*3	5'-AACTGACTAACTAGGTGTAAATGATACTGATCCATTAG-ATTCAAATGTAGCA- 3'	53
<i>NQO1</i>	*2	5'-AAGCATTCAAGAACCATCCACCTACCC- 3'	26
<i>GPX1</i>	*1	5'-AAATAACTAACTAGGTGCGGCGCCCTAGGCACAGCTG- 3'	38

* Neutral non-binding tails are in italics (see material and methods)

3 groups: 6 genes involved in carcinogen metabolism (biotransformation), 4 DNA repair genes and 2 genes involved in oxidative stress processes.

PCR Primer design and multiplex PCR amplification

Primer 3 software (www.broad.mit.edu/cgi-bin/primer/primer3_www.cgi) and Net-primer software (www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html) were used to design PCR primers. All PCR primers were designed to have a melting temperature (T_m) of $60 \pm 2^\circ\text{C}$ and a primer length of 19-23 nucleotides (Table 1). See Knaapen et al. (9) for more detailed information about primer design.

PCR was performed in three separate multiplex PCR reactions: one 8-plex, one 7-plex and one 4-plex reaction (indicated as 1-3 respectively in Table 1). PCR was carried out in a Tgradient 96 well Thermal cycler (Biometra, Goettingen, Germany) in a 10 μl volume, containing 1x PCR buffer (Invitrogen, Breda, The Netherlands), 0.2 mM dNTP's (Invitrogen), 0.5 mM MgCl_2 (Invitrogen), 0.25 U Platinum[®] Taq-Polymerase (Invitrogen), and 40 ng template DNA. The final concentrations of the primer pairs in this reaction mixture were 2 μM . PCR conditions were 94°C for 3 min (denaturation); 30 cycles of 94°C for 30 sec, 56°C for 30 sec (for multiplex 1; 60°C for 2 and 57°C for 3) and 72°C for 30 sec; and a final extension for 5 minutes at 72°C .

PCR product purification

Contaminating dNTP's and PCR primers were removed after PCR-amplification to avoid participation in the subsequent primer extension reaction. 10 μl of multiplex PCR

product were mixed with 4 µl Exo-SAP-IT (containing exonuclease I and shrimp alkaline phosphatase, Amersham, Roosendaal, The Netherlands) and incubated at 37°C for 45 min. Purified multiplex PCR products were stored at -20°C until genotyping.

Multiplex Genotyping using Single Base Extension

Genotyping was performed by Single Base Extension (SBE) using SnaPshot™ (Applied Biosystems, Nieuwekerk a.d. IJssel, The Netherlands) as described (9). Single base extension primers were designed using Primer 3 and Netprimer software, with a template specific part of 20-33bp and a T_m of 66-69°C (Table 1). The extension-primers were designed to bind immediately adjacent 5' to the specific SNP. To facilitate detection of all the polymorphisms in one single run, the length of the extension primers was adjusted to a distinct size, by addition of a neutral non-binding sequence to their 5' side, as described by Lindblad-Toh et al., 2000 (14) (Table 1). Three µl of the ExoSAP-IT purified PCR product were mixed with 3.5 µl SNaPshot multiplex ready reaction mix (containing AmpliTaq® polymerase, PCR buffer and fluorescently-labelled ddNTP's), 1 µl of the pooled SBE primers and 2.5 µl of PCR water (perform in dimmed light). The final concentration of the SBE primers was 0.2 µM. Single Base Extension was performed in a thermal cycler (T gradient, Biometra) using 25 cycles of 96°C (10 sec) and 60°C (30 sec). After Single Base extension, the samples were incubated at 37°C (1 hr) with 1 U shrimp alkaline phosphatase (SAP, Amersham) in order to degrade the unincorporated ddNTP's. Single base extension reactions were performed in three separate multiplex genotyping experiments on the multiplex PCR reactions as described above.

Genescan analysis

Single base extension products were diluted 5x in water. One µl of the diluted sample was mixed with 13 µl deionized formamide and 0.4 µl of Genescan™ 120 LIZ™ size standard and denatured at 95°C for 5 minutes. Subsequently, the samples were analyzed on an ABI Prism® 3100 genetic analyzer using Genescan™ Analysis software (Version 3.7). For more detailed information see Knaapen et al., 2004 (9).

Statistical Analysis

Linear regression analysis was conducted to investigate the relationship between DNA adduct levels and the amount of cigarettes smoked per day. To investigate which genetic polymorphisms have the highest contribution to the inter-individual variation within this relationship, the genotypes were coded based on the number of polymorphic alleles: 0 (two wild type alleles), 1 (heterozygous, one polymorphic variant allele) and 2 (homozygous mutant, two polymorphic alleles). In case of deletions (*GSTM1* and *GSTT1*), the wild type was coded 0 and the deletion was coded 2. This time, no *a priori* knowledge on expected biological effects of the polymorphisms (their effect on enzyme activity in

relation to DNA adduct formation) was considered when coding the genotypes. Subsequently, SNPs in the same gene yielding the same phenotypic effect, were merged to one single variable for that gene. This was done for *NAT2*, *BRCA2* and *GSTP1*. Note that for *mEH2* and *XRCC1* also more than one SNP was investigated; however, these SNPs have opposite phenotypic effects and therefore can not be combined. This eventually led to 15 genotypes. To evaluate the association of each single polymorphism with respect to DNA adduct level, conventional methods, like Mann-Whitney U tests (for two groups) and Jonckheere-Terpstra tests (for more than two groups) were performed. Second, to investigate the association of multiple polymorphisms, the individuals were divided into subgroups based on their response to exposure (amount of cigarettes smoked per day) in terms of DNA adduct levels (see Figure 1). Three subgroups (classes) were formed based on the regression line: class 1 if the observed adduct level is ≤ 0.66 times the expected value according to the regression line; class 3 if the observed adduct level is ≥ 1.5 times the expected value according to the regression line and class 2 with all other individuals (Figure 1). Using discriminant analysis each individual can be classified. Stepwise discriminant analysis was performed with the dependent variable “class” as grouping variable and all genotypes, age, gender and cigarettes per day as independent variables. Subsequently, results were cross-validated by the leave-one-out method. In all statistical tests, $P < 0.05$ was considered statistically significant. All statistics were performed using SPSS for Windows (version 11.5). Results are expressed as mean $\pm$ SD.

Results

Overall analysis of DNA adduct levels

DNA adduct analysis in peripheral blood lymphocytes demonstrated the existence of so-called diagonal radioactive zones, which represent bulky DNA adducts typically to result from exposure to cigarette smoke. The mean DNA adduct level was 1.40 ± 0.79 adducts per 10^8 nucleotides and ranged from <0.25 to 3.90 adducts per 10^8 nucleotides. A significant relationship was observed between the self-reported number of cigarettes smoked per day (exposure) and DNA adduct level ($P < 0.001$, Figure 1). As expected, large inter-individual variations were observed within this relationship; individuals with approximately similar exposures can have different DNA adduct levels.

Figure 1. Relationship between individual DNA adduct level and exposure (number of cigarettes smoked per day) ($R^2=0.714$, $P<0.001$). Individuals were divided according to the regression line into three classes for discriminant analysis: class 1 if the observed adduct level was ≤ 0.66 times the expected value according to the regression line (●; below the lower dashed line); class 3 if the observed adduct level was ≥ 1.5 times the expected value according to the regression line (○; above the upper dashed line) and class 2 with all other individuals (x; in-between the dashed lines). The regression line was forced through the origin, since in non-smokers, no DNA adducts are measurable.

Multiplex SBE Genotyping

Clear signals could be obtained for all polymorphisms in all individuals (data not shown). No signals were detected in negative control SBE reactions (using water as template), indicating that no false positive signals due to primer-dimers or hair-pins were generated. In Table 1, all observed genotype frequencies in the study population are shown. The distributions of all genotypes in the total population appeared to be in Hardy-Weinberg equilibrium.

Sum of total putative risk alleles in relation to DNA adduct levels

As a first approach to investigate whether the genotype affects DNA adduct levels, the polymorphisms were *a priori* categorized as low or high risk alleles based on their expected modulating effect on DNA adduct levels (Table 2). Subsequently, the sum of risk alleles was computed for each individual and linear regression demonstrated a significant association between these sums and DNA adduct levels ($P=0.001$; Figure 2). This was not due to differences in exposure, as the amount of cigarettes smoked per day did not correlate with the sum of risk alleles (Figure 2).

Table 2. Effects of the selected polymorphisms on enzyme function and DNA adduct levels

Polymorphism		Effect on enzymatic function	Expected effect on DNA adduct level	Reference
Biotransformation				
CYP1A2	*1F	Higher inducibility	Increased bioactivation, higher adduct levels	(25)
GSTM1	*0 del	Deletion, no enzyme activity	Decreased detoxification, higher adduct levels	(16, 26)
GSTP1	*2 I105V	} Decreased enzyme activity	} Decreased detoxification, higher adduct levels	(26, 27)
	*3 A114V			
GSTT1	*0 del	Deletion, no enzyme activity	Decreased detoxification, higher adduct levels	(26)
NAT2	*5 I114T	} Decreased enzyme activity	} Less N-acetylation; decreased detoxification; higher adduct levels	(13)
	*6 R197Q			
	*7 G286E			
mEH	*2 Y113H	Decreased enzyme activity	As phase II; decreased detoxification; increased adduct levels	(18, 28, 29)
	*3 H139R	Increased enzyme activity	As phase I; increased bioactivation; increased DNA adduct levels	(18, 28, 29)
DNA Repair				
XRCC1	*2 R194W	Increased enzyme activity	Increased repair capacity, lower adduct levels	(30-32)
	*3 R280H	} Decreased enzyme activity	} Reduced repair capacity, higher adduct levels	(30)
	*4 Q399R			(7, 33)
XRCC3	*1 T241M	Decreased enzyme activity	Reduced repair capacity, higher adduct levels	(7)
XPD	*5 K751Q	Decreased enzyme activity	Reduced repair capacity, higher adduct levels	(34, 35)
BRCA2	*1 D991N	} Decreased enzyme activity	} Reduced repair capacity, higher adduct levels	(36)
	*3 N372H			
Oxidative Stress				
NQO1	*2 P187S	Reduced enzyme activity	Higher DNA adduct levels	(26)
GPX1	*1 P198L	Less efficient final GSH peroxidase complex	Higher DNA adduct levels	(20)

Figure 2. Relationship between DNA adduct levels the sum of risk alleles (mean $\pm$ SE, $p=0.001$), Figure 3A. N indicates the number of individuals). The sum of risk alleles per individual was computed by adding the number of polymorphisms that putatively increase DNA adduct levels (see Table 2). There were no significant differences in exposure levels between the different groups (mean $\pm$ SE, $P=0.93$, Figure 3B).

Uni-variate analysis

Analysis of the association of each polymorphism separately with DNA adduct levels, only revealed a significant role for *GSTM1*. DNA adduct levels were significantly higher in *GSTM1* null individuals (1.59 ± 0.80 per 10^8 nucleotides) than in *GSTM1* positive subjects (1.05 ± 0.50 , $P<0.01$).

Multivariate analyses

In order to analyze the contribution of all genotypes simultaneously to the inter-individual variation in DNA adduct levels, discriminant analysis was conducted. The population was divided into three subgroups (low, medium or high responders). Discriminant analysis with all genotypes, age, gender and exposure as independent variables, was used to predict the classification into these three subgroups. The predictors that were significant for this discrimination appeared to be *GSTM1* ($P<0.001$), *mEH*2* ($P=0.001$), *GPX1* ($P<0.001$) and exposure ($P<0.001$). Classification results are shown in Table 3. 65.1% of the original grouped cases and 60.3% of the cross-validated grouped cases were correctly classified, whereas prediction by chance would have been 34% only.

Table 3. Classification results obtained by discriminant analysis. 65% of original grouped cases and 60% from cross-validated grouped cases were correctly classified

	Class	Predicted group membership			Total
		1	2	3	
Original (Counts)	1	13	3	1	17
	2	3	17	6	26
	3	2	7	11	20
Cross Validated* (Counts)	1	11	5	1	17
	2	3	17	6	26
	3	2	8	10	20

*In cross-validation, each case is classified by the functions derived from all cases other than that particular case (N=63-1, leave-one-out method).

*Effect of the sum of GSTM1, mEH*2 and GPX on DNA adduct levels*

Within this selection of putatively risk alleles, *GSTM1*, *mEH*2* and *GPX* were found to have the strongest effect on the prediction of the DNA adduct level in smokers. The combined effects of these three polymorphisms on DNA adduct level are shown in Figure 3. The sum of risk alleles of *GSTM1*0*, *mEH*2* and *GPX* was associated with bulky DNA adduct level ($P<0.001$). Individuals having four risk alleles for these three genes had higher DNA adduct levels (1.97 ± 1.026 per 10^8 nucleotides) than individuals not possessing these particular risk alleles (0.79 ± 0.49).

Figure 3. Effect of the combined genotypes of *GSTM1*, *mEH*2* and *GPX1* on bulky DNA adduct levels. The sum of risk alleles for the genes that were found to be the key effectors, is associated with DNA adduct level (a) ($R^2=0.219$, $P<0.001$). Exposure levels did not vary between groups (b) ($R^2=0.009$, $P=0.471$). *N* is the number of individuals per group.

Discussion

Several studies have shown that there is a significant relationship between the amount of cigarettes smoked per day and DNA adduct levels in smoking individuals (1,5,13). However, large inter-individual differences in DNA adduct levels were observed in individuals with apparently similar exposures. Our results support the hypothesis that genetic polymorphisms can explain at least part of this inter-individual variation. The total sum of putatively high risk alleles in 12 genes correlated with DNA adduct levels in smokers.. Subsequently, in our set of 19 polymorphisms we identified *GSTM1*, *mEH* and *GPX1* in combination with exposure as the most relevant determinants of lymphocytic DNA adduct levels in smokers. Noteworthy, all three genes are involved in processes of deactivation of reactive carcinogenic metabolites. Furthermore, this is the first demonstration of the involvement of *GPX1* in DNA adduct formation.

GSTM1 and *mEH* and *GPX1* are all metabolic enzymes involved in the biotransformation of several carcinogenic compounds, including PAHs (15-18). The *GPX* enzyme functions in the detoxification of organic peroxides, like hydrogen peroxide or the oxidative stress caused by cigarette smoke, and in the conjugation of PAH-diols to glutathione (1,19,20). In *GPX1*, the *Pro*¹⁹⁸*Leu* substitution has been associated with a lower enzyme activity and increased lung cancer risk (20). Therefore, this allelic variant leads to less detoxification and hence, higher DNA adduct levels. Indeed, we observe that individuals carrying the slow allelic variant for *GPX1* have higher adduct levels compared to wild type individuals.

Our data confirm previous studies which described that individuals lacking the *GSTM1* enzyme have higher DNA adduct levels as compared to *GSTM1* positive individuals (13,21,22). This agrees with the function of *GSTM1* of scavenging DNA-reactive chemical compounds. For *mEH*, the functional effects of the polymorphisms on DNA adduct levels are more complicated. The most intensively studied polymorphisms in this gene, which were also investigated in the present study, are the *Tyr*¹¹³*His* (*mEH**2; exon 3) and *His*¹³⁹*Arg* (*mEH**3; exon 4) variants. It has been described that the first variant results in a decreased *mEH* activity of approximately 40%, whereas the latter results in increased enzymatic activity (17). In our study, an association was found for *mEH**2, showing that individuals carrying the decreased activity variant had higher DNA adduct levels. In this perspective, *mEH* functions as a phase II enzyme, in which the slow allelic variant (*mEH**2) resulted in increased concentrations of epoxide intermediates and hence higher DNA adduct levels.

As was described above, a relationship was found between the sum of risk alleles and bulky DNA adduct levels (Figure 2). This was also shown in a comparable approach by Matullo et al. (23). When focusing on the sum of risk alleles of the three polymorphisms that were identified to be the key effectors using discriminatn analysis (*GSTM1*, *mEH**2 and *GPX1*), this relationship was enriched: individuals having a relatively high sum of risk alleles for these three genes, showed more than 2-fold higher

Chapter 3

adduct levels than individuals having no such risk alleles, at the same level of exposure (Figure 3).

All investigated polymorphisms in this study were selected on the basis of their high frequencies in Caucasian populations (>5%), expected effects on enzyme activities associated with cancer risk, and the expected effects on bulky DNA adduct levels as detected by ^{32}P -postlabeling. They comprise three groups: biotransformation genes (phase-1 genes like *CYP450s* and phase-2 genes like *GSTs*), DNA repair genes (e.g. *XRCC* and *XPB*), and oxidative stress related genes (e.g. *GPX1*). Although some of these genes contain additional polymorphisms and more genes may be involved in the modulation of DNA adduct levels, those included in this study comprise a representative set of polymorphisms matching our selection criteria. As a consequence, for instance *CYP1A1* which is expected to be an important enzyme in the biotransformation of polycyclic aromatic hydrocarbons (PAHs) present in cigarette smoke (24), was not included in the assay because the polymorphic frequency in Caucasians is low and the effects of *CYP1A1* polymorphisms on enzyme activity and cancer risk are controversial (25).

As we state in our introduction, single genes (or polymorphisms) will never completely explain the inter-individual variations in DNA adduct levels caused by cigarette smoking (23). We therefore focused on a combination of polymorphisms, using discriminant analysis. *mEH*2* and *GPX1*1*, which were found to be non-significant in the univariate analysis, may therefore be significant when investigating multiple polymorphisms simultaneously in the multivariate analysis, for instance because of interactions.

To conclude, our data indicate that assessing multiple genetic polymorphisms explain part of the inter-individual variations in DNA adduct levels and that the analysis of many genotypes simultaneously is important to obtain better insights in the mechanisms that modulate DNA adduct levels. Furthermore, for high-throughput genotyping studies, we consider discriminant analysis as a meaningful statistical tool to investigate the effects of multiple genotypes.

References

1. Hecht, S.S. (1999) Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst*, 91, 1194-210.
2. Veglia, F., Matullo, G. and Vineis, P. (2003) Bulky DNA adducts and risk of cancer: a meta-analysis. *Cancer Epidemiol Biomarkers Prev*, 12, 157-60.
3. Van Schooten, F.J., Godschalk, R.W., Breedijk, A., et al. (1997) 32P-postlabelling of aromatic DNA adducts in white blood cells and alveolar macrophages of smokers: saturation at high exposures. *Mutat Res*, 378, 65-75.
4. Phillips, D.H. (2002) Smoking-related DNA and protein adducts in human tissues. *Carcinogenesis*, 23, 1979-2004.
5. Dallinga, J.W., Pachon, D.M., Wijnhoven, S.W., et al. (1998) The use of 4-aminobiphenyl hemoglobin adducts and aromatic DNA adducts in lymphocytes of smokers as biomarkers of exposure. *Cancer Epidemiol Biomarkers Prev*, 7, 571-7.
6. Duell, E.J., Wiencke, J.K., Cheng, T.J., et al. (2000) Polymorphisms in the DNA repair genes XRCC1 and ERCC2 and biomarkers of DNA damage in human blood mononuclear cells. *Carcinogenesis*, 21, 965-71.
7. Matullo, G., Palli, D., Peluso, M., et al. (2001) XRCC1, XRCC3, XPD gene polymorphisms, smoking and (32)P-DNA adducts in a sample of healthy subjects. *Carcinogenesis*, 22, 1437-45.
8. Gupta, R.C., Arif, J.M. and Gairola, C.G. (1999) Enhancement of pre-existing DNA adducts in rodents exposed to cigarette smoke. *Mutat Res*, 424, 195-205.
9. Knaapen, A.M., Ketelslegers, H.B., Gottschalk, R.W., et al. (2004) Simultaneous genotyping of nine polymorphisms in xenobiotic-metabolizing enzymes by multiplex PCR amplification and single base extension. *Clin Chem*, 50, 1664-8.
10. Godschalk, R.W., Feldker, D.E., Borm, P.J., et al. (2002) Body mass index modulates aromatic DNA adduct levels and their persistence in smokers. *Cancer Epidemiol Biomarkers Prev*, 11, 790-3.
11. Boyum, A. (1976) Isolation of lymphocytes, granulocytes and macrophages. *Scand J Immunol*, Suppl 5, 9-15.
12. Godschalk, R.W., Maas, L.M., Van Zandwijk, N., et al. (1998) Differences in aromatic-DNA adduct levels between alveolar macrophages and subpopulations of white blood cells from smokers. *Carcinogenesis*, 19, 819-25.
13. Godschalk, R.W., Dallinga, J.W., Wikman, H., et al. (2001) Modulation of DNA and protein adducts in smokers by genetic polymorphisms in GSTM1, GSTT1, NAT1 and NAT2. *Pharmacogenetics*, 11, 389-98.
14. Lindblad-Toh, K., Winchester, E., Daly, M.J., et al. (2000) Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nat Genet*, 24, 381-6.
15. Coles, B. and Ketterer, B. (1990) The role of glutathione and glutathione transferases in chemical carcinogenesis. *Crit Rev Biochem Mol Biol*, 25, 47-70.
16. Ketterer, B., Harris, J.M., Talaska, G., et al. (1992) The human glutathione S-transferase supergene family, its polymorphism, and its effects on susceptibility to lung cancer. *Environ Health Perspect*, 98, 87-94.

17. Hassett, C., Robinson, K.B., Beck, N.B., et al. (1994) The human microsomal epoxide hydrolase gene (EPHX1): complete nucleotide sequence and structural characterization. *Genomics*, 23, 433-42.
18. Hassett, C., Lin, J., Carty, C.L., et al. (1997) Human hepatic microsomal epoxide hydrolase: comparative analysis of polymorphic expression. *Arch Biochem Biophys*, 337, 275-83.
19. Foureman, G.L. and Eling, T.E. (1989) Peroxidase-mediated formation of glutathione conjugates from polycyclic aromatic dihydrodiols and insecticides. *Arch Biochem Biophys*, 269, 55-68.
20. Ratnasinghe, D., Tangrea, J.A., Andersen, M.R., et al. (2000) Glutathione peroxidase codon 198 polymorphism variant increases lung cancer risk. *Cancer Res*, 60, 6381-3.
21. Butkiewicz, D., Grzybowska, E., Hemminki, K., et al. (1998) Modulation of DNA adduct levels in human mononuclear white blood cells and granulocytes by CYP1A1 CYP2D6 and GSTM1 genetic polymorphisms. *Mutat Res*, 415, 97-108.
22. Rojas, M., Alexandrov, K., Cascorbi, I., et al. (1998) High benzo[a]pyrene diol-epoxide DNA adduct levels in lung and blood cells from individuals with combined CYP1A1 MspI/Msp-GSTM1*0/*0 genotypes. *Pharmacogenetics*, 8, 109-18.
23. Matullo, G., Peluso, M., Polidoro, S., et al. (2003) Combination of DNA repair gene single nucleotide polymorphisms and increased levels of DNA adducts in a population-based study. *Cancer Epidemiol Biomarkers Prev*, 12, 674-7.
24. Shimada, T., Yun, C.H., Yamazaki, H., et al. (1992) Characterization of human lung microsomal cytochrome P-450 1A1 and its role in the oxidation of chemical carcinogens. *Mol Pharmacol*, 41, 856-64.
25. Bartsch, H., Nair, U., Risch, A., et al. (2000) Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer Epidemiol Biomarkers Prev*, 9, 3-28.
26. Wu, X., Zhao, H., Suk, R., et al. (2004) Genetic susceptibility to tobacco-related cancer. *Oncogene*, 23, 6500-23.
27. Watson, M.A., Stewart, R.K., Smith, G.B., et al. (1998) Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution. *Carcinogenesis*, 19, 275-80.
28. Zhou, W., Thurston, S.W., Liu, G., et al. (2001) The interaction between microsomal epoxide hydrolase polymorphisms and cumulative cigarette smoking in different histological subtypes of lung cancer. *Cancer Epidemiol Biomarkers Prev*, 10, 461-6.
29. Georgiadis, P., Demopoulos, N.A., Topinka, J., et al. (2004) Impact of phase I or phase II enzyme polymorphisms on lymphocyte DNA adducts in subjects exposed to urban air pollution and environmental tobacco smoke. *Toxicol Lett*, 149, 269-80.
30. Ratnasinghe, D., Yao, S.X., Tangrea, J.A., et al. (2001) Polymorphisms of the DNA repair gene XRCC1 and lung cancer risk. *Cancer Epidemiol Biomarkers Prev*, 10, 119-23.
31. David-Beabes, G.L. and London, S.J. (2001) Genetic polymorphism of XRCC1 and lung cancer risk among African-Americans and Caucasians. *Lung Cancer*, 34, 333-9.
32. Lunn, R.M., Langlois, R.G., Hsieh, L.L., et al. (1999) XRCC1 polymorphisms: effects on aflatoxin B1-DNA adducts and glycophorin A variant frequency. *Cancer Res*, 59, 2557-61.
33. Duell, E.J., Holly, E.A., Bracci, P.M., et al. (2002) A population-based study of the Arg399Gln polymorphism in X-ray repair cross-complementing group 1 (XRCC1) and risk of pancreatic adenocarcinoma. *Cancer Res*, 62, 4630-6.
34. Benhamou, S. and Sarasin, A. (2002) ERCC2/XPD gene polymorphisms and cancer risk. *Mutagenesis*, 17, 463-9.

Genetic susceptibility in smokers

35. Palli, D., Russo, A., Masala, G., et al. (2001) DNA adduct levels and DNA repair polymorphisms in traffic-exposed workers and a general population sample. *Int J Cancer*, 94, 121-7.
36. Coughlin, S.S., Khoury, M.J. and Steinberg, K.K. (1999) BRCA1 and BRCA2 gene mutations and risk of breast cancer. Public health perspectives. *Am J Prev Med*, 16, 91-8.

Chapter 4

Potential role of CYP1B1 in the metabolic activation of 4- AminoBiPhenyl in humans

Hans B. Ketelslegers¹, Roger W.L. Godschalk, Bart J.M. Eskens, Jan W. Dallinga,
Ralph W.H. Gottschalk, Frederik J. van Schooten, Jos C.S. Kleinjans
and Joost H.M. van Delft

Department of Health Risk Analysis and Toxicology,
Maastricht University, Maastricht, the Netherlands

Abstract

Metabolites of the human carcinogen 4-aminobiphenyl (4-ABP), one of the major carcinogenic constituents of cigarette smoke, form hemoglobin (Hb) adducts. Therefore, 4-ABP-Hb adducts represent a useful biomarker for exposure to 4-ABP, which has been associated with the occurrence of bladder cancer. However, with respect to 4-ABP-Hb adduct formation, not every individual responds to a similar degree to 4-ABP exposure and consequently, an individual's risk of ultimately developing effects in response to exposure may be influenced by genetic polymorphisms in biotransformation enzymes. The hypothesis of the current study therefore is that genetic polymorphisms in genes coding for enzymes with a putative effect on 4-ABP-Hb metabolism explain the variations in 4-ABP-Hb adduct formation in response to cigarette smoking. This study aims to determine the most important genetic polymorphisms with respect to 4-ABP metabolism, and their influence on the dose-response relationship between cigarette smoking and 4-ABP-Hb adduct formation. 4-ABP-Hb adducts were measured in blood samples from 57 smoking and 10 non-smoking volunteers. An association was found between cigarette smoking and 4-ABP-Hb adduct levels in smokers ($R^2=0.5$, $P<0.001$). Subsequently, subjects were genotyped for 12 polymorphisms in 7 genes involved in biotransformation reactions, all selected based on their putative involvement in 4-ABP metabolism and/or 4-ABP-Hb adduct formation. From this selection of polymorphisms, a significant impact was found for the *CYP1B1* *Leu*⁴³²*Val* polymorphism ($P=0.021$), which has been reported to lead to a decrease in enzyme activity. Indeed higher levels of 4-ABP-Hb adducts were observed in homo- and heterozygous carriers of the *CYP1B1* ⁴³²*Leu* as compared to the double *CYP1B1* ⁴³²*Val* genotype. A significant interaction between these *CYP1B1* genotypes and the level of exposure was found ($P=0.003$). Noteworthy, a saturation effect was observed for 4-ABP-Hb adduct formation at high smoking doses in carriers of the *CYP1B1* ⁴³²*Leu* allele, whereas a linear relationship without any evidence of saturation was observed in subjects homozygously carrying the *CYP1B1* ⁴³²*Val* polymorphism. No effect of polymorphisms in other genes were found. This is the first study in humans suggesting a crucial role of the CYP1B1 enzyme in 4-ABP metabolism, indicating a protective effect of the *CYP1B1* *Leu*⁴³²*Val* polymorphism against the formation of 4-ABP-Hb adduct levels, depending on the smoking dose.

Introduction

Inhalation of mainstream cigarette smoke is the major source of human exposure to the arylamine 4-aminobiphenyl (4-ABP) (1). 4-ABP has been classified by the International Agency for Research on Cancer (IARC) as carcinogenic to humans and is related to several types of cancers, with urinary bladder cancer being most established (1-3). This carcinogenic effect is mainly exerted by the reactive *N*-hydroxy-ABP intermediate which is formed by metabolic transformation of 4-ABP in the liver. The main enzymes involved in this so-called *N*-hydroxylation are the cytochrome P450's (CYP450). It is generally accepted that CYP1A2 is the primarily responsible enzyme for metabolic activation of 4-ABP to *N*-hydroxy-ABP (4,5). *N*-hydroxy-ABP is known to bind to macromolecules such as proteins, including hemoglobin. Therefore, levels of 4-ABP-hemoglobin (4-ABP-Hb) adducts in blood can be used as a biomarker for exposure to 4-ABP (1,2,6,7). Indeed, good associations between cigarette smoke exposure and the levels of 4-ABP-Hb adducts were observed (8), but also, relatively large inter-individual variations were found, which could be due to genetic variations in enzymes involved in metabolic activation and detoxification (mainly by Glutathione-S-Transferases (e.g. GSTM1) and *N*-acetyltransferases (NAT1 and NAT2) (9)) of 4-ABP.

Single Nucleotide Polymorphisms (SNPs) in the *CYP1A2* gene have been associated with variations in enzyme activity and expression of this gene is induced in smokers, leading to even higher activities of the enzyme (4,10). Therefore, individuals exhibiting high CYP1A2 activity and exposed to 4-ABP, are expected to have increased levels of *N*-hydroxy-ABP and hence, to show higher 4-ABP-Hb adduct levels. However, in recent studies in laboratory animals the role of CYP1A2 in this 4-ABP metabolic activation has been challenged, because also in *CYP1A2* knockout animals, 4-ABP related adducts were detected, at an even higher level compared to *CYP1A2* wild-type mice. As a result it has been suggested that CYP1A2 may not be the primary enzyme for 4-ABP *N*-hydroxylation, implying that other (CYP450-) enzymes may be involved (11,12).

Analysis of SNPs in metabolic genes may provide information about candidate enzymes involved in the biotransformation of 4-ABP, and thus about the modulation of the dose-response relationship between cigarette smoking and 4-ABP-Hb adduct formation. Furthermore, it has been described earlier that saturation of 4-ABP-Hb adduct formation occurs at high exposure levels. It was postulated that this saturation effect might be explained by different metabolic processes (8).

The hypothesis of the current study therefore is that genetic polymorphisms in genes coding for enzymes with a putative effect on 4-ABP-Hb metabolism, e.g. cytochrome P450s other than CYP1A2, explain the variations in 4-ABP-Hb adduct formation in response to cigarette smoking. To identify the crucial mediators in 4-ABP metabolism in humans, the impact of 12 Single Nucleotide Polymorphisms (SNPs) in 7 genes, including *CYP1A2*, involved in biotransformation processes, on 4-ABP-Hb ad-

Chapter 4

duct levels was examined in smokers. Subsequently, their effect on the dose-response relationship between the number of cigarettes smoked per day and 4-ABP-Hb adduct levels was investigated.

Materials and Methods

Study Population

The study population consisted of 33 female and 24 male cigarette smokers, with an average age of 43 ± 9 years. Mean smoking dose was 26 ± 9 cigarettes per day (range: 5-50 cigarettes per day). This study was conducted in smokers, however, only to adjust for background 4-APB-Hb adduct levels, 10 non-smoking volunteers, matched on gender and age, were included. Blood samples were donated for biomarker analysis. Inform consent was obtained from all individuals. Sample size of the study population was large enough to obtain more than 80% power, using an $\alpha=0.05$, $\delta=100$ pg/g and $\sigma=75$ pg/g, for all genotypes. The study was approved by the medical ethical commission, and inform consent was obtained from all individuals.

Isolation of Hb

Peripheral blood erythrocytes were isolated after venipuncture into EDTA tubes according to the method described by Bryant et al. (13), with minor modifications, as described earlier (8).

Analysis of 4-ABP-Hb levels

4-ABP-Hb levels were determined by GC-MS as described previously (8). Briefly, erythrocytes were lysed with de-ionized water, centrifuged, and Hb-containing solutions were dialyzed against de-ionized water for 48 hours at 4°C. The Hb concentration was determined using Drabkin's method (Sigma kit 525a). 4-ABP was released from the Hb by addition of 10 N NaOH. Subsequently, 4-ABP was extracted with hexane and derivatized by the addition of triethylamine and pentafluoropropionic anhydride. After evaporation of the hexane, the residue was re-dissolved in dichloromethane, and analyzed by GC-MS (13). For samples in which no 4-ABP-Hb was detectable, adduct levels were arbitrarily set at 1 pg/g Hb.

DNA Isolation

Blood was sampled by venapuncture into EDTA-tubes and lymphocytes were subsequently isolated by centrifugation on lymphoprep according to Bøyum (14), washed with PBS and stored at -20°C until DNA isolation. Subsequent DNA isolation was performed using standard phenol extraction procedures.

Selection of Polymorphisms

All functional polymorphisms that were included in this study, were selected based on their expected involvement in 4-ABP metabolism and/or 4-ABP-Hb adduct formation (Table 1). Main pathways of interest were Phase 1 (CYP1A1, 1A2, 1B1 and 3A4 and Ah-receptor) and Phase 2 (NAT2 and GSTM1) biotransformation reactions. Sequences and frequencies of occurrence of the SNPs in a Caucasian population were obtained from the Cancer SNP 500 database (<http://snp500cancer.nci.nih.gov>).

Multiplex PCR and Genotyping

DNA sequences containing the SNP of interest were amplified by multiplex PCR and subsequently genotyped by Single Base Extension (SBE) as described earlier (15). Briefly, PCR was carried out in a 10µL reaction volume containing water (Acros), 10x PCR buffer (Invitrogen, Breda, The Netherlands), 0,2 mM/L dNTP's (Invitrogen), 0,5 mM MgCl₂ (Invitrogen), 1,25 U Platinum[®] Taq Polymerase (Invitrogen), 40 ng template DNA and 0,2 µmol/L PCR primer. After a denaturation period of 3' at 95°C, amplification was accomplished in 30 cycles of 94°C (30sec), 60°C (30sec) and 72°C (30 sec). Final extension was done for 5 min. at 72°C.

Subsequent genotyping was done using the SNaPshot method (Applied Biosystems, Nieuwekerk a.d. IJssel, the Netherlands), and SBE products were analyzed on an ABI Prism 3100 genetic analyzer using Genescan Analysis software (version 3.7, (15)). Both PCR and SBE primers were designed using primer 3 ([http://www-genome.wi.mit.edu/cgi-bin/primer3_www.cgi](http://www-genome.wi.mit.edu/cgi-bin/primer3/www.cgi)) and net primer (<http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html>) software (Table 1).

Statistics

Linear regression analysis was conducted to investigate the relationship between numbers of cigarettes smoked per day and 4-ABP-Hb adduct levels. To explain inter-individual variations in the relationship between cigarette smoking and 4-ABP-Hb adduct levels, a slightly adjusted method described previously was used (16). Briefly, the population was divided into low and high responders on basis of the regression line (below and above the regression line, respectively). Subsequently, binary logistic regression analysis was conducted for classification of subjects into one of these groups, using gender, age, number of cigarettes smoked per day and all polymorphisms as independent variables. Interactions between significant variables were subsequently investigated as new variables by binary logistic regression. Non-parametric analyses (Mann-Whitney-U for 2 groups and Jonckheere-Terpstra for more than 2 groups) were conducted to test the effect of each variable separately on 4-ABP-Hb adduct levels.

Table 1. Selection of SNPs included in the study, together with their PCR and SBE primers.

Polymorphism	Nucleotide Change	Amino Acid Change	PCR Primers	Product (bp)	SBE Primer*	Frequency** (%)	
CYP1A2	*1F	A>C -154	None	F5'-GAGGCTCCTTTCCAGCTCTC-3' R5'-CTCCCAGCTGGATACCAGA-3'	106	5'-AACTGACTAAACTAGGTGCCACTCAAAGGGTGAGCTC TGTGGGC-3'	51/42/7
CYP1A1	m1	T>C 3801	None	F5'-TGCAGCTGGTACCATTTTGT-3' R5'-AAAGCTGTGGTGGCGTGT-3'	114	5'-AACTGATCATTTTGTTCCTACTGTAACCTCCACCTCC-3'	91/9/0
	m2	A>G 2455	I462V	F5'-CTGATGGTGCTATCGACAAG-3' R5'-CAGGATAGCCAGGAAGAGAA-3'	123	5'-AAAGACCTCCCAGCGGGCAA-3'	96/4/0
CYP1B1	*3	C>G 4326	V432L	F5'-CCCAAGGACACTGTGGTTTT-3' R5'-AGAAAGTTCTTCGCCAATGC-3'	186	5'-AACAAGTTCTCCGGGTTAGGCCACTTCA-3'	26/56/18
	*4	A>G 4390	N453S	F5'-CCCAAGGACACTGTGGTTTT-3' R5'-AGAAAGTTCTTCGCCAATGC-3'	186	5'-AACTGTTTCATCACTCTGCTGGTCAGGTCCTTG-3'	72/28/0
CYP3A4	*1B	A>G -391	None	F5'-GCACACTCCAGGCATAGGTA-3' R5'-GGAGCCATTGGCATAAAATC-3'	137	5'-AACTGACTAAACTAGGACTTACGGAGGACAGCCATAG AGACAAGGGCA-3'	96/4/0
GSTM1	*0	C/0	Del	F5'-CAGAGTTTCTGGGGAAGCGG-3' R5'-CTTGGGCTCAAATATACGGTGG-3'	191	5'-AACTGACTAAACTAGGTGCCACGCGAGGAGTGATATG GGGAATGAGAT-3'	59/41
NAT2	*5	T>C 341	I114T	F5'-CAAATACAGCACTGGCATGG-3' R5'-GGCTGATCCTTCCCAGAAAT-3'	135	5'-AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTAT TCACCTTCTCCTGCAGGTGACCA-3'	29/55/16
	*6	G>A 590	R197Q	F5'-CCTGCCAAAGAAGAAACACC-3' R5'-GGGTCTGCAAGGAACAAAAT-3'	143	5'-CCTACCAAAAAATATACTTATTTACGCTTGAACCTC-3'	51/42/7
	*7	G>A 857	G286E	F5'-TCCTTGGGGAGAAATCTCGT-3' R5'-GGGTGATACATACACAAGGGTTT-3'	92	5'-AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGA CAGCCCTCGTGCCCAAACCTGGTGATG-3'	91/9
AHR	*1	A>G 501	R554K	F5'-CCCAGGGCTCTTTCAAGATA-3' R5'-AGCCAAGGACTGTTGCTGTT-3'	238	5'-ATCTGACTAAACTAGGTGCCTTTCTGAAAAATTTTCA' TTCTGCATGTGT-3	86/14/0
	*6	A>C 517	3'UTR	F5'-CACATTATTCTGGGCACCAC-3' R5'-AAAAAGCACCAAACCATTTATTTT-3'	244	5'-AACTGACTGCGCTAGGTTTGAATTATTTCTGTACAGT AAAAATAAAATACTTTG-3'	93/7/0

*Neutral nonbinding tails are shown in Italics (see Knaapen et al. (15) for more detailed information)
**Frequencies are presented as wildtype/heterozygous/double variants. All frequencies are in concordance with those expected from the literature (see text).

To detail effects of significant predictor(s) on dose-response relationships, moving average analysis was conducted. Using this statistical approach, the variation around the regression line is “smoothened”, revealing more clearly the underlying trend of the dose-response relationship, and how it is modulated by the different genotypes. Therefore, mean levels of cigarettes per day and of 4-ABP-Hb adducts are computed for each of the genotypes comprising 10 subsequent samples, sorted for the amount of cigarettes smoked per day. For each new variable, this average is “moved” by 1 value, i.e. $(x_0 + \dots + x_n)/n$, $(x_0 + 1 + \dots + x_{n+1})/n$, ..., $(x_0 + i + \dots + x_{n+i})/n$ (where $n=10$ (# of values in mean), and $i=10, 33, 16$ for wild types, heterozygous and homozygous individuals respectively (# of steps the average is “moved”, i.e. the number of new variables). A new “smoothened” dose-response relationship can then be plotted between these new variables. No further statistics are allowed using these variables, because the variables are no longer independent.

All statistical analyses were performed using SPSS 13.0 for Windows, and $P < 0.05$ was considered statistically significant. Results are expressed as mean $\pm$ standard deviation (SD).

Results

Overall Analysis of 4-ABP-Hb adduct levels

The mean 4-ABP-Hb adduct level in smokers was 182 ± 94 pg/g Hb. This level was significantly higher than in non-smokers (28 ± 16 pg/g Hb, $P < 0.001$). A positive relationship was observed between the number of cigarettes smoked per day and 4-ABP-Hb adduct levels in smokers ($R^2 = 0.46$, $P < 0.001$).

Effect of each SNP separately on 4-ABP-Hb adduct levels

Overall, 4-ABP-Hb adduct levels were significantly lower ($P = 0.005$) in individuals carrying the low activity *CYP1B1* ⁴³²Val polymorphism (130 ± 71 pg/g Hb) as compared to homozygous (226 ± 72 pg/g Hb) and heterozygous *CYP1B1* ⁴³²Leu individuals (192 ± 100 pg/g Hb). None of the other selected SNPs were found to have an effect on 4-ABP-Hb adduct levels.

Effect of multiple SNPs, gender, age and cigarette smoking on 4-ABP-Hb adduct levels

First, the population was divided into low and high responders based on the regression line. Using binary logistic regression, only a significant effect for *CYP1B1* ⁴³²Leu⁴³²Val on the classification of each individual into one of the two subgroups was observed ($P = 0.020$). None of the other investigated polymorphisms were found to have a significant impact on this classification. The effect of *CYP1B1* ⁴³²Leu⁴³²Val was observed particularly in combination with daily smoking dose ($P = 0.003$). To investigate this interac-

tion in detail, subjects were divided into two equal large groups of moderate and high smokers. Comparing 4-ABP-Hb adduct levels between different *CYP1B1* *Leu*⁴³²*Val* genotypes within the subgroups of moderate or high smokers revealed that 4-ABP-Hb adduct levels were significantly lower in moderate smokers carrying the *CYP1B1* *432Val* variant (98 ± 69 pg/g Hb) as compared to wild type (*432Leu*) individuals (225 ± 79 pg/g Hb, $P=0.008$; Figure 1).

Figure 1. Combined effect of the *CYP1B1* *Leu*⁴³²*Val* polymorphism and cigarette smoking dose on 4-ABP-Hb adduct levels. 4-ABP-Hb adduct levels were higher in heavy smokers compared to low-smokers in individuals homozygously carrying the *CYP1B1* *432Val* variant allele. For the *CYP1B1* *432Leu* wild-type and heterozygous individuals, no such effect was observed. Moreover, 4-ABP-Hb adduct levels were higher in these individuals compared to the *CYP1B1* *432Val* variants in low-smoking individuals, whereas this effect was less clear in heavy smokers. For visualization of the results, connection lines were used between the data-points.
 ● *CYP1B1* *432Leu* wild types, ○ *CYP1B1* *Leu*⁴³²*Val* heterozygous individuals,
 ● *CYP1B1* *432Val* homozygous variants.

In heavy smokers, no such difference in adduct levels between homozygous carriers of the *CYP1B1* *432Val* polymorphism and *CYP1B1* *432Leu* wild-type individuals was observed (158 ± 65 pg/g Hb versus 229 ± 55 pg/g Hb respectively, $P=0.245$, Figure 1). Upon combining these results, no difference in 4-ABP-Hb adduct levels exists between moderately and heavily smoking wild types (225 ± 79 pg/g Hb versus 229 ± 55 pg/g Hb, $P=1.0$) and between moderately and heavily smoking heterozygous individuals (191 ± 92 pg/g Hb versus 194 ± 120 pg/g Hb, $P=0.80$), as is shown in Figure 1. In contrast, in carriers of the *CYP1B1* *432Val* variant, higher adduct levels (158 ± 65 pg/g Hb) are observed in heavy smokers as compared to moderate smoking individuals (98 ± 69 pg/g Hb, Figure 1), although not statistically significant.

Effect of CYP1B1 Leu⁴³²Val on the Dose-Response Relationship

It may be concluded that in wild type and heterozygous *CYP1B1*⁴³²Leu individuals a saturation effect seems to occur in 4-ABP-Hb adduct formation at high dose exposures (Figure 1). To investigate this saturation effect in more detail, moving average analysis was used to visualize the dose-response relationship. A plateau appears to be reached in wild type and heterozygous *CYP1B1*⁴³²Leu individuals at >15 cigarettes per day. This plateau was not observed in homozygous carriers of the *CYP1B1*⁴³²Val variant allele; in this subgroup a distinctly linear dose-response relationship between cigarette smoke dose and 4-ABP-Hb adduct formation was observed, which was less steep as compared to the linear part of the relationship in the *CYP1B1*⁴³²Leu wild type- and heterozygous individuals (Figure 2). In the latter, a positive linear relationship between mean 4-ABP-Hb adduct levels and dose was found below the cut-off of 15 cigarettes per day.

Figure 2. *Effect of CYP1B1 Leu⁴³²Val on the dose-response relationship between mean cigarette dose and 4-ABP-Hb adduct levels visualized by moving averages. For the individuals having the CYP1B1⁴³²Val variant allele, a dose-response relationship between cigarette smoking and 4-ABP-Hb adduct levels was observed. A relationship between smoking and 4-ABP-Hb adduct levels was also found in low-smoking wild-type and heterozygous CYP1B1⁴³²Leu individuals, with a steeper slope compared to the CYP1B1⁴³²Val variants, whereas in heavy smokers a steady state situation occurs in these individuals. Mean smoking dose and mean adduct levels were computed by moving average.*
 ● *CYP1B1*⁴³²Leu wild-type and heterozygous individuals, ▲ *CYP1B1*⁴³²Val homozygous variants.

Discussion

To evaluate the impact of low-penetrance gene polymorphisms on cigarette smoking-related 4-ABP-Hb adduct levels, 57 smoking individuals were genotyped for 12 phenotypically effective SNPs in 7 genes with a hypothesized involvement in the biotransformation of 4-ABP. Only *CYP1B1* *Leu*⁴³²*Val* was found to have a statistically significant effect on 4-ABP-Hb levels in smokers. Moreover, 4-ABP-Hb adduct levels seem to level off at high doses in *CYP1B1* *Leu*⁴³² wild type and heterozygous individuals, whereas for homozygous carriers of the *CYP1B1* *Val*⁴³² variants, no such plateau was reached.

The arylamine 4-ABP is one of the major carcinogenic components of cigarette smoke. It has been previously shown that 4-ABP-Hb adduct levels are higher in smokers as compared to non-smokers, and that levels increase with increasing numbers of cigarettes smoked per day (6,13,17). Also in the current study, higher levels of 4-ABP-Hb adducts were found in smokers compared to non-smokers, and overall, a positive relationship between the dose of cigarette smoke exposure and 4-ABP-Hb adduct levels was demonstrated. Within this relationship, variations in 4-ABP-Hb adduct levels were observed ($R^2=0.46$) and it has been postulated that these inter-individual differences may be partly explained by genetic polymorphisms in low penetrance genes (18-20). From the set of polymorphisms studied in the current study, an effect for *CYP1B1* *Leu*⁴³²*Val* on 4-ABP-Hb adduct levels was found; higher adduct levels were observed in homozygous wild-type and heterozygous individuals as compared to homozygous variant individuals, which suggests a protective effect of the *CYP1B1* *Val*⁴³² variant allele. Thus, carriers of one or more wildtype *Leu*⁴³² alleles have a higher metabolic activity towards 4-ABP than the *Val*⁴³² double variants (21), which may lead to higher levels of the reactive *N*-hydroxy-ABP intermediate and hence, to higher 4-ABP-Hb adduct levels.

This role of CYP1B1 in 4-ABP metabolism has not yet been described. In addition, a putative role for CYP1B1 in different types of cancer has been suggested, however no evidence for a role in bladder cancer can be found (22-24). It has generally been assumed from *in vitro* studies, that CYP1A2 is the most important enzyme in the bioactivation of 4-ABP (4). In contrast, several recent reports suggest that CYP1A2 may not be the primary enzyme responsible for metabolizing 4-ABP and/or other enzymes may be involved (11,12), including CYP1A1 and CYP1B1, which have been described (at least for CYP1A1) to be able to metabolize 4-ABP to some degree (11,25). Polymorphisms in these genes cause altered functions of the enzyme, and the functional analysis of these low penetrance polymorphisms can therefore provide information about the involvement in metabolic processes. However, polymorphisms in CYP1A1 have a low occurrence in a caucasian population (<5%), as has been confirmed in this study (frequency of *CYP1A1* variants was 1%). Also, no effect of the *CYP1A2**1F polymorphism on 4-ABP-Hb adduct levels was found. *CYP1A2**1F is a functional

polymorphism, leading to a higher inducibility, which is sufficiently present in the Caucasian population (41% in heterozygous and 7% in double variant individuals) to be able to detect an effect if it would be present. This negative outcome supports the results of a study in mice by Kimura et al., that CYP1A2 may not be the primary enzyme for 4-ABP metabolism (12). Several other studies have assessed the possible effect of different genotypes, predominantly polymorphisms in phase II metabolizing enzymes such as glutathione *S*-transferases (GSTs) and *N*-acetyltransferases (NATs), on 4-ABP-Hb adduct formation in humans, with contradicting results (8,9,26-28). In the current study, no significant effect of polymorphisms in GST and NAT2 enzymes on 4-ABP-Hb adduct formation was found.

Nine SNPs have been reported in the human *CYP1B1* gene, of which five cause amino-acid substitutions (29). Seven variant alleles have been identified until now, namely *CYP1B1**1 - *7 (29,30). The role of these *CYP1B1* allelic variants in the biotransformation of various substances has been investigated, with contradictory results (21,31-33), and their association with the occurrence of different forms of cancers has been analyzed (23,34-36). Among these *CYP1B1* alleles, *CYP1B1* *Leu*⁴³²*Val* is the most prominent. An important role has been described for the metabolism of benzo(a)pyrene to *trans*-7,8-dihydrodiol, the metabolite just prior to the one that can interact with DNA or proteins to form adducts (32,37,38). Lower oxidation rates are observed in the variant *432Val* isoform of the *CYP1B1* *Leu*⁴³²*Val* allele (21,31). Our results are consistent with these findings, and contribute to the existing literature that the *CYP1B1* *Leu*⁴³²*Val* polymorphism indeed leads to a decrease in enzyme activity regarding the oxidation of various pro-carcinogens, such as 4-ABP, and hence lower adduct levels in smokers carrying the double *432Val* variant allele.

As was shown in Figure 2, the observed effect of *CYP1B1* *Leu*⁴³²*Val* on 4-ABP-Hb adduct levels occurs in interaction with smoking behaviour. The described differences in the level of 4-ABP-Hb adducts between the 3 *CYP1B1* *Leu*⁴³²*Val* genotypes (wild type, heterozygous or variants) were found to be predominant in moderate smokers, but not in heavy smokers. This can be explained by the fact that 4-ABP-Hb adduct levels levels off above a dose of 15 cigarettes per day in wildtype and heterozygous individuals, whereas only in *CYP1B1* *432Val* variants higher 4-ABP-Hb adduct levels were found at high smoking dose as compared to moderate smokers carrying the *CYP1B1* *432Val* variant alleles. Apparently, 4-ABP-Hb adduct formation is less effective at high doses, as the levels of 4-ABP-Hb adducts are independent from the number of cigarettes smoked per day at such high dose (8).

In conclusion, it was postulated that the described variations in the dose-response relationship between cigarette smoking and 4-ABP-Hb adduct levels is to be explained by different metabolic routes (8). Our results confirm this hypothesis, and demonstrate that at least part of this variation can indeed be explained by the *CYP1B1* *Leu*⁴³²*Val* polymorphism, as the double *432Val* variant allele has a protective effect on 4-ABP-Hb adduct levels in response to cigarette smoking. The level of 4-ABP-Hb adducts in smokers is associated with the number of cigarettes smoked per day, with a saturation

Chapter 4

effect at high doses of exposures limited to individuals having the high activity *CYP1B1*⁴³²*Leu* genotype, implying that CYP1B1 plays a significant role in the metabolism of 4-ABP. In addition, these data show that studying the impact of genetic polymorphisms on dose-response relationships may provide valuable information on the involvement of certain enzymes in toxicologically relevant metabolic pathways.

References

1. Smith, C.J., Perfetti, T.A., Garg, R., et al. (2003) IARC carcinogens reported in cigarette mainstream smoke and their calculated log P values. *Food Chem Toxicol*, 41, 807-17.
2. Vineis, P. (1992) Epidemiological models of carcinogenesis: the example of bladder cancer. *Cancer Epidemiol Biomarkers Prev*, 1, 149-53.
3. Lower, G.M., Jr., Nilsson, T., Nelson, C.E., et al. (1979) N-acetyltransferase phenotype and risk in urinary bladder cancer: approaches in molecular epidemiology. Preliminary results in Sweden and Denmark. *Environ Health Perspect*, 29, 71-9.
4. Eaton, D.L., Gallagher, E.P., Bammler, T.K., et al. (1995) Role of cytochrome P4501A2 in chemical carcinogenesis: implications for human variability in expression and enzyme activity. *Pharmacogenetics*, 5, 259-74.
5. Guengerich, F.P., Humphreys, W.G., Yun, C.H., et al. (1995) Mechanisms of cytochrome P450 1A2-mediated formation of N-hydroxy arylamines and heterocyclic amines and their reaction with guanyl residues. *Princess Takamatsu Symp*, 23, 78-84.
6. Yu, M.C., Ross, R.K., Chan, K.K., et al. (1995) Glutathione S-transferase M1 genotype affects aminobiphenyl-hemoglobin adduct levels in white, black and Asian smokers and nonsmokers. *Cancer Epidemiol Biomarkers Prev*, 4, 861-4.
7. Bartsch, H., Malaveille, C., Friesen, M., et al. (1993) Black (air-cured) and blond (flue-cured) tobacco cancer risk. IV: Molecular dosimetry studies implicate aromatic amines as bladder carcinogens. *Eur J Cancer*, 29A, 1199-207.
8. Dallinga, J.W., Pachon, D.M., Wijnhoven, S.W., et al. (1998) The use of 4-aminobiphenyl hemoglobin adducts and aromatic DNA adducts in lymphocytes of smokers as biomarkers of exposure. *Cancer Epidemiol Biomarkers Prev*, 7, 571-7.
9. Godschalk, R.W., Dallinga, J.W., Wikman, H., et al. (2001) Modulation of DNA and protein adducts in smokers by genetic polymorphisms in GSTM1, GSTT1, NAT1 and NAT2. *Pharmacogenetics*, 11, 389-98.
10. Sim, S.C. and Ingelman-Sundberg, M. (2006) The human cytochrome P450 Allele Nomenclature Committee Web site: submission criteria, procedures, and objectives. *Methods Mol Biol*, 320, 183-91.
11. Tsuneoka, Y., Dalton, T.P., Miller, M.L., et al. (2003) 4-aminobiphenyl-induced liver and urinary bladder DNA adduct formation in Cyp1a2(-/-) and Cyp1a2(+/+) mice. *J Natl Cancer Inst*, 95, 1227-37.
12. Kimura, S., Kawabe, M., Ward, J.M., et al. (1999) CYP1A2 is not the primary enzyme responsible for 4-aminobiphenyl-induced hepatocarcinogenesis in mice. *Carcinogenesis*, 20, 1825-30.
13. Bryant, M.S., Skipper, P.L., Tannenbaum, S.R., et al. (1987) Hemoglobin adducts of 4-aminobiphenyl in smokers and nonsmokers. *Cancer Res*, 47, 602-8.
14. Boyum, A. (1976) Isolation of lymphocytes, granulocytes and macrophages. *Scand J Immunol, Suppl* 5, 9-15.
15. Knaapen, A.M., Ketelslegers, H.B., Gottschalk, R.W., et al. (2004) Simultaneous genotyping of nine polymorphisms in xenobiotic-metabolizing enzymes by multiplex PCR amplification and single base extension. *Clin Chem*, 50, 1664-8.

16. Ketelslegers, H.B., Gottschalk, R.W., Godschalk, R.W., et al. (2006) Interindividual variations in DNA adduct levels assessed by analysis of multiple genetic polymorphisms in smokers. *Cancer Epidemiol Biomarkers Prev*, 15, 624-9.
17. Bryant, M.S., Vineis, P., Skipper, P.L., et al. (1988) Hemoglobin adducts of aromatic amines: associations with smoking status and type of tobacco. *Proc Natl Acad Sci U S A*, 85, 9788-91.
18. Georgiadis, P., Demopoulos, N.A., Topinka, J., et al. (2004) Impact of phase I or phase II enzyme polymorphisms on lymphocyte DNA adducts in subjects exposed to urban air pollution and environmental tobacco smoke. *Toxicol Lett*, 149, 269-80.
19. Bartsch, H. and Hietanen, E. (1996) The role of individual susceptibility in cancer burden related to environmental exposure. *Environ Health Perspect*, 104 Suppl 3, 569-77.
20. Vineis, P. (2004) Individual susceptibility to carcinogens. *Oncogene*, 23, 6477-83.
21. Shimada, T., Watanabe, J., Kawajiri, K., et al. (1999) Catalytic properties of polymorphic human cytochrome P450 1B1 variants. *Carcinogenesis*, 20, 1607-13.
22. Hung, R.J., Boffetta, P., Brennan, P., et al. (2004) GST, NAT, SULT1A1, CYP1B1 genetic polymorphisms, interactions with environmental exposures and bladder cancer risk in a high-risk population. *Int J Cancer*, 110, 598-604.
23. Roos, P.H. and Bolt, H.M. (2005) Cytochrome P450 interactions in human cancers: new aspects considering CYP1B1. *Expert Opin Drug Metab Toxicol*, 1, 187-202.
24. Vineis, P., Veglia, F., Garte, S., et al. (2007) Genetic susceptibility according to three metabolic pathways in cancers of the lung and bladder and in myeloid leukemias in nonsmokers. *Ann Oncol*.
25. Turesky, R.J., Constable, A., Richoz, J., et al. (1998) Activation of heterocyclic aromatic amines by rat and human liver microsomes and by purified rat and human cytochrome P450 1A2. *Chem Res Toxicol*, 11, 925-36.
26. Yu, M.C., Skipper, P.L., Taghizadeh, K., et al. (1994) Acetylator phenotype, aminobiphenyl-hemoglobin adduct levels, and bladder cancer risk in white, black, and Asian men in Los Angeles, California. *J Natl Cancer Inst*, 86, 712-6.
27. Vineis, P., Caporaso, N., Tannenbaum, S.R., et al. (1990) Acetylation phenotype, carcinogen-hemoglobin adducts, and cigarette smoking. *Cancer Res*, 50, 3002-4.
28. Bartsch, H., Caporaso, N., Coda, M., et al. (1990) Carcinogen hemoglobin adducts, urinary mutagenicity, and metabolic phenotype in active and passive cigarette smokers. *J Natl Cancer Inst*, 82, 1826-31.
29. Aklillu, E., Oscarson, M., Hiderstrand, M., et al. (2002) Functional analysis of six different polymorphic CYP1B1 enzyme variants found in an Ethiopian population. *Mol Pharmacol*, 61, 586-94.
30. Ingelman-Sundberg, M. (2004) Human drug metabolising cytochrome P450 enzymes: properties and polymorphisms. *Naunyn Schmiedeberg's Arch Pharmacol*, 369, 89-104.
31. Shimada, T., Watanabe, J., Inoue, K., et al. (2001) Specificity of 17beta-oestradiol and benzo[a]pyrene oxidation by polymorphic human cytochrome P4501B1 variants substituted at residues 48, 119 and 432. *Xenobiotica*, 31, 163-76.
32. Shimada, T., Oda, Y., Gillam, E.M., et al. (2001) Metabolic activation of polycyclic aromatic hydrocarbons and other procarcinogens by cytochromes P450 1A1 and P450 1B1 allelic variants and other human cytochromes P450 in *Salmonella typhimurium* NM2009. *Drug Metab Dispos*, 29, 1176-82.
33. Abnet, C.C., Fagundes, R.B., Strickland, P.T., et al. (2006) The influence of genetic polymorphisms in Ahr, CYP1A1, CYP1A2, CYP1B1, GST M1, GST T1, and UGT1A1 on urine

- 1-hydroxypyrene glucuronide concentrations in healthy subjects from Rio Grande do Sul, Brazil. Carcinogenesis.
34. Watanabe, J., Shimada, T., Gillam, E.M., et al. (2000) Association of CYP1B1 genetic polymorphism with incidence to breast and lung cancer. Pharmacogenetics, 10, 25-33.
 35. Cecchin, E., Russo, A., Campagnutta, E., et al. (2004) Lack of association of CYP1 B1*3 polymorphism and ovarian cancer in a Caucasian population. Int J Biol Markers, 19, 160-3.
 36. Ko, Y., Abel, J., Harth, V., et al. (2001) Association of CYP1B1 codon 432 mutant allele in head and neck squamous cell cancer is reflected by somatic mutations of p53 in tumor tissue. Cancer Res, 61, 4398-404.
 37. Aklillu, E., Ovrebo, S., Botnen, I.V., et al. (2005) Characterization of common CYP1B1 variants with different capacity for benzo[a]pyrene-7,8-dihydrodiol epoxide formation from benzo[a]pyrene. Cancer Res, 65, 5105-11.
 38. Shimada, T., Gillam, E.M., Oda, Y., et al. (1999) Metabolism of benzo[a]pyrene to trans-7,8-dihydroxy-7, 8-dihydrobenzo[a]pyrene by recombinant human cytochrome P450 1B1 and purified liver epoxide hydrolase. Chem Res Toxicol, 12, 623-9.

Chapter 5

Multiplex genotyping as a biomarker for susceptibility to identify subgroups at increased risk in carcinogenic exposure-effect relationships in the FLEHS biomonitoring study

Hans B. Ketelslegers¹, Ralph W.H. Gottschalk¹, Gudrun Koppen², Greet Schoeters², Willy Bayens³, Joost H.M. van Delft¹ and Jos C.S. Kleinjans¹

¹Department of Health Risk Analysis and Toxicology, Maastricht University, Maastricht, the Netherlands

²Flemish Institute for Technological Research, Mol, Belgium

³Analytical and Environmental Chemistry Department, ANCH, Vrije Universiteit Brussel, Brussels, Belgium

Abstract

Cancer may result from interactions between genetic and environmental factors, and certain subgroups in the general population are at increased risk due to relatively higher susceptibility. The current study, part of a large biomonitoring study conducted in Flanders in 2002-2006 (FLEHS), aims to determine these more susceptible subpopulations regarding different environmental exposures, based on genotypic differences between individuals. Therefore, a random selection of 429 adolescents (aged 14-16 years) and 361 adults (50-65 years old) was genotyped for 36 polymorphisms in 23 genes involved in xenobiotic metabolism, DNA repair and oxidative stress. In both age groups, exposure-effect relationships between exposure to several organochloride substances (PCB, HCB, DDE), heavy metals (Cadmium, Lead), and urinary metabolites (1-OHP, tt-MA), and biomarkers of adverse effect (DNA damage levels, represented by Comet and 8-OHdG) were investigated. Additionally, in the study among adults, the relationship of these exposures with several tumour markers (PSA, CEA and p53) was tested. A total of 8 significant exposure-effect relationships between the investigated markers for exposure and effect were observed. Subsequently, the effect of the genotype, sex and cigarette smoking on these exposure-effect relationships was assessed. Therefore, individuals were divided into two subgroups: above (indicating relatively higher susceptibility) or below the regression line. Three novel associations were found, including an effect of the genotype on the specific relationships. First of all, a positive relationship was observed between ethylbenzene exposure and 8-OHdG DNA damage levels. More susceptible subgroups within this relationship were defined by *GSTT1* and *CAT* genotypes. Second, Cadmium exposure correlated positively with serum p53 levels. This relationship was affected by *GSTM1* and *BRCA2* genotypes. Finally, higher CEA levels were found upon Lead exposure. An effect on this relationship was found for the *BRCA2* genotype and cigarette smoking. Overall, however, the effects due to genotype were relatively marginal. In conclusion, this study shows that at least part of the inter-individual differences in the specific exposure-effect relationships between biomarkers for carcinogen exposure and genotoxic effect could be explained by genotypic differences, enabling the identification of more susceptible subgroups for environmental cancer risks, and suggesting that it is of social importance to continue the responsible care debate about inter-individual differences in susceptibility for developing harmful effects in response to environmental exposures.

Introduction

Over several decades now, molecular epidemiology research has provided evidence that the development of cancer may result from environmental exposures to carcinogenic compounds interacting with both genetic and acquired susceptibility factors (1,2). It has been estimated that the incidence of cancer can be reduced by as much as 85% upon preventing exposure to these environmental components (3). Therefore, continued efforts are needed to investigate the effect of hazardous environmental exposures on environmental cancer risk, and eventually, to translate this knowledge into preventive policy measures. Molecular epidemiology approaches have the advantage of applying methodologies which enable investigating the preclinical effects of exposure by using biological markers (biomarkers) present in human cells, tissues and body fluids (1). These biomarkers comprise measurements of the so-called internal dose of an environmental carcinogen (e.g. urinary metabolites resulting from carcinogen metabolism), biologically effective dose (e.g. DNA damage reflecting certain genotoxic exposures) and early biological effects (e.g. deleterious mutations). The EU Environmental Health Action plan stimulates biomarkers studies among relevant human populations (<http://www.eu-humanbiomonitoring.org/sub/back/ehap.htm>).

The Flanders Environment and Health Study (FLEHS) is a large-scale biomonitoring study conducted in Flanders in the period 2002-2006, and assigned by the Flemish Government (www.milieu-en-gezondheid.be). Its aim is to obtain a better perspective on health risks associated with exposure to harmful environmental compounds among inhabitants of different regions of Flanders. The outcomes of this research project may eventually lead to new policy measures regarding environmental exposures to hazardous compounds, e.g. from traffic/industrial exhaust.

Conventional risk assessment methods may underestimate the individual's risk of exposures to environmental carcinogens because of the default assumption that all individuals within a certain population possess equal susceptibility to a specific carcinogen dose (3). Although the established risk assessment procedure for setting exposure limits to environmental carcinogens is considered sufficiently safe due to application of the relatively rigid linear, non-threshold, extrapolation model and the inclusion additional safety factors, environmental health policy making may benefit from more information on the degree of inter-individual variations in environmental cancer risk.

Because genetic polymorphisms may influence every step in the continuum from endogenous exposure to a specific compound to the eventual health effect, investigating genetic variations in monitored populations may provide information about the inter-individual differences in genetic susceptibility for environmental diseases such as cancer. In the FLEHS, individuals were monitored for exposure to organochlorides, heavy metals, Polycyclic Aromatic Hydrocarbons (PAH) and benzene, to investigate their potentially hazardous effect on human health. Organochlorides and heavy metals are known to produce free radicals and thereby lead to oxidative stress (4-8). Moreover,

Chapter 5

the former, as well as the carcinogenic Polycyclic Aromatic Hydrocarbons (PAH) and benzene, are metabolically transformed to exert their carcinogenic effect. The highly reactive metabolites formed during this biotransformation process, are known to induce DNA damage. Moreover, these metabolic processes are also known to cause oxidative stress leading to (oxidative) DNA damage. If DNA repair is not sufficient, this DNA damage may lead to mutations in critical genes and eventually initiate or affect the carcinogenic process. In order to elucidate the effects of these exposures among the Flemish populations, DNA damage levels as measured by the Comet assay, 8-hydroxy-2'-deoxyguanosine (8-OHdG) and several tumor markers were used as biomarkers for effect.

Within the putative relationships between the exposure and effect markers, large inter-individual variations are expected. It is postulated that these differences in response can be explained by genetic variation in the involved pathways between individuals, in combination with smoking habits and gender. Therefore, the aim of the current study was to determine these more susceptible subpopulations regarding different environmental exposures. In this perspective, a random selection of 790 individuals from the FLESH study population was genotyped for 36 polymorphisms in 23 critical genes involved in biotransformation, oxidative stress and DNA repair, with the aim to investigate the role of these genotypes on various exposure-effect relationships.

Material and Methods

Study Population

Volunteers were all recruited upon advertisement by the FLESH Biomonitoring and Surveillance team. The campaign was approved by the ethical committee of the University of Antwerp, and every individual signed an informed consent prior to sample collection. Blood samples were donated by all individuals for analysis of exposure (internal dose) and effect biomarkers. Information on age, sex and smoking habits was obtained by questionnaire. Two age groups were defined, consisting of 429 adolescents and 361 adults. Both groups represent a selection from 4800 subjects who participated in the complete FLESH biomonitoring study, run in 2002 to 2006.

The adolescents were sampled between October 2003 and July 2004 and consisted of 256 males and 173 females, with an average age of 15 years (± 0.5 year). 54 individuals reported smoking 7 ± 7 cigarettes per day (range: 1-41).

Subsequently, 361 adults, aged 50-65 years at the time of inclusion, were sampled between September 2004 and June 2005. This study population comprised of 185 males and 176 females, of whom 65 individuals reported to smoke 17 ± 9 cigarettes per day (range: 1-40).

The participants were recruited in 8 different regions of Flanders. In order to achieve statistical power, the impact of genetic polymorphisms on exposure-effect relationships was assessed in the whole groups of all individuals of one age.

White Blood Cell and DNA Isolation

White blood cells were isolated by destroying red blood cells using lysis buffer containing 155 mM NH_4Cl , 10 mM KHCO_3 and 10 mM EDTA of 0°C (blood: buffer ratio is 1:2). After 2 rounds of i) incubation with lysis buffer on ice for 30 minutes and ii) centrifugation (10 minutes at 1000 rpm at 4°C), the cells were washed with 1 ml 150 mM NaCl and stored at -80°C prior to DNA isolation.

DNA isolation was performed in a 96 wells format using the Invisorb® Blood Midi HTS 96 Kit / C (Invitek, Berlin, Germany), and stored in 96 wells plates at -4°C until genotyping.

Selection of Polymorphisms

36 Single Nucleotide Polymorphisms in 23 genes involved in Biotransformation, DNA Repair and Oxidative Stress were analyzed in this study. Selection of these polymorphisms was based on their expected contribution to the formation and repair of DNA damage induced by environmental (carcinogenic) chemicals (Table 1). Moreover, they had to occur at a frequency in the population of >5% and should be likely to affect the enzyme activity. An exception has been made for polymorphisms in the phase I biotransformation gene *CYP1A1* (as their frequency of occurrence is lower than 5% in most populations), because of the prominent role of this enzyme in the bioactivation/detoxification of numerous xenobiotics.

All DNA sequences and allele frequencies were obtained from the SNP 500 cancer database (<http://snp500cancer.nci.nih.gov>).

Primer design and multiplex PCR

PCR and Single Base Extension (SBE) primers were designed as described previously. PCR was carried out in 4 different multiplex PCR reactions (Table 1) on a Tgradient 96 well Thermal cycler (Biometra, Goettingen, Germany), each in a 10 µl volume using 96 wells plates (described previously (9)). Values for T_m were optimized for every multiplex PCR reaction individually: 56°C for multiplex 1, 60°C for multiplex 2, 57°C for multiplex 3 and 60°C for multiplex 4 (Table 1). PCR products were subsequently incubated (37°C for 45 min) with 4 µl Exo-SAP-IT (Amersham, Roosendaal, The Netherlands) to digest contaminating dNTPs and PCR primers. Enzymes were deactivated at 75°C (15 min).

Table 1. Overview of the included SNPs and their phenotypic characteristics. Between brackets are sets of SNPs analysed together.

Polymorphism		db SNP ID		Effect on enzymatic function
Biotransformation				
CYP1A2	A>C-154	(1)	rs762551	Higher inducibility
CYP1A1	T>C3801	(4)	rs5030838	} Higher enzyme activity
	I462V	(4)	rs1048943	
	T461N	(4)	rs1799814	
CYP1B1	V432L	(4)	rs1056836	} Higher enzyme activity
	N453S	(4)	rs1800440	
CYP2E1	G>T-70	(4)	rs6413420	Higher enzyme activity
CYP3A4	A>G-391	(4)	rs2740574	Higher enzyme activity
GSTM1	Del	(1)		Deletion, no enzyme activity
GSTP1	I105V	(1)	rs947894	} Decreased enzyme activity
	A114V	(1)	rs1799811	
GSTT1	Del	(1)		Deletion, no enzyme activity
NAT2	I114T	(1)	rs1801280	} Decreased enzyme activity
	R197Q	(1)	rs1799930	
	G286E	(1)	rs1799931	
mEH	Y113H	(3)	rs1051740	Decreased enzyme activity
	H139R	(3)	rs2234922	Increased enzyme activity
MTHFR	A429E	(1)	rs1801131	} Decreased enzyme activity
	A222V	(1)	rs1801133	
COMT	V108M	(4)	rs4680	Decreased enzyme activity
DNA Repair				
XRCC1	R194W	(2)	rs1799782	Increased enzyme activity
	R280H	(2)	rs25489	Decreased enzyme activity
	Q399R	(2)	rs25487	Decreased enzyme activity
XRCC3	T241M	(2)	rs861539	Decreased enzyme activity
XPB	K751Q	(2)	rs1052559	} Decreased enzyme activity
	R156R	(2)	rs238406	
BRCA2	G>A-26 (5'UTR)	(2)	rs1799943	} Decreased enzyme activity
	N372H	(2)	rs144848	
APE1	D148G	(3)	rs3136820	Decreased enzyme activity
OGG1	S326C	(2)	rs1052133	Decreased enzyme activity
Oxidative Stress				
mnSOD	V16A	(3)	rs1799725	Decreased enzyme activity
CAT	C>T-262	(3)	rs1001179	Lower catalase activity
MPO	G>A-463	(4)	rs2333227	Higher enzyme activity
NQO1	P187S	(3)	rs1800566	} Reduced enzyme activity
	R139W	(3)	rs4986998	
GPX1	P198L	(3)	rs1050450	Lower enzyme activity

(1) 10-plex PCR, (2) 9-plex PCR, (3) 8-plex PCR, (4) 9-plex PCR
(See Material and Methods)

Genotyping

Genotyping was performed by Primer Extension using the Snapshot methodology (Applied Biosystems, Nieuwekerk a.d. IJssel, the Netherlands). 4 different multiplex SBE reactions were conducted on the corresponding multiplex PCR products using 96 wells plates. The protocol used for SBE was described previously (9). After genotyping, SBE products were mixed with 13µl deionized formamide containing 0,4µl Genescan™ 120 LIZ™ size standard, denatured at 95°C for 5 min, and thereafter analyzed on an ABI Prism® 3100 genetic analyzer using Genescan™ Analysis software (Version 3.7).

Statistical Analysis

First, the putative (confounding) effect of smoking was tested by comparing biomarker levels between smokers and non-smokers using a Mann Whitney U test. In case smoking significantly affected biomarker levels, subsequent analyses needed to be conducted among non-smokers, as primary focus was on the exposure to environmental toxic compounds, such as traffic/industrial exhaust. In case no effect of smoking was observed, further statistical analyses were conducted in the total population, cigarette smoking being regarded as an independent variable. The subsequent analysis is a variant on the statistical methodology described earlier (10), and summarized in Figure 1.

Figure 1. *Overview of the Statistical analysis. If smoking significantly affected the exposure marker level (Mann-Whitney U test), the analysis was conducted in non-smokers only. Exposure-effect relationships were analyzed by linear regression, and 2 subgroups were made based on the regression line. Subsequent classification of the individuals into one of these subgroups (low- or high responders) was conducted using binary logistic regression analysis.*

Shortly, after possible elimination of smokers, relationships between the investigated markers for exposure and effect (or internal and biologically effective dose) were investigated by linear regression analysis. For each individual the deviation from this relationship was determined, by dividing the level of response that was actually observed (O) by the level that was expected (E) according to the regression line. Using this approach, the population was divided into two subgroups: 0) so called low responders, i.e. all individuals below the regression line (thus an $O/E < 1$; protective) and 1) high responders, i.e. the more sensitive subjects, above the regression line ($O/E > 1$; more susceptible). Subsequently, to identify the most important factors for discrimination of the different subgroups, binary logistic regression analysis is conducted to classify each individual into one of the two subgroups based on the putative variables cigarettes per day (only if smokers were included), sex and all SNPs. The investigated genotypes were coded 2 in case of double variants, 0 for wild type and 1 for heterozygous individuals (10). Results are presented as mean $\pm$ SE.

Results

A total of 28,440 genotypes was generated in the overall study population. Statistical analyses assessing the impact of these genotypes on various exposure-effect relationships (Figure 1), were conducted in the adolescent and adult population separately, and obtained results will therefore be separately described for each subpopulation.

Adolescent Population

Several exposure markers were measured in adolescents, as shown in Table 2A. Among these exposure markers, blood Cadmium (Cd) ($P=0.000$) and urinary 1-OH-P ($P=0.013$) were found to be significantly elevated in smokers (Table 2A). Therefore, subsequent analyses of these markers were conducted solely in non-smokers. Of all possible exposure-effect relationships, only chlorobenzene ($P=0.046$) and ethylbenzene ($P=0.022$) exposures were significantly related to levels of DNA damage levels as determined by COMET assay (Table 2A). The effect of benzene exposure on urinary tt-MA levels was also tested; however no significant relationships were observed. Dosimeter data for benzene was only available in 54 individuals.

In case of the exposure-effect relationship between chlorobenzene exposure and Comet DNA damage, no significant predictors for the observed inter-individual variation were found among evaluated SNPs (Table 3). In contrast, *CAT* C²⁶²T ($P=0.027$) and *GSTT1* null ($P=0.035$) polymorphisms appeared the most important variables of genetic susceptibility for Comet DNA damage levels in association with exposure to ethylbenzene (Table 3). For the *CAT* C²⁶²T polymorphism, individuals carrying the double variant allele had 1.4 times higher DNA damage levels than

Table 2. Results of the correlation analyses, presented as P-values, in the adolescent (A) and adult (B) population. If smoking significantly affected the exposure marker level, the correlation analysis was conducted on non-smokers. Significant correlations are marked (P<0.05).

A

	Smoking*	Comet**	tt-MA**
p'p'DDE	0.858	0.521	
HCB	0.343	0.552	
Cd_Blood	0.000	0.353***	
Cd_Urine	0.209	0.146	
Pb	0.375	0.080	
PCB	0.548	0.755	
1OHP	0.013	0.303***	
ttMA	0.256	0.602	
Benzene (N=54)	0.165	0.462	0.808
ClBenzene (N=54)	0.837	0.046	0.747
EthylBenzene (N=54)	0.808	0.022	0.892

B

	Smoking*	Comet**	8-OHdG**	PSA**	CEA**	p53**
PCB	0.544	0.676	0.755	0.843	0.520	0.230
HCB	0.517	0.239	0.003	0.623	0.546	0.082
p'p'DDE	0.562	0.651	0.613	0.873	0.836	0.150
Cd_Blood	0.000	0.156***	0.666***	0.139***	0.530***	0.908***
Cd_Urine	0.000	0.790***	0.007***	0.158***	0.546***	0.003***
Pb	0.510	0.438	0.088	0.663	0.004	0.356
1OHP	0.000	0.777***	0.029***	0.078***	0.371***	0.680***
tt-MA	0.074	0.238	0.001	0.766	0.075	0.944
Comet	0.400	***	0.683	0.122	0.141	0.694
8-OHdG	0.887	0.683	***	0.861	0.136	0.703

* Effect of moking, Mann Whitney-U
** Exposure-effect relationships, Linear Regression
*** Relationship in non-smokers
**** N = 54

CAT C²⁶²T wild type individuals ($P=0.018$, Figure 2A). Individuals lacking the *GSTT1* gene had 1.2 times higher DNA damage levels as compared to subjects carrying the wild type *GSTT1* gene variant, in response to ethylbenzene exposure (Figure 2B), although not statistically significant ($P=0.139$).

Figure 2. Effect of *CAT* C-262T (Figure 2A) and *GSTT1* null (Figure 2B) genotypes on the relationship between ethyl-benzene exposure and Comet levels in adolescents, presented as observed effect / expected effect (O/E Ratio). The dashed line represents the situation where the expected response equals the observed response. Results are presented as mean $\pm$ SE.

Adult Population

Investigated biomarkers in adults, the effect of smoking on these biomarkers, and exposure-effect relationships are shown in Table 2B. Only Cd, in blood as well as in urine, and urinary 1-OH-P levels were significantly elevated by smoking ($P=0.000$ for each biomarker, Table 2B). Therefore, all subsequent statistical analyses of these biomarkers were conducted in non-smokers. Significant exposure-effect relationships were established between levels of 8-OHdG, and four of the investigated exposure markers, namely HCB ($P=0.003$), urinary Cd ($P=0.007$, in non-smokers), urinary 1-OH-P ($P=0.029$, in non-smokers) and urinary tt-MA ($P=0.001$) (Table 2B). Moreover, two tumor markers, namely CEA and P53 were significantly associated with Lead (Pb) in blood ($P=0.004$) and urinary Cd ($P=0.003$, in non-smokers) respectively (Table 2B). These associations will be detailed below:

HCB vs 8-OHdG

A positive exposure-effect relationship was found between HCB exposure and 8-OHdG DNA damage levels. Females have 1.1 times higher 8-OHdG DNA damage levels upon HCB exposure than males ($P=0.01$, Figure 3A). Additionally, inter-individual variations

in developing DNA damage after exposure to HCB significantly depends on the *mEH* genotype (Table 3). As Figure 3B shows, HCB-related 8-OHdG levels were 0.9 times lower in carriers of one or more *mEH* $H^{139}R$ variant alleles compared to wild type individuals, however, this difference was not statistically significant ($P=0.055$).

Figure 3. Effect of Sex (Figure 3A) and the *mEH* $H^{139}R$ (Figure 3B) polymorphism on the relationship between HCB exposure and 8-OHdG levels in adults, presented as observed effect / expected effect (O/E Ratio). The dashed line represents the situation where the expected response equals the observed response. Results are presented as mean ± SE.

Urinary Cd vs 8-OHdG

Exposures to Cd as measured in urine, was found to be associated with levels of 8-OHdG in non-smokers. The deletion of the *GSTT1* gene was found to be a significant predictor of inter-individual differences within this relationship (Table 3). As can be seen in Figure 4, 0.9 times lower DNA damage levels were observed for individuals carrying the wild type allele for *GSTT1* compared to individuals lacking the *GSTT1* gene ($P=0.039$).

1-OHP vs 8-OHdG

An association between urinary 1-OHP and 8-OHdG DNA damage levels was found in non-smokers. Inter-individual variations within this relationship was affected by sex ($P=0.012$), *mEH* $H^{139}R$ ($P=0.023$) and *CYP1A1* $T^{461}N$ ($P=0.05$) (Table 3). According to their urinary 1-OHP levels, females had 1.1 times higher DNA damage levels as compared to males ($P=0.008$, Figure 5A). Moreover, as can be seen in Figure 5B, individuals carrying the *mEH* $H^{139}R$ polymorphism have 0.9 times lower 1-OHP related DNA damage levels than wild type individuals ($P=0.044$). Lastly, the level of 8-OHdG DNA damage, based on urinary 1-OHP dose, was 1.2 times higher in individuals carrying one variant allele for the *CYP1A1* $T^{461}N$ polymorphism as compared to wild type

Figure 4. Effect of the *GSTT1* null genotype on the relationship between cadmium exposure and 8-OHdG levels in adults, presented as observed effect / expected effect (O/E Ratio). The dashed line represents the situation where the expected response equals the observed response. Results are presented as mean ± SE.

Figure 5. Effect of sex (Figure 5A) and the *mEH H139R* (Figure 5B) and *CYP1A1 T461N* (Figure 5C) genotypes on the relationship between urinary 1-hydroxy-pyrene and 8-OHdG levels in adults, presented as observed effect / expected effect (O/E Ratio). The dashed line represents the situation where the expected response equals the observed response. Results are presented as mean ± SE.

individuals ($P=0.033$). As can be seen in Figure 5C, no *CYP1A1* $T^{461}N$ double variant individuals were detected in this population.

tt-MA vs 8-OHdG

Only Sex had a significant effect on the positive relationship between urinary *tt*-MA and 8-OHdG DNA damage levels ($P=0.002$, Table 3). Females had 1.1 times higher DNA damage levels compared to males ($P=0.002$), based on the exposure-effect curve between urinary *tt*-MA and 8-OHdG, as shown in Figure 6.

Figure 6. Effect of sex on the relationship between urinary *tt*-muconic acid and 8-OHdG levels in adults, presented as observed effect / expected effect (O/E Ratio). The dashed line represents the situation where the expected response equals the observed response. Results are presented as mean ± SE.

Pb vs CEA

The *BRCA2* $N^{372}H$ polymorphism ($P=0.046$) together with Smoking ($P=0.000$) significantly affected the positive relationship between Pb exposure and CEA levels (Table 3). Although smoking was not found to elevate the level of Pb exposure ($P=0.510$), smokers had 2.7 times higher CEA levels as compared to non-smokers ($P<0.001$) in response to Pb exposure, as can be seen in Figure 7A. Additionally, variations around the exposure-effect relationship were also found to be predicted by the *BRCA2* $N^{372}H$ polymorphism. Observed Pb-associated CEA levels were respectively 1.9 and 1.6 times higher in individuals carrying the double variant allele as in heterozygous and wild type individuals ($P=0.003$, Figure 7B).

Cd vs p53

A positive exposure-effect relationship between urinary Cd and serum p53 was observed in non-smokers. *GSTM1**0 ($P=0.004$) and *BRCA2* $N^{372}H$ ($P=0.024$) were found to contribute significantly to inter-individual variations within this relationship (Table 3). Wild type *GSTM1* individuals had 1.1 times higher p53 levels in association with urinary Cd excretion than individuals lacking the *GSTM1* gene, although not statistically

significant ($P=0.106$, Figure 8A). Moreover, serum p53 levels in individuals carrying the double variant allele for the *BRCA2* N³⁷²H polymorphism were observed to be respectively 1.3 and 1.4 times higher as compared to wild type and heterozygous individuals ($P=0.024$, Figure 8B).

Figure 7. Effect of smoking (Figure 7A) and the *BRCA2* N372H polymorphism (Figure 7B) on the relationship between Pb exposure and CEA levels in adults, presented as observed effect / expected effect (O/E Ratio). The dashed line represents the situation where the expected response equals the observed response. Results are presented as mean $\pm$ SE.

Figure 8. Effect of the *GSTM1* null (Figure 8A) and *BRCA2* N372H (Figure 8B) genotypes on the relationship between Cd exposure and serum P53 levels in adults, presented as observed effect / expected effect (O/E Ratio). The dashed line represents the situation where the expected response equals the observed response. Results are presented as mean $\pm$ SE.

Table 3. Results of genotype, gender and smoking habits on the different exposure-effect relationships (Table 2A and B) in the adolescent and adult populations. When smoking significantly affected the level of the exposure marker, the relationship was investigated in non-smokers. Corresponding P-value is given between brackets.

Relationship	Significantly differently distributed variables between high and low responders	Response	
		Effect	O/E Ratio
ADOLESCENTS			
Chlorine-Benzene * Comet	no significant variables	***	***
Ethyl-Benzene * Comet	CAT C-262T (0.027)	Variant allele associated with increased Comet levels	1.30
	GSTT1*0 (0.035)	Increased Comet levels in null-individuals	1.11
ADULTS			
HCB * 8-OHdG	sex (0.008)	Higher 8-OHdG levels in Females	1.05
	mEH H139R (0.033)	Lower 8-OHdG levels in individuals with at least one variant allele	0.96 (hz)**; 0.94 (dv)**
Cadmium-Urine * 8-OHdG*	GSTT1*0 (0.041)	Increased 8-OHdG levels in null-individuals	1.10
1-OHP * 8-OHdG*	Sex (0.012)	Females have higher 8-OHdG levels	1.05
	mEH H139R (0.023)	Lower 8-OHdG levels in individuals with at least one variant allele	0.96 (hz)**; 0.93 (dv)**
	CYP1A1 T461N (0.05)	Variant allele associated with increased 8-OHdG levels	1.20
tt-MA * 8-OHdG	Sex (0.002)	Higher 8-OHdG levels in Females	1.06
Lead * CEA	Smoking (0.000)	Smokers have higher CEA levels	2.02
	BRCA2 N372H (0.046)	Higher CEA levels in double variants	1.60
Cadmium-Urine * p53*	GSTM1*0 (0.004)	Lower p53 levels in null-individuals	0.94
	BRCA2 N372H (0.024)	Higher p53 levels in individuals with double variants	1.30

*Analysis in non-smokers

**hz= heterozygous individuals, dv=double variants

Multiple comparisons adjustment for significance

Discussion

The Flanders Environment and Health Study (FLEHS), assigned by the Flemish Government in 2001, aims to obtain a better perspective on exposure levels to environmental chemicals for inhabitants of different regions in Flanders and possible harmful effects of these exposures on human health. The results of this study are supposed to eventually be used to evaluate current environmental risk assessment procedures and to result in new regulations regarding toxic environmental exposures. The current study aims to investigate the role of genetic variability between individuals to identify subgroups at increased susceptibility regarding environmental health effects. Therefore, a random selection taken from the FLEHS study population, comprising 429 adolescents and 361 adults, was genotyped for 36 SNPs in 23 genes involved in relevant pathways with respect to environmental cancer risk assessment. In both subpopulations, several exposure-effect relationships between environmental toxicants and effect markers were established. With one exception, more susceptible subgroups could be defined within all significant relationships by analysis of genetic variants in relevant genes. This stresses the significant role which genetic-based inter-individual differences play in biomonitoring projects studying environmental health effects, such as FLEHS.

Adolescents

DNA damage levels measured by Comet analysis were used as a biomarker for genotoxic effects in association with several environmental exposures in the adolescent subpopulation. Of the investigated toxicants, only chloro- and ethylbenzene were significantly associated with this biomarker, both associations have not yet been described in the literature. The main pathway of benzene biotransformation involves formation of metabolites which are known to generate reactive oxygen species (ROS) through redox cycling (11,12). These ROS eventually induce oxidative DNA damage; exposure to benzene in mice has been shown to induce oxidative DNA damage measured by the Comet assay (13,14). Inter-individual variations around the regression line were observed. In case of chlorobenzene exposure, however, no genetic susceptible subgroups could be defined. This may be due to the low number of individuals in which this exposure marker was actually measured; dosimeter data was only available for 54 individuals within the adolescent subgroup. In case of ethylbenzene exposure, *Catalase* and *GSTT1* genotypes significantly affected the relationship between ethylbenzene exposure and DNA damage. To our knowledge, this is the first report suggesting a genotoxic effect of ethylbenzene in humans. Moreover, increased levels compared to the expected response based on the regression line were observed in carriers of the *CAT* C²⁶²T polymorphism and in individuals lacking *GSTT1* activity. *Catalase* and *GSTT1* are involved in the primary defense against (environmentally induced) oxidative stress. Both poly-

morphisms have been associated with reduced enzyme activity which leads to higher levels of reactive metabolites and increased levels of ROS, and hence, higher DNA damage levels (15,16). Therefore, *GSTT1-null* individuals and/or individuals carrying the *CAT C⁻²⁶²T* polymorphism, may be more susceptible for developing health effects induced by ROS resulting from environmental exposure to ethylbenzene, which is in agreement with our observations; *GSTT1-null* individuals were found to have 1.2 times higher DNA damage levels compared to *GSTT1* wild types, whereas the *CAT C⁻²⁶²T* polymorphism was observed to be associated with a 1.4 times higher DNA damage levels compared to wild type individuals.

Adults

In the adult population, smoking was found to significantly increase the level of Cd exposure and urinary 1-hydroxy-pyrene (1-OHP). The latter is a commonly used biomarker for exposure to Polycyclic Aromatic Hydrocarbons (PAH) (17,18). Tobacco smoke indeed represents a well-known source of both Cd and PAH, and increased levels of Cd and urinary 1-OHP have been previously observed in smokers (19-21). In order to assure that putatively observed effects are solely due to environmental exposures to Cd and PAHs, subsequent statistical analysis of these biomarkers were restricted to non-smokers. Because *trans-trans*-muconic acid (tt-MA) has been reported as a marker for cigarette smoke exposure (22), it is remarkable that tt-MA levels were not found to be higher in smokers compared to non-smokers. However, contradictory results have been reported, regarding the effect of smoking on urinary tt-MA levels (23).

As biomarker of early effect, DNA damage levels in peripheral lymphocytes as measured by Comet assay as well as analysis of urinary 8-hydroxy-2'-deoxyguanosine (8-OHdG) levels were assessed. In addition, several tumour markers (PSA, CEA and p53) were examined as biomarkers of effect.

Urinary excretion of the DNA repair product 8-OHdG is the most commonly used biomarker for assessing oxidative DNA damage as induced by (environmental) genotoxicants, in combination with the efficiency of repair of such damage (24). Several studies have shown an effect of environmental exposures to various agents on urinary 8-OHdG levels (25,26). In the current study, an association of hexachlorobenzene (HCB), urinary Cd, 1-OHP and tt-MA with 8-OHdG levels was observed.

1-OHP- and tt-MA-related levels of DNA damage were found to be higher in females compared to males. Previous studies found increased levels of both exposure markers in females (27,28).

Additionally, an effect of *CYP1A1 T⁴⁶¹N* and *mEH H¹³⁹R* on the exposure-effect relationship between 1-OHP and DNA damage levels was found. CYP1A1 is a Phase I biotransformation enzyme, which activates certain carcinogenic compounds, like several PAHs, into highly reactive metabolites. A key role for CYP1A1 in pyrene metabolism has been described (29), and variants of the CYP1A1 gene have been associated with increased enzyme activity (30-33). The *CYP1A1 T⁴⁶¹N* polymorphism de-

scribed here, results in a greater capacity of metabolizing PAHs, and thus higher levels of reactive metabolites are formed. These reactive metabolites are known to induce oxidative stress and eventually oxidative DNA damage (34). Therefore, within the described exposure-effect relationship between 1-OHP and 8-OHdG levels, it may be expected that individuals carrying the *CYP1A1 T⁴⁶¹N* polymorphism are more susceptible to develop oxidative DNA damage upon PAH exposure compared to wild-type individuals. Indeed, the observed 1-OHP induced DNA damage levels were 1.2 times higher in individuals carrying the *CYP1A1 T⁴⁶¹N* polymorphism than in *CYP1A1 T⁴⁶¹N* wild type individuals. The role of mEH is more complex, as it may catalyze both Phase I and Phase II biotransformation reactions (35). In general, mEH has been described as a very important Phase II enzyme (35). At least two functional polymorphisms have been determined in mEH, of which a high activity variant of the enzyme is coded by the *mEH H¹³⁹R* polymorphism in exon 4 (36). In the current study, individuals carrying the *mEH H¹³⁹R* polymorphism had 0.9 times lower 8-OHdG levels as compared to *mEH H¹³⁹R* wild types. Although the level of this effect is very low, the observed pattern is as expected, as the high activity variant of the enzyme (*mEH H¹³⁹R*, *mEH R¹³⁹R*) leads to more detoxification, and hence, to lower DNA damage levels.

A similar effect of mEH on 8-OHdG levels was found upon HCB exposure. HCB is a persistent organochloric pesticide which accumulates in humans and is widespread in the environment because it is not well degraded (37). The oxidative metabolites of HCB are able to bind to macromolecules, such as DNA (38). Urinary levels of 8-OHdG were observed to be increased in association with HCB exposure, and *mEH H¹³⁹R* wild type individuals were defined as the more susceptible subgroup. Similar to the effect observed in response to PAH exposure, the level of the effect was marginal (0.9 times lower levels in hetero- and homozygous carriers of the *mEH H¹³⁹R* polymorphism compared to wild types), however, the observed pattern indicates again a protective effect of the *mEH H¹³⁹R* polymorphism, where the low activity of the wild type enzyme may lead to less detoxification of reactive metabolites and hence, to higher DNA damage levels.

Furthermore, an impact on dose-relationships between Cd and urinary 8-OHdG levels by a genetic variant from another Phase II enzyme, namely GSTT1, was found. Cd is a highly toxic metal, and a well-known occupational (e.g. heavy metal mining, waste incinerators and farm fertilizers) and environmental pollutant (water, food and air contaminations) (19). Occupational exposure to Cd has been associated with lung, prostate, pancreas and kidney cancer, and it has been correspondingly classified as a category 1 carcinogen by IARC (39). Cd is not directly genotoxic, but it may affect genome instability by inhibition of DNA repair, depletion of cellular anti-oxidants like Glutathione (GSH) and indirect formation of ROS (through the replacement of iron, zinc and copper from various proteins, like ferritin, thereby increasing the level of free metal-ions which participate in oxidative stress via fenton reactions) (40-42). This Cd-induced oxidative stress has been described to cause 8-OHdG-related DNA strand breaks (43). Also in the current study, a positive exposure-effect relationship between

Cd exposure and 8-OHdG levels was found. Within this relationship, individuals lacking GSTT1 activity, have 1.1 times higher oxidative DNA damage levels after Cd exposure as compared to GSTT1 positive individuals. GSTT1 is an enzyme involved in the GSH pathway, one of the most essential antioxidant defense mechanisms. Thus, *GSTT1-null* individuals have an impaired antioxidant defense mechanism which results in an enhanced effect of oxidative stress, induced indirectly by Cd exposure, and hence, to higher 8-OHdG DNA damage levels. This is confirmed by our observations.

Exposure to Cd was also found to be related to an increase in serum p53 levels, an association described for the first time in this study. It has been suggested by Xie and Shaikh (2006) that Cd can activate p53 by the indirect generation of oxidative stress (44). For example, Cd has been reported to generate hydrogen peroxide (8), which in turn is known to trigger p53 transcriptional activity (45). Moreover, Cd is known to interfere with DNA repair, which in turn may also induce apoptosis (19). Analysis of the p53 tumor suppressor gene is proposed as a biomarker for effect in biomonitoring studies (46,47). Within the observed exposure-effect relationship between Cd exposure and p53 levels, wild type individuals for *GSTM1* (1.1 times higher than in *GSTM1* variants) and individuals homozygously carrying the *BRCA2 N³⁷²H* polymorphism (1.3 and 1.4 times higher compared to respectively wild-type and heterozygous individuals) were found to have the highest serum p53 levels. *GSTM1*-like *GSTT1*- is a member of the GST super-family and uses GSH in the defense against oxidative stress. GSH in turn, is known to protect cells against apoptosis by various multi-factorial mechanisms including interaction with pro-apoptotic signals (8). A possible explanation for the observed effect of *GSTM1* on the exposure-effect relationship between Cd exposure and serum p53 levels is that, in *GSTM1*- positive individuals, cellular GSH levels are depleted by *GSTM1* activity in reaction on the oxidative stress generated by Cd exposure. This depletion of GSH increases cellular damage (e.g. chromosome aberrations), triggers p53 induced apoptosis and thus increases serum p53 levels (48). In addition to the interference of Cd with the DNA repair system, carrying the *BRCA2 N³⁷²H* polymorphism, associated with a decrease in DNA repair, may lead to even higher DNA damage levels and hence an increase in Cd-induced serum p53 levels compared to average levels.

Another tumor-marker, Carcino Embryonic Antigen (CEA), was also found to be increased upon exposure to a heavy metal, namely Pb. This is the first report that describes this relationship. Pb is indirectly acting genotoxic metal, classified as carcinogenic by IARC, and (occupational) exposure to Pb has been reported to cause DNA damage as determined by COMET assay (49-51). It has been proposed that the DNA damage related evaluation of blood levels of proteins like CEA may be useful as markers to identify health risk after exposure to environmental pollutants (52). In the current study, Pb induced CEA levels were respectively 1.9 and 1.6 times increased in carriers of the *BRCA2 N³⁷²H* polymorphism as compared to heterozygous and wild type individuals. In addition to the interference of Pb with DNA repair leading to higher DNA damage levels, the decrease in DNA repair activity in carriers of the *BRCA2 N³⁷²H* polymorphism may lead to an increase in DNA damage, and hence, in the level of this

tumor marker. Moreover, elevated levels of serum CEA have been reported in smokers compared to non-smokers (53). In agreement to these findings, also in our study, smokers were observed to have 2.7 times higher CEA levels than non-smokers upon Pb exposure.

Overall

In both subpopulations the genotype was shown to significantly influence the inter-individual susceptibility for biomarker response to harmful environmental exposures. However, in most cases, observed effects were relatively low. This may be the result of the rather low exposure levels of the investigated compounds, which lead to weak exposure-effect relationships and, as a consequence, make it difficult to detect an effect of the genotype on these specific associations.

Regarding the complexity of carcinogen exposure and the number of pathways that is involved in the modification of their effects, it is unlikely that one single SNP can predict these inter-individual differences in response to these exposures. Investigating multiple polymorphisms in different pathways simultaneously should therefore be considered. Moreover, as these low-penetrance polymorphisms are risk factors for developing health effects in response to (environmental) exposures, future studies assessing genotyping as a biomarker for susceptibility must shift interest from exploring the effect of the genotype on the specific biomarker, towards the investigation of the effect of the genotype on the exposure-effect relationship between markers of interest. This research proposes a framework for the application of investigating multiple genetic polymorphisms in pathways of interest as an important tool to identify more susceptible subgroups in biomonitoring studies assessing human health risk in response to hazardous environmental exposures.

In conclusion, this innovative study has provided new insights in exposure-effect relationships in humans exposed to certain environmental contaminants, and secondly, has quantified the role of genetic polymorphisms in these relationships enabling the identification of more susceptible subgroups. These findings suggest a critical role of inter-individual differences in environmental cancer risk assessment. To be effective in the protection of diseases such as cancer at the individual's level, risk assessment methods must challenge the assumption that all individuals are uniformly vulnerable to environmental pollutants. It is generally considered that the linear non-threshold extrapolation model is conservative enough and that exposure levels to (genotoxic) carcinogens are sufficiently low. However, results of the current study show that more susceptible subgroups with a higher risk for environmental carcinogenesis can be defined, based on the effect of specific genotypes on the corresponding exposure-effect relationships. Although the results of this study should be interpreted carefully as overall effects were relatively modest and they comprise biomarker analysis instead of ultimate cancer incidence, the outcome of this work underlines the societal relevance of

Multiplex genotyping as biomarker for susceptibility

continuing the responsible care debate about inter-individual differences in genetically predisposed susceptibility for environmental cancer risk.

References

1. Perera, F.P. and Weinstein, I.B. (1982) Molecular epidemiology and carcinogen-DNA adduct detection: new approaches to studies of human cancer causation. *J Chronic Dis*, 35, 581-600.
2. Higginson, J. (1976) Importance of environmental factors in cancer. *IARC Sci Publ*, 15-23.
3. Perera, F.P. (1996) Molecular epidemiology: insights into cancer susceptibility, risk assessment, and prevention. *J Natl Cancer Inst*, 88, 496-509.
4. Tharappel, J.C., Lee, E.Y., Robertson, L.W., et al. (2002) Regulation of cell proliferation, apoptosis, and transcription factor activities during the promotion of liver carcinogenesis by polychlorinated biphenyls. *Toxicol Appl Pharmacol*, 179, 172-84.
5. Safe, S.H. (1994) Polychlorinated biphenyls (PCBs): environmental impact, biochemical and toxic responses, and implications for risk assessment. *Crit Rev Toxicol*, 24, 87-149.
6. Mariussen, E. and Fonnum, F. (2006) Neurochemical targets and behavioral effects of organohalogen compounds: an update. *Crit Rev Toxicol*, 36, 253-89.
7. Ezendam, J., Vissers, I., Bleumink, R., et al. (2003) Immunomodulatory effects of tetrachlorobenzoquinone, a reactive metabolite of hexachlorobenzene. *Chem Res Toxicol*, 16, 688-94.
8. Valko, M., Rhodes, C.J., Moncol, J., et al. (2006) Free radicals, metals and antioxidants in oxidative stress-induced cancer. *Chem Biol Interact*, 160, 1-40.
9. Knaapen, A.M., Ketelslegers, H.B., Gottschalk, R.W., et al. (2004) Simultaneous genotyping of nine polymorphisms in xenobiotic-metabolizing enzymes by multiplex PCR amplification and single base extension. *Clin Chem*, 50, 1664-8.
10. Ketelslegers, H.B., Gottschalk, R.W., Gottschalk, R.W., et al. (2006) Interindividual variations in DNA adduct levels assessed by analysis of multiple genetic polymorphisms in smokers. *Cancer Epidemiol Biomarkers Prev*, 15, 624-9.
11. Midorikawa, K., Uchida, T., Okamoto, Y., et al. (2004) Metabolic activation of carcinogenic ethylbenzene leads to oxidative DNA damage. *Chem Biol Interact*, 150, 271-81.
12. Subrahmanyam, V.V., Kolachana, P. and Smith, M.T. (1991) Hydroxylation of phenol to hydroquinone catalyzed by a human myeloperoxidase-superoxide complex: possible implications in benzene-induced myelotoxicity. *Free Radic Res Commun*, 15, 285-96.
13. Tuo, J., Loft, S., Thomsen, M.S., et al. (1996) Benzene-induced genotoxicity in mice in vivo detected by the alkaline comet assay: reduction by CYP2E1 inhibition. *Mutat Res*, 368, 213-9.
14. Tuo, J., Deng, X., Loft, S., et al. (1999) Dexamethasone ameliorates oxidative DNA damage induced by benzene and LPS in mouse bone marrow. *Free Radic Res*, 30, 29-36.
15. Ahn, J., Gammon, M.D., Santella, R.M., et al. (2005) Associations between breast cancer risk and the catalase genotype, fruit and vegetable consumption, and supplement use. *Am J Epidemiol*, 162, 943-52.
16. Hayes, J.D. and Strange, R.C. (2000) Glutathione S-transferase polymorphisms and their biological consequences. *Pharmacology*, 61, 154-66.
17. Jongeneelen, F.J., Anzion, R.B., Leijdekkers, C.M., et al. (1985) 1-hydroxypyrene in human urine after exposure to coal tar and a coal tar derived product. *Int Arch Occup Environ Health*, 57, 47-55.

18. Jongeneelen, F.J. (1997) Methods for routine biological monitoring of carcinogenic PAH-mixtures. *Sci Total Environ*, 199, 141-9.
19. Bertin, G. and Averbeck, D. (2006) Cadmium: cellular effects, modifications of biomolecules, modulation of DNA repair and genotoxic consequences (a review). *Biochimie*, 88, 1549-59.
20. Saieva, C., Tumino, R., Masala, G., et al. (2003) Urinary 1-hydroxypyrene and t,t-muconic acid as biomarkers of exposure to environmental pollutants in two areas in Italy (EPIC-Florence and Ragusa). *Tumori*, 89, 679-86.
21. Jongeneelen, F.J. (1994) Biological monitoring of environmental exposure to polycyclic aromatic hydrocarbons; 1-hydroxypyrene in urine of people. *Toxicol Lett*, 72, 205-11.
22. Hecht, S.S. (2002) Human urinary carcinogen metabolites: biomarkers for investigating tobacco and cancer. *Carcinogenesis*, 23, 907-22.
23. de Paula, F.C., Silveira, J.N., Junqueira, R.G., et al. (2003) [Assessment of urinary trans, trans-muconic acid as a biomarker of exposure to benzene]. *Rev Saude Publica*, 37, 780-5.
24. Hong, Y.C., Park, H.S. and Ha, E.H. (2000) Influence of genetic susceptibility on the urinary excretion of 8-hydroxydeoxyguanosine of firefighters. *Occup Environ Med*, 57, 370-5.
25. Loft, S., Vistisen, K., Ewertz, M., et al. (1992) Oxidative DNA damage estimated by 8-hydroxydeoxyguanosine excretion in humans: influence of smoking, gender and body mass index. *Carcinogenesis*, 13, 2241-7.
26. Kim, Y.D., Lee, C.H., Nan, H.M., et al. (2003) Effects of genetic polymorphisms in metabolic enzymes on the relationships between 8-hydroxydeoxyguanosine levels in human leukocytes and urinary 1-hydroxypyrene and 2-naphthol concentrations. *J Occup Health*, 45, 160-7.
27. Cocco, P., Tocco, M.G., Ibba, A., et al. (2003) trans,trans-Muconic acid excretion in relation to environmental exposure to benzene. *Int Arch Occup Environ Health*, 76, 456-60.
28. Merlo, F., Andreassen, A., Weston, A., et al. (1998) Urinary excretion of 1-hydroxypyrene as a marker for exposure to urban air levels of polycyclic aromatic hydrocarbons. *Cancer Epidemiol Biomarkers Prev*, 7, 147-55.
29. Nerurkar, P.V., Okinaka, L., Aoki, C., et al. (2000) CYP1A1, GSTM1, and GSTP1 genetic polymorphisms and urinary 1-hydroxypyrene excretion in non-occupationally exposed individuals. *Cancer Epidemiol Biomarkers Prev*, 9, 1119-22.
30. Cascorbi, I., Brockmoller, J. and Roots, I. (1996) A C4887A polymorphism in exon 7 of human CYP1A1: population frequency, mutation linkages, and impact on lung cancer susceptibility. *Cancer Res*, 56, 4965-9.
31. Spurr, N.K., Gough, A.C., Stevenson, K., et al. (1987) Msp-1 polymorphism detected with a cDNA probe for the P-450 I family on chromosome 15. *Nucleic Acids Res*, 15, 5901.
32. Crofts, F., Cosma, G.N., Currie, D., et al. (1993) A novel CYP1A1 gene polymorphism in African-Americans. *Carcinogenesis*, 14, 1729-31.
33. Hayashi, S., Watanabe, J., Nakachi, K., et al. (1991) Genetic linkage of lung cancer-associated MspI polymorphisms with amino acid replacement in the heme binding region of the human cytochrome P450IA1 gene. *J Biochem (Tokyo)*, 110, 407-11.
34. Loft, S. and Poulsen, H.E. (1996) Cancer risk and oxidative DNA damage in man. *J Mol Med*, 74, 297-312.
35. Georgiadis, P., Demopoulos, N.A., Topinka, J., et al. (2004) Impact of phase I or phase II enzyme polymorphisms on lymphocyte DNA adducts in subjects exposed to urban air pollution and environmental tobacco smoke. *Toxicol Lett*, 149, 269-80.

36. Hassett, C., Robinson, K.B., Beck, N.B., et al. (1994) The human microsomal epoxide hydrolase gene (EPHX1): complete nucleotide sequence and structural characterization. *Genomics*, 23, 433-42.
37. Sala, M., Sunyer, J., Otero, R., et al. (1999) Health effects of chronic high exposure to hexachlorobenzene in a general population sample. *Arch Environ Health*, 54, 102-9.
38. Li, Y.F., Shi, N., Li, H.Y., et al. (2006) [Toxicity and oxidative stress on rats by hexachlorobenzene]. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi*, 24, 601-4.
39. IARC (1982) Some industrial chemicals and dyestuffs. IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans. IARC, Lyon, 29, 93-148.
40. Price, D.J. and Joshi, J.G. (1983) Ferritin. Binding of beryllium and other divalent metal ions. *J Biol Chem*, 258, 10873-80.
41. Casalino, E., Sblano, C. and Landriscina, C. (1997) Enzyme activity alteration by cadmium administration to rats: the possibility of iron involvement in lipid peroxidation. *Arch Biochem Biophys*, 346, 171-9.
42. McMurray, C.T. and Tainer, J.A. (2003) Cancer, cadmium and genome integrity. *Nat Genet*, 34, 239-41.
43. Mikhailova, M.V., Littlefield, N.A., Hass, B.S., et al. (1997) Cadmium-induced 8-hydroxydeoxyguanosine formation, DNA strand breaks and antioxidant enzyme activities in lymphoblastoid cells. *Cancer Lett*, 115, 141-8.
44. Xie, J. and Shaikh, Z.A. (2006) Cadmium induces cell cycle arrest in rat kidney epithelial cells in G2/M phase. *Toxicology*, 224, 56-65.
45. Polyak, K., Xia, Y., Zweier, J.L., et al. (1997) A model for p53-induced apoptosis. *Nature*, 389, 300-5.
46. Badawi, A.F. (1996) Molecular and genetic events in schistosomiasis-associated human bladder cancer: role of oncogenes and tumor suppressor genes. *Cancer Lett*, 105, 123-38.
47. Collins, A.R. (1998) Molecular epidemiology in cancer research. *Mol Aspects Med*, 19, 359-432.
48. Ray, S. and Chatterjee, A. (2007) Influence of glutathione on the induction of chromosome aberrations, delay in cell cycle kinetics and cell cycle regulator proteins in irradiated mouse bone marrow cells. *Int J Radiat Biol*, 83, 347-54.
49. Steinmetz-Beck, A., Szahidewicz-Krupska, E., Beck, B., et al. (2005) [Genotoxicity effect of chronic lead exposure assessed using the comet assay]. *Med Pr*, 56, 295-302.
50. Karakaya, A.E., Ozcagli, E., Ertas, N., et al. (2005) Assessment of abnormal DNA repair responses and genotoxic effects in lead exposed workers. *Am J Ind Med*, 47, 358-63.
51. Danadevi, K., Rozati, R., Saleha Banu, B., et al. (2003) DNA damage in workers exposed to lead using comet assay. *Toxicology*, 187, 183-93.
52. Perera, F.P., Hemminki, K., Gryzbowska, E., et al. (1992) Molecular and genetic damage in humans from environmental pollution in Poland. *Nature*, 360, 256-8.
53. van Larebeke, N., Husson, B., De Coen, W., et al. (2003) Immunologic and other biological parameters as a function of smoking status and of residence in areas differing in terms of air pollution. *Int J Environ Health Res*, 13, 55-69.

Chapter 6

The prevalence of an at-risk genotype for environmental carcinogenesis decreases with age; survival of the fittest?

Hans B. Ketelslegers¹, Roger W.L. Godschalk¹, Ralph W.H. Gottschalk¹, Ad M. Knaapen¹, Gudrun Koppen², Greet. Schoeters², Willy Bayens³, Joost H.M. van Delft¹
and Jos C.S. Kleinjans¹

¹Department of Health Risk Analysis and Toxicology, Maastricht University,
Maastricht, the Netherlands

²Flemish Institute for Technological Research, Mol, Belgium

³Analytical and Environmental Chemistry Department, ANCH, Vrije Universiteit
Brussel, Brussels, Belgium

Abstract

Certain individuals are at higher risk for developing environmental cancer as compared to the averaged population, which may be due to genetic polymorphisms in genes involved in biotransformation of environmental carcinogens and in modulation of their adverse effects. Assuming that an at risk genotype coding for enzymes involved in environmental carcinogenesis is associated with an increased risk for cancer development and possible death, it may be postulated that the prevalence of the number of risk alleles in these genes will decrease with age in a random selection of healthy individuals. In the current study, 442 adolescents (average age: 15 years) and 380 adults (average age: 60 years), all healthy Caucasians originating from Flanders, Belgium, were genotyped for 29 polymorphisms in 19 toxicologically relevant genes involved in biotransformation of environmental carcinogens, oxidative stress defence and DNA repair. For statistical analysis, genotypes were coded based on the number of risk alleles and, subsequently, a sum of risk alleles was computed for each individual. Comparison of this overall sum of risk alleles between the two age groups showed that relatively high number of risk alleles were less frequent among adult subjects as compared to the adolescent population (39% versus 49% of subjects carried >18 risk alleles, $P=0.005$). To evaluate this in more detail, sums of risk alleles for xenobiotic metabolizing enzymes (subdivided into phase I and II biotransformation reactions), oxidative stress and DNA repair related genes were separately computed. The sum of risk alleles in phase I genes (*CYP450*'s), oxidative stress or DNA repair did not differ significantly between the 'elderly' and the 'young' study populations. However, a high sum of risk alleles for biotransformation (>5 risk alleles), and more specifically, for phase II enzymes (GSTs, NATs; >4 risk alleles), was significantly less frequent among adults (44% for the total sum of risk alleles in biotransformation enzymes and 42% for phase II enzymes) as compared to adolescents (54% for all biotransformation polymorphisms and 56% for phase II enzymes, $P=0.002$ and <0.001 respectively). In addition, analyses of the single polymorphisms showed a significantly lower prevalence of the *GSTP1* $I^{105}V$, *NAT2* $I^{114}T$, *NQO1* $R^{139}W$ and *BRCA2* 5'UTR genotypes in the 'elderly' compared to the 'young' age group. It is concluded that the number of carcinogenesis related at-risk alleles is lower in a random selection among the general population of 'elderly' individuals as compared to 'young' individuals, which supports the hypothesis that an at-risk genotype for chemically induced cancer affects life expectancy: Survival of the Fittest?

Introduction

We are continuously exposed to a wide variety of potentially harmful chemicals as present in our environment and as consequence of our daily life style, some of which have been classified as human carcinogens by the International Agency for Research on Cancer (IARC). Such xenobiotics may contribute to most chronic human diseases, especially cancer. Humans, as they age, have to cope with a lifetime-accumulating exposure to such (carcinogenic) agents and corresponding damage to cellular functions induced by these compounds. The Flanders Environment and Health Study (FLEHS), is a large biomonitoring campaign that aims to investigate the impact of these hazardous exposures on human health across the general population of Flanders. Environmental carcinogens, such as Benzene and Cadmium (known to be human carcinogens (1)) and Polycyclic Aromatic Hydrocarbons, Polychlorinated Biphenyls and Hexachlorobenzene (reasonably anticipated to be human carcinogens (1)), have been established in both Flemish adolescents and adults in this context (www.milieu-en-gezondheid.be).

Many xenobiotics to which humans are exposed, would not be health-threatening if they were not bioactivated through metabolism. Xenobiotic metabolism can be divided into two important phases (2). During phase I, potentially harmful compounds are activated by oxidation reactions resulting in the formation of reactive intermediates, which may damage DNA and other intracellular targets. In addition, these metabolic processes may contribute to the formation of reactive oxygen species (ROS), mainly as a result of the use of oxygen in Phase I reactions (3). Besides indirect formation of ROS, several environmental pollutants are known to contribute directly to the intracellular level of ROS. ROS cause multiple forms of oxidative stress including oxidative DNA damage. A steady state level is normally maintained by a balance between the rate at which oxidative damage is produced and by which it is efficiently counteracted or repaired. This is partly dependent on inactivation of the ROS by anti-oxidant defense mechanisms (4). Electrophilic metabolites may be detoxified in phase II, by the conjugation of endogenous ligands to the activated compounds, thereby increasing their hydrophilic nature and facilitating urinary excretion. An imbalance between oxidative metabolism and detoxification causes highly reactive intermediate products to accumulate, which may form DNA lesions and eventually cause mutations, initiating or promoting the carcinogenic process. This may be counteracted by several DNA repair pathways which are capable of restoring DNA damage thereby reducing mutation frequencies of cancer-related genes (5).

Overall, disadvantageous polymorphisms in these pathways may constitute an at-risk genotype increasing an individual's susceptibility for chemical carcinogenesis, and eventually affect human longevity (Figure 1)(6). In a study focusing on GST genotypes on cancer risk conducted by Voho et al. (7) it was hypothesized that if these polymorphisms are indeed true risk factors for environmental cancer development, their prevalence should decrease with age in the cancer-free fraction of the general popula-

Figure 1. Human longevity and age related diseases such as cancer are strongly dependent on interactions between environmental- and genetic factors. Risk alleles (RA) in genetic pathways that maintain the defenses against these potentially hazardous environmental exposures and the effective repair of the damage caused by these compounds, may affect the inter-individual variations in human life expectancy.
'Age=15' represents the 'young', whereas 'age>50' represents the 'elderly' age group.

tion; the results of the conducted study indeed supported this hypothesis (7). In addition, few studies comparing prevalences of different genotypes between age groups exist. In addition, contradicting results about the effect of the genotype on human longevity were reported (7-10). A possible explanation for this may be that these studies focus on single genotypes. It is however very unlikely that one single SNP is pivotal in causing inter-individual differences in environmental disease development. As a consequence, in studies on environmental health and disease, it is crucial to assess multiple polymorphisms simultaneously. In this way it is possible to obtain a better insight in human susceptibility in response to exposure to potentially harmful agents. In the current study, the most relevant polymorphisms in genes coding for enzymes involved in xenobiotic metabolism, oxidative stress defence and DNA repair, were analyzed in healthy Caucasian adolescents and adults from the FLEHS study cohort, to test the hypothesis that the prevalence of at-risk alleles -associated with an increased risk for chemically induced carcinogenesis- decreases with age in a randomly selected population.

Material and Methods

Study Population

The study population consisted of a random selection of 442 healthy adolescents (182 females and 260 males, average age 15 ± 0.5 years) and 380 healthy adults (176 females and 204 males, aged 50 to 65 years at the time of inclusion) originating from Flanders, Belgium, and taken from the FLEHS study cohort (<http://www.milieu-en-gezondheid.be>). These regions were selected because of their different environmental burden. A small proportion ($N=47$; 11%) of the adolescents reported to smoke 7 ± 7 cigarettes per day and 62 adult individuals (16%) currently smoked 17 ± 10 cigarettes per day. Individuals who reported to have an ethnic background other than Caucasian, were excluded from analysis, because the prevalence of certain genotypes may differ between races (11). Also, individuals who reported to suffer or have suffered from cancer ($N=19$) were initially excluded, as the study was to be conducted in a cancer-free population. Population size was sufficient to obtain statistical power of more than 95%, minimizing the chance for false positive findings; using a $\alpha=0.05$, $\delta=0.05$ and $\sigma=0.15$, a population size of at least $N=235$ is required (<http://www.biostat.mc.vanderbilt.edu>). The study was approved by the ethical committee of the University of Antwerp, and every individual signed an informed consent prior to sample collection.

White blood cell and DNA isolation

Immediately after blood collection, white blood cells were isolated by lysing erythrocytes using standard procedures and total white blood cells were stored as cell pellets at -80°C until DNA isolation. DNA was isolated in 96 wells plates using the Invisorb® Blood Midi HTS 96 Kit / C (Invitex, Berlin, Germany), and DNA was subsequently stored in 96 wells plates at -20°C prior to genotyping at a concentration between 40 and 100 ng/ μl .

Genotyping

Genotyping was performed using the SNaPShot multiplex genotyping method described earlier (12). All individuals were genotyped for a total of 29 SNPs in 19 genes involved in biotransformation of exogenous compounds (13 SNPs), 10 DNA repair polymorphisms and 6 SNPs in oxidative stress related genes. SNP's were determined in 4 different multiplex reactions (Table 1). For each multiplex PCR reaction, T_m was optimized: 56°C for multiplex 1, 60°C for multiplex 2, 57°C for multiplex 3 and 60°C for multiplex 4. Each of the multiplex PCR products was subsequently genotyped in a corresponding multiplex Single Base Extension reaction and analyzed on an ABI Prism™ 3100 genetic analyzer using Genescan™ Analysis software (Version 3.7).

Table 1. Overview of the selected SNPs, their effect on enzyme function and definition of risk alleles.

Polymorphism		Expected effect on enzymatic function		Putative Risk Allele	Risk Allele leading to
Biotransformation					
CYP1A2	A>C-154	(1)	Higher inducibility	-154 C	Increased formation of reactive metabolites
CYP1A1	T>C3801	(4)		3801 C	
	I462V	(4)		462 V	
	T461N	(4)	Higher enzyme activity	461 N	
CYP2E1	G>T-70	(4)		-70 T	
CYP3A4	A>G-391	(4)		-391 G	

GSTM1	Del	(1)	Deletion, no enzyme activity	Del	Decreased detoxification of reactive metabolites
GSTT1	Del	(1)		Del	
GSTP1	I105V	(1)		105 V	
	A114V	(1)	114 V		
NAT2	I114T	(1)	Decreased enzyme activity	114 T	
	R197Q	(1)		197 Q	
	G286E	(1)		286 E	
DNA Repair					
BRCA2	5'UTR (A-26G)	(2)	Modulated expression	A-26	Decreased DNA repair
	N372H	(2)	Decreased enzyme activity	372 H	
XRCC1	R194W	(2)	Increased enzyme activity	R194	
	R280H	(2)		280 H	
	Q399R	(2)		399 R	
XRCC3	T241M	(2)		241 M	
XPB	K751Q	(2)	Decreased enzyme activity	751 Q	
	R156R	(2)		156 R	
APE1	D148G	(3)		148 G	
OGG1	S326C	(2)		326 C	
Oxidative Stress					
mnSOD	V16A	(3)	Decreased enzyme activity	16 A	Decreased oxidative stress defense
CAT	C>T-262	(3)	Lower catalase activity	-262 T	
NQO1	P187S	(3)	Decreased enzyme activity	187 S	
	R139W	(3)		139 W	Higher production of free radicals
MPO	G>A-463	(4)	Higher enzyme activity	G-463	
GPX1	P198L	(3)	Decreased enzyme activity	198 L	Higher oxidative stress

(1)-(4): different multiplex PCRs (See Material and Methods)

Statistical Analysis

All polymorphisms were selected based on their putative involvement in environmental carcinogenesis, and classified into different biological pathways as described previously (Table 1)(13). Corresponding risk alleles were defined based on phenotypic effects of the polymorphisms, as known from literature (Table 1). Genotypes were coded based on the number of risk alleles: homozygously carriers of alleles that were expected to enhance cancer risk were coded 2, carriers of 2 alleles with a putative protective health effect were coded 0 and heterozygous genotypes were coded 1. In case of *GSTT1* and *GSTM1*, wild-types were coded 0 and deletions were coded 2; heterozygous individuals could not be identified in this particular cohort. For each individual, a sum of risk alleles was computed for all SNPs and for different modes of action, comprising biotransformation, oxidative stress and DNA repair. Moreover, biotransformation genes were further segregated into genes involved in phase I (*CYP450s*), and in phase II (*GSTs* and *NATs*) biotransformation reactions. Difference in prevalence of these specific sums of risk alleles between the two age groups were initially examined using the Mann-Whitney U test. Subsequently, to investigate this effect in more detail, 2 subgroups were created for each specific biological pathway representing a relatively low- or high sum of risk alleles; the cut-off point was defined based on the median number of risk alleles. The prevalence of these low and high sums of risk alleles for each biological pathway were compared between the two age-groups by Chi-Square analysis using a 2x2 cross table. Finally, to investigate the contribution of single polymorphisms within each pathway, the difference in distribution of individual SNPs between the two age groups was tested by Chi-Square analysis.

All statistical analyses were conducted using SPSS for windows, version 13.0

Results

A total of 29 polymorphisms was genotyped in 822 individuals, resulting in 23,838 genotypes. The theoretical maximum sum of risk alleles present in an individual, computed from this selection of SNPs, is: 29 SNPs genotyped per individual * 2 putative risk alleles = 58 risk alleles. In the total population, the mean of the overall sum of risk alleles was 18 ± 3 with a range from 8 to 29 risk alleles. For both age groups separately, the range was 11 to 29 risk alleles (18.4 ± 3) for the 'young' and 8 to 26 risk alleles (17.6 ± 3) for the 'elderly' subpopulation respectively. The distribution of the respective sums of risk alleles did not differ between males and females, or smokers and non-smokers, in both the total population and the 'young' and 'elderly' age groups (Mann-Whitney U).

The mean sum of risk alleles for all investigated polymorphisms was found to be significantly lower in the 'elderly' compared to the 'young' population (respectively 17.6 vs 18.4 ($P=0.003$)). To investigate this in more detail, two subgroups were defined,

Chapter 6

discriminating individuals for having either a low- or a high sum of risk alleles. The cut-off point depended on the median sum of risk alleles either with respect to the total sum of all SNPs, or to individual pathways: biotransformation, phase I, phase II, oxidative stress and DNA repair.

All SNPs

Overall, for the total sum of risk alleles in all selected genes, the cut-off point was defined at 18 risk alleles. Distributions of the overall sum of risk alleles in all 29 selected SNPs in both age groups are presented in Table 2. The prevalence of the high sum of risk alleles (>18 risk alleles) was significantly less frequent in the 'elderly' (39%) as compared to the 'young' (49%) subpopulation ($P=0.005$).

Biotransformation pathway

For all polymorphisms in biotransformation-related genes, the mean number of risk alleles was significantly lower in the 'elderly' age group (5.3) as compared to the 'young' age group (5.8, $P=0.003$). Subsequently, the cut-off sum of risk alleles in biotransformation enzymes was set at 5 (range: 1-12 risk alleles). Distributions of these sums of risk alleles in both age groups are presented in Table 2. The presence of a relatively high sum of risk alleles in all selected biotransformation enzymes (>5 risk alleles) was less frequent in the 'elderly' (44%) compared to the 'young' population (54%, $P=0.002$).

The total sum of SNPs coding for biotransformation enzymes was further subdivided into alleles coding for phase I and phase II biotransformation reactions. No difference in the mean number of risk alleles was observed for phase I-related genes. For comparing the distributions of the relatively low and high sums of risk alleles between the two age groups, the cut-off point was defined at 1 high-risk allele (range: 0-4 risk alleles). As is shown in Table 2, no significant differences in prevalence of the high sum of risk alleles were found between both age groups ($P=0.147$). Regarding phase II enzymes, the mean number of risk alleles was significantly lower in the 'elderly' (4.3 risk alleles) compared to the 'young' age group (4.8 risk alleles, $P<0.001$, Figure 2). In this case, 4 risk alleles (range: 0-11 risk alleles) were used as a cut-off point. Comparison of the prevalence of the high sum of risk alleles in phase II enzymes-coding genes showed an increased frequency of high sums of risk alleles in the 'young' (56%) as compared to the 'elderly' age group (42%, $P<0.001$, Table 2).

Oxidative Stress pathway

For the oxidative stress-related genes, no difference in the mean number of risk alleles was observed between the two age groups. For the subsequent analysis of prevalences of low and high number of risk alleles, carrying 4 risk alleles (mean: 4 ± 1 , range: 0-9

Table 2. Prevalence (%) of persons with low and high sum of risk alleles in the ‘young’ and ‘elderly’ age groups for all investigated pathways (total), and for the biotransformation, phase I, phase II, oxidative stress related and DNA repair genes. The cut-off point between low and high sum of risk alleles was based on the median number of risk alleles for each specific pathway.

	Total ¹		Phase I+II ²		Phase I ³		Phase II ⁴		Oxidative Stress ⁵		DNA Repair ⁶	
	‘Young’	‘Elderly’	‘Young’	‘Elderly’	‘Young’	‘Elderly’	‘Young’	‘Elderly’	‘Young’	‘Elderly’	‘Young’	‘Elderly’
Low Sum of Risk Alleles	51	61	46	56	74	70	44	58	60	62	40	45
High Sum of Risk Alleles	49	39	54	44	26	30	56	42	40	38	60	55
Chi-Square	<i>P</i> = 0.005		<i>P</i> = 0.002		<i>P</i> = 0.147		<i>P</i> < 0.001		<i>P</i> = 0.235		<i>P</i> = 0.086	

Cut-off points: ≤18<⁽¹⁾, ≤5<⁽²⁾, ≤1<⁽³⁾, ≤4<⁽⁴⁾, ≤4<⁽⁵⁾, ≤9<⁽⁶⁾ risk alleles

Table 3. Distribution (%) of the number of risk alleles for NQO R139W, GSTP I105V, NAT2 I114T and BRCA2 5’UTR polymorphisms in the ‘young’ and ‘elderly’ age groups.

	NQO (R139W)		GSTP (I105V)		NAT2 (I114T)		BRCA2 (5’UTR)	
	‘Young’	‘Elderly’	‘Young’	‘Elderly’	‘Young’	‘Elderly’	‘Young’	‘Elderly’
No Risk Alleles	85	96	30	43	19	30	4	6
1 Risk Allele	14	3	53	47	50	48	30	38
2 Risk Alleles	1	1	17	10	31	22	66	56
Ch-Square	<i>P</i> < 0.001		<i>P</i> = 0.001		<i>P</i> < 0.001		<i>P</i> = 0.006	

Chapter 6

risk alleles) was used as cut-off point. A comparable pattern in the distribution of these sums of risk alleles was observed between both age groups as was seen in the overall sum of risk alleles and for genes involved in biotransformation, although not significantly different between the two age groups ($P=0.235$, Table 2).

DNA Repair pathway

Regarding DNA repair genes, again the mean number of risk alleles did not differ between both subpopulations. Here, genotypes at high risk were cut-off at 9 risk alleles (mean: 9 ± 2 , range: 4-14 risk alleles). No significant difference in distribution of the low- and high sum of risk alleles in DNA repair genes was found between the two age groups ($P=0.086$), although again a similar pattern as for genes involved in the biotransformation and oxidative stress pathways was observed (Table 2).

Age-related distribution of Single Genotypes

Finally, distributions of 5 genotypes in 4 different genes were found to be significantly different between the two age groups. A significantly lower prevalence in the 'elderly' compared to the 'young' subpopulation was found for the heterozygous carriers of the *NQO1* $R^{139}W$ polymorphism (3% vs 14%, $P<0.001$), *GSTP1* $I^{105}V$ heterozygous and homozygous variant individuals (47% vs 53% and 10% vs 17% respectively, $P=0.001$), homozygous carriers of the *NAT2* $I^{114}T$ polymorphism (22% vs 31%, $P=0.001$) and for *BRCA2* 5'UTR wild type individuals (56% vs 66%, $P=0.006$, Table 3). These observed genotype frequencies were in agreement with frequencies as predicted by the Hardy-Weinberg equilibrium.

Discussion

In the western population, both morbidity and mortality rates of age-related diseases such as cancer, increase rapidly between 15 and 50 years of age: cancer incidence increases from 3% at age 20 to 21% between 55 and 64 years old, whereas for instance in the US death of cancer of all sites increases from 1% in Caucasians at an age of 20 to 17% at 60 years old (14). In Flanders, similar trends have been observed (15). Furthermore, some individuals within the general population appear to be at higher risk for developing cancer after a lifetime exposure to potentially hazardous compounds to which we are exposed in our environment or as a consequence of our personal habits. Because environmental carcinogenesis is a result of an interaction between genetic and exogenous factors, these inter-individual differences in cancer risk may be explained to some degree by genetic variations between individuals in genes involved in the carcinogenic process (16). Individuals with a putatively unfavorable combination of exposure to xenobiotic carcinogens and a high number of risk alleles in these genes modifying responses to such exposures, might have an increased risk for developing cancer and

eventually mortality. This provides a rationale for the hypothesis that the prevalence of the number of risk alleles will decrease with age in a random selection of healthy individuals (Figure 1)(6,7).

In the present study conducted among a Caucasian population in Flanders, a significant decrease of the total sum of risk alleles was indeed observed in an 'elderly' population as compared to a 'young' population. A more detailed evaluation of this effect showed that high risk alleles in biotransformation genes, especially Phase II, had the largest contribution to this decrease. During aging, the overall load of toxic metabolites and adverse effects caused by reactive oxygen species accumulate in humans (17), partially due to an enduring exposure to potentially harmful xenobiotic compounds. Polymorphisms in genes involved in Phase II biotransformation of these exogenous toxicants and leading to relatively less detoxification of chemical carcinogens, are thus likely to affect human health and longevity by modifying the individual susceptibility to develop environmental cancer (8,18). The age-related decrease in the sum of risk alleles in biotransformation genes as observed in this study, and in particular those related to phase II conjugation reactions (Figure 2), supports this hypothesis. In addition, several case-control studies have reported an association between polymorphisms in CYP450, NAT and GST enzymes and various types of cancer (19-21).

Figure 2. *Distribution of the different sums of risk alleles in Phase II genes (GSTs and NATs, Table 1) in the 'young' and in the 'elderly' population. The dashed line represents the cut-off point between the relatively low (≤ 4) and high (> 4) sum of risk alleles (see text). The distribution range at the side of the relatively high sums of risk alleles (above the dashed cut-off line) is smaller, and moreover, the mean number of risk alleles is significantly lower in the 'elderly' population compared to the 'young' population ($P < 0.001$).*

Remarkably, upon restricting the statistical evaluation to the cancer patients in the 'elderly' age group ($N=19$), the distribution of a high sum of risk alleles in phase I and II enzymes appeared similar in the 'young' age group (54%) and the 'elderly' cancer patients (56%). Although this effect was not confirmed for genes involved in the

other biological pathways which we analyzed, probably due to the low N of the cancer patients, these results support the hypothesis that enzymes involved in the biotransformation of xenobiotic compounds are associated with environmental cancer development.

In the current study, assessing the distribution of individual polymorphisms in both age groups revealed a lower prevalence of polymorphisms in two phase II genes, namely *GSTP1* and *NAT2*, in the 'elderly' as compared to the 'young' age group. The phenotypic effect of both polymorphisms is associated with an increased risk for chemical carcinogenesis (22). Moreover, the results concerning *GSTP1* are in agreement with previous studies focusing on single genotypes and longevity, in which an involvement of *GSTP1* and other phase II enzymes with human longevity has been suggested (7,8,10,18). However, most of these studies focus on *CYP450* and *GST* polymorphisms, while only 2 studies can be found on *NAT2*, both describing a lack of association with human longevity (9,23). The *NAT1*^{I14}*T* polymorphism described here, is associated with a decrease in phase II enzyme activity (slow-acetylator status (24,25)) and various types of cancer, including breast (26) and bladder cancer (27). Therefore, regarding the important role of detoxification in cancer development and aging, a lower frequency of the *NAT1*^{I14}*T* polymorphism in the 'elderly' compared to the 'young' age groups might be expected.

In addition to these 2 polymorphisms, a significant decrease in prevalence of the *NAD(P)H:quinone oxidoreductase1* (*NQO1*) *R*¹³⁹*W* polymorphism and *BRCA2* 5'*UTR* wild types was observed, which are both novel findings. *NQO1* normally converts quinone to more stable hydroquinones, bypassing the production of reactive oxygen species and protecting the cell from toxic and carcinogenic effects by quinone and related chemicals (28,29). The *NQO1* *R*¹³⁹*W* polymorphism is associated with a decrease in enzyme activity (30,31) and various types of cancer (21,32-34). Moreover, in patients with intermediate malignant urinary bladder neoplasms, followed up for 5 years, individuals carrying the *NQO1* *R*¹³⁹*W* polymorphism were suggested to have a shorter disease free survival rate than non-carriers (35). Therefore, the observed decrease in prevalence of the *NQO1* *R*¹³⁹*W* polymorphism with age seems to be in concordance with the suggested crucial role of oxidative stress and associated free radical generation in degenerative disease development, such as cancer, and aging (36). Interestingly, regarding the *BRCA2* 5'*UTR* polymorphism, wild-type individuals were less frequent in the 'elderly' compared to the 'young' subpopulation. This polymorphism at position A-26G of the *BRCA2* gene has been associated with a decrease risk of breast cancer (37). Therefore, the wild-type variant of the *BRCA2* 5'*UTR* polymorphism is regarded as the risk allele (Table 1), and the observed genotype distribution for this polymorphism is as expected.

The distribution of the sum of risk alleles for genes involved in Phase I biotransformation, oxidative stress and DNA repair did not differ significantly between the age groups. Regarding the fact that DNA damage accumulates during lifetime, an effect on longevity of polymorphisms in genes involved in DNA repair would be expected

(38). Indeed, a decrease in the prevalence of high risk alleles in DNA repair enzymes was found in the 'elderly' population, although not quite statistically significant. An explanation may be, that the average age of the adult population selected in the current study may still be too young. Other studies investigating the genetics of longevity, mainly focus on centenarians, as these individuals are survivors lacking the risk alleles associated with early mortality (39). For example, it has been postulated earlier that an efficient DNA repair system is only of selective advantage, with respect to HPRT mutation frequency, at an advanced age (40). Therefore, a stronger effect of the genotype with respect to this pathway might be detected at ages higher than the upper age limit in the current study. On the other hand, the age limits used in the current study are meaningful, because it has also been suggested that early deaths and morbidity (before 50 years old) are of particular interest as they represent a large decrease in expected lifespan and quality of life (41). A stronger genetic component might be expected in these age groups as compared to older ages because of severe sickness or early death resulting from genetic diseases such as cancer, of which both morbidity and mortality rates increase rapidly with age. For example incidences of prostate and breast cancer, the most common cancers in Flanders in respectively men and women, increase rapidly after 50 years old and have been reported to be among the highest in Europe (15).

Overall, the cross-sectional design of the research presented here, may be considered as a limitation of the current study. Obviously, a longitudinal design is suggested for a better testing of the hypothesis. In this perspective, it might be useful to also include body burden of carcinogenic compounds into the study design.

In conclusion, these findings show that the number of at-risk alleles in genes, involved in the biotransformation of chemical carcinogens and in the modulation of their toxic effects, decreases with age in a randomly selected general population. Because the study population, which according to our statistical power calculation was large enough to exclude the chance of false positive findings, was selected from an identical "gene-pool", that is, individuals were selected within the same regions in Flanders and no confounding effects of race (only Caucasians were included) and gender were to be expected, the results presented in this study may provide further evidence that an at-risk genotype, coding for enzymes involved in the carcinogenic process, may modifies the risk for chemical carcinogenesis and has a selective disadvantage for human longevity.

In a way, Charles Darwin might have been the first describing human longevity and quality of life. Now, the "book of life" (42) provides insights in the mechanisms for variation and selection. Assessment of genetics in molecular epidemiology studies might lead to novel insights in these principles. Individuals who inherited a protective genotype may have an increased life expectancy with a higher quality of life compared to other individuals within the same population bearing an at-risk genotype: *survival of the fittest?*

References

1. Report on Carcinogens, Eleventh Edition; U.S. Department of Health and Human Services, Public Health Service, National Toxicology Program.
2. Kelada, S.N., Eaton, D.L., Wang, S.S., et al. (2003) The role of genetic polymorphisms in environmental health. *Environ Health Perspect*, 111, 1055-64.
3. Collins, A.R. (1999) Oxidative DNA damage, antioxidants, and cancer. *Bioessays*, 21, 238-46.
4. Forsberg, L., de Faire, U. and Morgenstern, R. (2001) Oxidative stress, human genetic variation, and disease. *Arch Biochem Biophys*, 389, 84-93.
5. Bohr, V.A. (1995) DNA repair fine structure and its relations to genomic instability. *Carcinogenesis*, 16, 2885-92.
6. Kirkwood, T. (2006) Ageing: too fast by mistake. *Nature*, 444, 1015-7.
7. Voho, A., Impivaara, O., Jarvisalo, J., et al. (2006) Distribution of glutathione S-transferase M1, P1 and T1 genotypes in different age-groups of Finns without diagnosed cancer. *Cancer Detect Prev*, 30, 144-51.
8. Taioli, E., Mari, D., Franceschi, C., et al. (2001) Polymorphisms of drug-metabolizing enzymes in healthy nonagenarians and centenarians: difference at GSTT1 locus. *Biochem Biophys Res Commun*, 280, 1389-92.
9. Muir, M.L., Verasdonck, P., Cottet, F., et al. (1998) Lack of association between human longevity and genetic polymorphisms in drug-metabolizing enzymes at the NAT2, GSTM1 and CYP2D6 loci. *Hum Genet*, 102, 526-32.
10. Christiansen, L., Brasch-Andersen, C., Bathum, L., et al. (2006) A longitudinal study of the effect of GSTT1 and GSTM1 gene copy number on survival. *Mech Ageing Dev*, 127, 597-9.
11. Garte, S., Gaspari, L., Alexandrie, A.K., et al. (2001) Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev*, 10, 1239-48.
12. Knaapen, A.M., Ketelslegers, H.B., Gottschalk, R.W., et al. (2004) Simultaneous genotyping of nine polymorphisms in xenobiotic-metabolizing enzymes by multiplex PCR amplification and single base extension. *Clin Chem*, 50, 1664-8.
13. Ketelslegers, H.B., Gottschalk, R.W., Godschalk, R.W., et al. (2006) Interindividual variations in DNA adduct levels assessed by analysis of multiple genetic polymorphisms in smokers. *Cancer Epidemiol Biomarkers Prev*, 15, 624-9.
14. Ries LAG, M.D., Krapcho M, Mariotto A, Miller BA, Feuer EJ, Clegg L, Horner MJ, Howlader N, Eisner MP, Reichman M, Edwards BK (eds) (2007) SEER Cancer Statistics Review, 1975-2004. National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2004/, based on November 2006 SEER data submission, posted to the SEER web site,.
15. E. Van Eycken, N.D.W. (2006) Wever, Cancer Incidence and Survival in Flanders, 2000-2001. Flemish Cancer Registry Network, VLK, Brussels.
16. Venitt, S. (1994) Mechanisms of carcinogenesis and individual susceptibility to cancer. *Clin Chem*, 40, 1421-5.
17. Lu, C.Y., Lee, H.C., Fahn, H.J., et al. (1999) Oxidative damage elicited by imbalance of free radical scavenging enzymes is associated with large-scale mtDNA deletions in aging human skin. *Mutat Res*, 423, 11-21.

18. Pesch, B., Dusing, R., Rabstein, S., et al. (2004) Polymorphic metabolic susceptibility genes and longevity: a study in octogenarians. *Toxicol Lett*, 151, 283-90.
19. Agudo, A., Sala, N., Pera, G., et al. (2006) Polymorphisms in metabolic genes related to tobacco smoke and the risk of gastric cancer in the European prospective investigation into cancer and nutrition. *Cancer Epidemiol Biomarkers Prev*, 15, 2427-34.
20. Hung, R.J., Boffetta, P., Brennan, P., et al. (2004) GST, NAT, SULT1A1, CYP1B1 genetic polymorphisms, interactions with environmental exposures and bladder cancer risk in a high-risk population. *Int J Cancer*, 110, 598-604.
21. Vineis, P., Veglia, F., Garte, S., et al. (2007) Genetic susceptibility according to three metabolic pathways in cancers of the lung and bladder and in myeloid leukemias in nonsmokers. *Ann Oncol*.
22. Taningher, M., Malacarne, D., Izzotti, A., et al. (1999) Drug metabolism polymorphisms as modulators of cancer susceptibility. *Mutat Res*, 436, 227-61.
23. Agundez, J.A., Rodriguez, I., Olivera, M., et al. (1997) CYP2D6, NAT2 and CYP2E1 genetic polymorphisms in nonagenarians. *Age Ageing*, 26, 147-51.
24. Blum, M., Demierre, A., Grant, D.M., et al. (1991) Molecular mechanism of slow acetylation of drugs and carcinogens in humans. *Proc Natl Acad Sci U S A*, 88, 5237-41.
25. Hickman, D., Risch, A., Camilleri, J.P., et al. (1992) Genotyping human polymorphic arylamine N-acetyltransferase: identification of new slow allotypic variants. *Pharmacogenetics*, 2, 217-26.
26. Sillanpaa, P., Hirvonen, A., Kataja, V., et al. (2005) NAT2 slow acetylator genotype as an important modifier of breast cancer risk. *Int J Cancer*, 114, 579-84.
27. Marcus, P.M., Vineis, P. and Rothman, N. (2000) NAT2 slow acetylation and bladder cancer risk: a meta-analysis of 22 case-control studies conducted in the general population. *Pharmacogenetics*, 10, 115-22.
28. Ross, D. and Siegel, D. (2004) NAD(P)H:quinone oxidoreductase 1 (NQO1, DT-diaphorase), functions and pharmacogenetics. *Methods Enzymol*, 382, 115-44.
29. Talalay, P. and Dinkova-Kostova, A.T. (2004) Role of nicotinamide quinone oxidoreductase 1 (NQO1) in protection against toxicity of electrophiles and reactive oxygen intermediates. *Methods Enzymol*, 382, 355-64.
30. Pan, S.S., Forrest, G.L., Akman, S.A., et al. (1995) NAD(P)H:quinone oxidoreductase expression and mitomycin C resistance developed by human colon cancer HCT 116 cells. *Cancer Res*, 55, 330-5.
31. Pan, S.S., Han, Y., Farabaugh, P., et al. (2002) Implication of alternative splicing for expression of a variant NAD(P)H:quinone oxidoreductase-1 with a single nucleotide polymorphism at 465C>T. *Pharmacogenetics*, 12, 479-88.
32. Cho, C.G., Lee, S.K., Nam, S.Y., et al. (2006) Association of the GSTP1 and NQO1 polymorphisms and head and neck squamous cell carcinoma risk. *J Korean Med Sci*, 21, 1075-9.
33. Lyn-Cook, B.D., Yan-Sanders, Y., Moore, S., et al. (2006) Increased levels of NAD(P)H:quinone oxidoreductase 1 (NQO1) in pancreatic tissues from smokers and pancreatic adenocarcinomas: A potential biomarker of early damage in the pancreas. *Cell Biol Toxicol*, 22, 73-80.
34. Ryk, C., Kumar, R., Sanyal, S., et al. (2006) Influence of polymorphism in DNA repair and defence genes on p53 mutations in bladder tumours. *Cancer Lett*, 241, 142-9.
35. Sanyal, S., Ryk, C., De Verdier, P.J., et al. (2007) Polymorphisms in NQO1 and the clinical course of urinary bladder neoplasms. *Scand J Urol Nephrol*, 41, 182-90.

Chapter 6

36. Beckman, K.B. and Ames, B.N. (1998) The free radical theory of aging matures. *Physiol Rev*, 78, 547-81.
37. Healey, C.S., Dunning, A.M., Teare, M.D., et al. (2000) A common variant in BRCA2 is associated with both breast cancer risk and prenatal viability. *Nat Genet*, 26, 362-4.
38. Hoeijmakers, J.H. (2001) Genome maintenance mechanisms for preventing cancer. *Nature*, 411, 366-74.
39. Lewis, S.J. and Brunner, E.J. (2004) Methodological problems in genetic association studies of longevity--the apolipoprotein E gene as an example. *Int J Epidemiol*, 33, 962-70.
40. Wilding, C.S., Rees, G.S., Relton, C.L., et al. (2006) Genotype profiles of loci encoding DNA repair enzymes in newborn and elderly populations: no evidence of association with longevity. *Biogerontology*, 7, 35-41.
41. Christensen, K., Johnson, T.E. and Vaupel, J.W. (2006) The quest for genetic determinants of human longevity: challenges and insights. *Nat Rev Genet*, 7, 436-48.
42. Venter, J.C., Adams, M.D., Myers, E.W., et al. (2001) The sequence of the human genome. *Science*, 291, 1304-51.

Chapter 7

Summary and General Discussion

Humans are continuously exposed to numerous potentially harmful pollutants, present in their daily life environment and as a consequence of hazardous individual lifestyles, with cigarette smoke being the best known example. Some of these xenobiotic compounds have been classified as carcinogenic by the International Agency of Research on Cancer (IARC). However, not every individual responds to the same degree to these exposures, and as a result, some individuals may be at higher risk for developing cancer. This variability in response can at least partly be explained by inter-individual differences in genetic makeup. Molecular epidemiology research has provided evidence that Single Nucleotide Polymorphisms (SNPs) in genes particularly involved in biotransformation of carcinogens, DNA repair and oxidative stress may modulate every step from carcinogen exposure to disease development (Chapter 1, Figure 2). The genotyping of SNPs can thus provide new mechanistic knowledge on the environmental carcinogenic process, and moreover, may provide important insights in the individual susceptibility to developing environmental health effects. Regarding the complexity of these exposures and the numerous pathways and enzymes that are involved in the biological response, it is unlikely that one single SNP is decisive in these processes. Studies investigating associations between genetic polymorphisms and disease thereby restricting analyses to single polymorphisms may therefore fail to fully explore the role of genetic variation in an individual's susceptibility for developing health effects. It is thus important to assess a relatively large number of polymorphisms in different pathways simultaneously in large populations, to obtain a better prediction of the inter-individual differences in human health risk in response to environmental exposures. This underlines the need for new, low-cost high-throughput genotyping methodologies in combination with innovative statistical approaches. It is hypothesized that, instead of looking at single SNPs in relation to a single marker, the *interaction* between multiple genes encoding for susceptibility to environmental exposure, with biomarkers of exposure and early effect, explains inter-individual differences in responses to environmental exposure and may thus provide new insights on environmental health risk assessment.

The development and validation of a cost-effective and rapid method for genotyping multiple SNPs in large populations has been described in **Chapter 2** of this thesis. This technique, based on Single Base Extension (SBE) by SNaPShot technology, was firstly developed for genotyping nine SNPs in five xenobiotic metabolizing enzymes. For validation, 67 lymphocyte DNA samples from healthy Caucasian volunteers were genotyped using the described assay, which resulted in accurate and complete genotypes with frequencies as expected from literature. Twenty DNA samples were genotyped in duplicate to test for inter-assay variation, and thoroughly consisting results were obtained. Further validation was conducted by genotyping at least 10 different samples with conventional methods (Restriction Fragment Length Polymorphism (RFLP) and sequencing), resulting in 100% concordance between the SBE assay and the conventional methods. In addition, data showed that for *NAT2*5* genotyping, the SBE assay is more accurate than the conventionally used RFLP analysis. Moreover, this assay can also be applied for *GSTM1* and *GSTT1* gene deletions, although no distinction

could be made between heterozygous deletions and homozygous positive genotypes. For an optimal use of the high-throughput capabilities of the SNaPShot technology, a multiplex PCR-based amplification, initially for nine SNPs in five different genes, was developed. Genotype and phenotype information, initially for these 9 SNPs, together with PCR and SBE primer sequences, was collected in a SNP database. This database was further completed to a total of 93 SNPs in 53 genes divided over 10 series which are easily applicable in the multiplex genotyping assay (multiplex capability of up to 13 SNPs in 96 samples simultaneously in one single experiment). All SNPs were selected based on known effects on enzyme activity as well as on health effects, including carcinogenesis (Chapter 1, Figure 2).

A first application study was conducted for 3 of these sets of polymorphisms. In this study, described in **Chapter 3**, 63 healthy Caucasian smoking volunteers were genotyped for 19 SNPs in 12 genes involved in carcinogen metabolism, DNA repair and oxidative stress defence, to evaluate the hypothesis that the analysis of multiple SNPs simultaneously leads to a better prediction of the inter-individual variations in DNA adduct levels in response to cigarette smoke exposure. The total sum of putatively high risk alleles within this selection positively correlated with DNA adduct levels. This relationship was enriched, when focusing on the sum of risk alleles of the *GSTM1 null*, *mEH Y¹¹³H* and *GPX1 P¹⁹⁸L* genotypes, which were identified as the most relevant polymorphisms for the modulation of bulky DNA adduct levels in response to cigarette smoke exposure. Smoking individuals carrying these at-risk genotypes are more susceptible for developing DNA adducts in response to harmful cigarette smoke constituents, which is in agreement with the expected phenotypic effect of the polymorphisms. Regarding the fact that in the conventionally used univariate analyses, in which the effect of each SNP separately on variations in DNA adduct levels is assessed, only one polymorphism (*GSTM1 null*) was found to be significant, these results of the multivariate analysis indicate that the assessment of multiple polymorphisms simultaneously is important to obtain better insights in biomonitoring studies investigating the inter-individual variations in exposure-effect relationships. For the purpose of investigating the effect of multiple genotypes on the exposure-effect relationship, a novel statistical approach was developed which depends on the formation of subgroups based on the regression line. In this approach, the observed response level is divided by the expected response, computed from the regression line. A ratio of more than 1 (above the regression line) comprises the more susceptible individuals, or so called high responders, whereas a ratio below 1 indicates a protective effect and comprises the low responders. Using this ratio, the study population is divided into subgroups and the number of created subgroups determines the statistical method for discrimination of the individuals into these groups, based on their genotypic differences. In this study, 3 subgroups were formed (low, medium and high responders) and classification of each individual into one of the three subgroups was conducted by discriminant analysis, based on gender, age, cigarettes per day and genotype, revealing the most important predictors. This protocol can easily be adapted, for example for classification by binary logistic regres-

sion into two subgroups: below (protective) and above (more susceptible) the average exposure-effect relationship, as used in the studies described in **Chapters 4** and **5**.

Besides lymphocytic DNA adduct levels, protein adducts such as 4-aminobiphenyl-hemoglobin (4-ABP-Hb) adducts are also commonly analyzed for biological (effect) monitoring of smoking behavior. Good correlations have been shown between cigarette smoke exposure and 4-ABP-Hb adduct levels, with a saturation effect at high smoking dose, and moreover, a large extent of inter-individual variation in exposure-effect relationships. Both phenomena were suggested to be due to genetic differences between individuals in enzymes involved in the biotransformation of 4-ABP. The metabolic transformation of 4-ABP is mainly conducted by cytochrome P450's (CYP450) in the liver. It was generally accepted that CYP1A2 is the primary responsible enzyme for this oxidation of 4-ABP, however in more recent studies, this role of CYP1A2 in the metabolic activation of 4-ABP has been challenged, suggesting that other (CYP450) enzymes may be involved. Therefore, a study was conducted to identify crucial genetic mediators in 4-ABP metabolism and their putatively effect on the exposure-effect relationship between smoking and 4-ABP-Hb adduct levels (**Chapter 4**). To this end, the statistical approach described in Chapter 3 was applied, in this case using two subgroups for classification: above (more susceptible) and below (protective) the observed regression line between cigarette dose and 4-ABP-Hb adduct levels. This revealed a crucial role for CYP1B1 in 4-ABP metabolism. In particular, the *CYP1B1* $L^{432}V$ polymorphism was found to modulate the dose-response relationship between cigarette smoke exposure and 4-ABP-Hb adduct levels, revealing a saturation effect at high doses, but limited to individuals bearing the high activity *CYP1B1* $L^{432}V$ genotype (homozygous wild-type and heterozygous individuals). In case of low activity homozygous carriers of the *CYP1B1* $L^{432}V$ polymorphism, no such plateau was reached, and overall adduct levels were lower as compared to homozygous wild-type and heterozygous individuals. These results confirm previous hypotheses, that the observed saturation effect in the dose-response relationship between cigarette smoke exposure and protein adduct formation may be explained by different metabolic routes. Furthermore, also in this study, no involvement of CYP1A2 and the phase II enzymes belonging to the glutathione-S-transferase (GST) and *N*-acetyltransferase (NAT) families was observed, which have all been investigated previously in relation to 4-ABP metabolism with contradictory results. It is concluded from these data, that CYP1B1 plays an important role in 4-ABP metabolism. Moreover, the assessment of multiple polymorphisms simultaneously may provide valuable information on the involvement of certain enzymes in toxicologically relevant pathways, and provides further evidence that it is applicable as a biomarker for susceptibility in biomonitoring studies.

The application of genotyping SNPs in large-scale biological monitoring, using the developed techniques and knowledge obtained in the studies described in the previous Chapters, is described in **Chapter 5**. This biomonitoring study was conducted within the framework of the Flanders Environment and Health Study (FLEHS), which had been assigned by the Flemish Government, and aimed to obtain a better perspective

on health risks associated with exposure to harmful environmental compounds among inhabitants of different regions of Flanders. Because individuals within a certain population possess different susceptibilities to a specific dose of an environmental compound, environmental health policy makers request more information on the sources of inter-individual variations in response to environmental stress. The study presented in this Chapter aimed to describe these variations due to genetic susceptibility. Amongst others, three remarkable exposure-effect relationships were observed and described for the first time in this study. An increase in oxidative DNA damage levels (8-OHdG) in response to ethylbenzene exposure was observed in the adolescent population. Within this relationship, more susceptible subgroups could be defined based on the *GSTT1 null* and *CAT C²⁶²T* genotypes, showing a respectively 1.2 and 1.4 times increase in DNA damage levels as compared to wild-type individuals. In the adult population, a relationship between cadmium exposure and serum levels of the tumor marker p53 was observed. Highest p53 levels in response to cadmium exposure were observed in individuals carrying the *GSTM1 wild type* (1.1 times higher than in *GSTM1 null* individuals) and *BRCA2 N³⁷²H* genotypes (individuals carrying this genotype had respectively 1.3 and 1.4 times higher compared to wild-type and heterozygous carriers). Lastly, an increase in lead exposure was found to be associated with an increase in the levels of another tumor marker, CEA. More susceptible subgroups within this relationship could be defined based on the *BRCA2 N³⁷²H* genotype and cigarette smoking, which increased the CEA levels in response to lead exposure almost 2 and 3-fold respectively. In conclusion, the research presented in this chapter resulted in an innovative study on biomarkers for genotoxic risks, providing new insights in exposure-effect relationships in adolescent and adult individuals exposed to several environmental pollutants and moreover, showing an important role of different genotypes within these relationships enabling the identification of more susceptible subgroups.

Additional evidence regarding the involvement of an at-risk genotype in environmental health risk susceptibility is presented in **Chapter 6**. The hypothesis was tested that if an unfavourable combination of environmental (carcinogen) exposures and an at risk genotype involved in environmental carcinogenesis is indeed a true risk factor for cancer development and even mortality, it will affect human life span and it can be postulated that the prevalence of an carcinogenesis related at-risk genotype will decrease with age in a randomly selected healthy population. A significant decrease in the total sum of risk alleles in enzymes involved in the biotransformation of xenobiotic compounds, oxidative stress defence and DNA repair was indeed observed in a randomly selected 'elderly' (individuals above 50 years old) compared to a 'young' (individuals 14-16 years old) age group. Moreover, enzymes involved in xenobiotic metabolism, and in particular those involved in phase II detoxification reactions, contributed most to this decrease in the sum of risk alleles. Regarding the high power of the conducted study, minimizing the chance for false positives, and the fact that the study population was selected from an identical "gene-pool" (no effect of race (only Caucasians were included) and gender might be expected) within the same regions in Flanders, it is

concluded that the results of this study provide further evidence that a carcinogenesis related at-risk genotype may modify the risk for environmental cancer development and has a selective disadvantage on human longevity.

General Discussion

Based on knowledge provided by molecular epidemiology studies over the last two decades, showing that diseases such as cancer may be a result of an interaction between largely preventable hazardous environmental exposures and genetic factors, research interest has shifted slowly towards cancer prevention (1). Two levels of prevention have been suggested, namely (i) the adaptation of hazardous lifestyles of individuals and (ii) the reduction of involuntary exposures to carcinogens at population level, through legislations. Because not all individuals respond in the same way to these exposures, policy makers are considering the possibility to account for these inter-individual variations in susceptibility. In this thesis, inter-individual susceptibility has been assessed by investigating the effect of genetic differences between individuals on exposure-effect relationships. The harmful effect of cigarette smoking, the most well known example of a hazardous individual lifestyle, has been investigated in Chapters 3 and 4. In both chapters, good correlations were observed between cigarette dose and the investigated effect marker (DNA (Chapter 3) and protein adducts (Chapter 4)). Moreover, the genotype was shown to affect both exposure-effect relationships, implying that more susceptible subgroups can be identified. More susceptible subgroups for developing health effects in response to various environmental pollutants, were also identified at population level, as described in Chapter 5. However, these exposure-effect relationships were not as strong as those seen with respect to cigarette smoke exposure.

One important difference between these results is, that the exposure-effect relationships presented in Chapter 5 were related to single (carcinogenic) exposures, whereas the focus in both Chapter 3 and 4 was on a mixture of various (low dose) carcinogens as present in cigarette smoke (2, 3). Regarding the carcinogenic or genotoxic potency of these exposures, it was postulated earlier that the interaction between low dose carcinogens, for example the joint effect of genotoxic agents and promoters in cigarette smoke, may be more important than single compounds (4). Therefore, the increase in effect in response to these exposures may be much stronger in the case of exposure to cigarette smoke exposure compared to single agents. An alternative explanation was also proposed, in which it was postulated that genetic differences between individuals might explain the lack of effect in response to single carcinogen exposures (4). Indeed, an effect of several polymorphisms was observed on the exposure-effect relationships in Chapter 5, showing a stronger effect in individuals carrying a certain at-risk genotype in response to specific (carcinogenic) exposures. However, the levels of these effects are still relatively low (in most cases only between 1 and 1.5 times difference between the respective genotypes). It is suggested that the effects of these polymorphisms on the dose response relationships will be higher when focusing on the ef-

fect in response to different (low dose) exposures simultaneously, as these represent more accurately the daily life situation (with the exception of occupational exposures to high doses of particular (carcinogenic) single exposures).

Another important difference is, that the cigarette smoke exposures in Chapter 3 and 4 are relatively high, whereas all environmental exposures in Chapter 5 are very low. Therefore, health effects induced by these environmental compounds may be difficult to detect, although the exposure-effect relationships that were still observed despite of these low levels were highly significant. In all relationships a relatively small R-square was observed (data not shown), indicating large variations around the regression line. The hypothesis that at least part of the inter-individual variation in the observed exposure-effect relationships could be explained by the genotype was confirmed. In addition, it has been postulated that genetic susceptibility may be more important at low exposure doses (4). From a metabolic point of view, this effect of the genotype at low levels of exposure may be explained by saturation of the relevant enzymes at high doses. This was clearly shown in Chapter 4, in which the impact of the *CYP1B1* L⁴³²V polymorphism was diminished at a high level of exposure.

One limitation of the research presented in the current thesis is the selection of polymorphisms. It was aimed to select the most important genes in the most common toxicologically relevant pathways based on literature data. Obviously, this is somewhat arbitrary. As high-throughput genotyping assays are still rapidly evolving, eventually whole genome scans will become feasible. A major tool in this is the publication of the haplotype map, which reduces the number of common SNPs for the investigation of the entire genome dramatically (5). However, investigating the complete genome also leads to large amounts of data, the majority of which may be biologically irrelevant. It is postulated, that with respect to the study field of *toxicogenetics*, genes participating in the most common pathways involved in chemical carcinogenesis have been investigated in this thesis. In addition, the results of the current thesis confirm that xenobiotic metabolizing enzymes play an important role in environmental health risk susceptibility (6, 7), as they were involved in almost every investigated exposure-effect relationship in the studies described in this thesis. Nonetheless, with respect to carcinogenesis, polymorphisms in genes involved in cell cycle control and apoptosis, and those involved in the development and regulation of the immune system may be additionally included in future studies (8). Moreover, it may be useful for future studies to assess the effect of haplotypes of the selected genes rather than single SNPs, to fully characterize the functional effects of the combined variants.

Besides the selection of SNPs, the number of samples to be analyzed in bio-monitoring studies is an important issue. Although the study populations in the current thesis were overall sufficiently large with respect to the number of investigated variables, a larger sample size will definitely increase the statistical power of the conducted analyses.

Chapter 7

With respect to the policy-supporting relevance of the research presented in this thesis, it is important to consider differences in safety assessment procedures between carcinogenic and non-carcinogenic compounds. Regarding non-carcinogens and also non-genotoxic carcinogens, the concept of a so called “zero tolerance risk model” is crucial. In this model, a threshold dose (the dose at and below which no toxic effect is detected) is used to compute the Acceptable Daily Intake (ADI). Within this ADI, a safety factor 10 is always included to account for the possible inter-individual variations in susceptibility (9). However, all investigated biomarkers described in this thesis represent exposures which are classified as (possible) carcinogenic to humans by the IARC. These exposures are regulated in a more stringent way. The extremely conservative approach of a linear non-threshold high-to-low dose extrapolation is applied. In this case a threshold dose is not applicable, based on the theory that each single carcinogenic molecule can cause cancer and every individual is exposed to some level. In this perspective, no correction factor for inter-individual variation is applied, because of the default assumption that the conventional model is conservative enough and thus, limits exposure levels to (genotoxic) carcinogens are set sufficiently low. However, this linear high-to-low dose extrapolation has been questioned, based on evidence from epidemiologic studies suggesting an interaction between polymorphisms involved in carcinogenesis and environmental factors, discriminating certain subpopulations at higher risk for cancer (4). The results of the current thesis show that in all investigated exposure-effect relationships more susceptible subgroups could be identified, showing an increased response in levels of biomarkers for genotoxic damage and thus, having a higher risk for developing health effects in response to environmental genotoxic exposures. Although results presented in this study should be interpreted carefully as they comprise biomarkers instead of actual cancer incidence, nonetheless, it is concluded that it is of social importance to continue the debate about the relevance of inter-individual differences in susceptibility with regard to environmental cancer risk assessment.

Conclusion

The research presented in this thesis provides a rapid, accurate and thoroughly validated method for genotyping multiple SNPs in large populations, easily applicable as a new biomarker for genetic susceptibility in molecular epidemiology studies. A strong aspect of the research presented in this thesis is the large number of genotypes, in pathways of interest, that can be assessed (ranging from 684 in Chapter 4 to 24,986 in Chapter 5), in contrast to the relatively low number that was feasible using conventional methodologies such as RFLP. A second aspect is that the impact of multiple polymorphisms is simultaneously investigated on the relationship between exposure- and effect markers, where conventional methods focus on the effect of a single polymorphism on a specific biomarker. This was shown by the results obtained from the developed multivariate statistical methodology, which provided additional information as compared to the outcome of conventional univariate analyses. For future studies, it is therefore recom-

mended to emphasize integrative effects of multiple polymorphisms on gene-environment interactions.

References

1. Perera, F. P. Molecular epidemiology: insights into cancer susceptibility, risk assessment, and prevention. *J Natl Cancer Inst*, 88: 496-509, 1996.
2. Smith, C. J., Perfetti, T. A., Garg, R., and Hansch, C. IARC carcinogens reported in cigarette mainstream smoke and their calculated log P values. *Food Chem Toxicol*, 41: 807-17, 2003.
3. Hecht, S. S. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst*, 91: 1194-210, 1999.
4. Vineis, P. Molecular epidemiology: low-dose carcinogens and genetic susceptibility. *Int J Cancer*, 71: 1-3, 1997.
5. The International HapMap Project. *Nature*, 426: 789-96, 2003.
6. Autrup, H. Genetic polymorphisms in human xenobiotica metabolizing enzymes as susceptibility factors in toxic response. *Mutat Res*, 464: 65-76, 2000.
7. Bartsch, H., and Hietanen, E. The role of individual susceptibility in cancer burden related to environmental exposure. *Environ Health Perspect*, 104 Suppl 3: 569-77, 1996.
8. Brennan, P. Gene-environment interaction and aetiology of cancer: what does it mean and how can we measure it? *Carcinogenesis*, 23: 381-7, 2002.
9. Knudsen, L. E., Loft, S. H., and Autrup, H. Risk assessment: the importance of genetic polymorphisms in man. *Mutat Res*, 482: 83-8, 2001.

Samenvatting en algemene discussie

Mensen worden in hun dagelijks leven continu blootgesteld aan stoffen die mogelijk schadelijk zijn voor hun gezondheid, aanwezig in het milieu of als gevolg van riskante individuele leefgewoonten, roken is hiervan het meest bekende voorbeeld. Een deel van deze lichaamsvreemde stoffen zijn geclassificeerd als carcinogeen door het 'International Agency of Research on Cancer' (IARC). Echter, niet iedereen reageert op dezelfde manier op deze blootstellingen, waardoor sommige mensen een verhoogd risico hebben op het krijgen van kanker. Dit verschil in gevoeligheid kan, in ieder geval deels, worden verklaard door inter-individuele verschillen tussen mensen in genetische achtergrond. Uit moleculair epidemiologische studies blijkt, dat "Single Nucleotide Polymorphisms" (SNPs) in genen die betrokken zijn bij de biotransformatie van carcinogenen, DNA repair en oxidatieve stress, iedere stap vanaf carcinogene blootstelling tot en met de uiteindelijke ontwikkeling van de ziekte kunnen beïnvloeden (Hoofdstuk 1, Figuur 2). Het genotyperen van SNPs kan daarom waardevolle kennis opleveren met betrekking tot de mechanistische achtergrond van het carcinogene proces en daarnaast belangrijke inzichten verschaffen in de individuele gevoeligheid voor het ontwikkelen van milieugerelateerde gezondheidseffecten. Omdat de blootstelling aan deze stoffen vaak erg complex is en gezien de vele pathways en enzymen die betrokken zijn in de biologische respons, is het onwaarschijnlijk dat één enkele SNP in één bepaald gen doorslaggevend is in deze processen. Daarom zullen studies die de associatie onderzoeken tussen polymorfismen en het ontstaan van ziekte, en die zich daarbij richten op de analyse van individuele polymorfismen, de rol van genetische variatie in de individuele gevoeligheid voor het ontwikkelen van gezondheidseffecten nooit volledig in kaart kunnen brengen. Het is dus van belang om een relatief groot aantal polymorfismen in verschillende biologische pathways tegelijkertijd te onderzoeken in grote populaties, om meer inzicht te krijgen in de verschillen tussen mensen met betrekking tot het ontwikkelen van gezondheidseffecten na blootstelling aan schadelijke stoffen. Dit onderstreept de vraag voor nieuwe, kost-effectieve 'high throughput' genotyperings-methoden in combinatie met innovatieve statistische benaderingen. De hypothese van het onderzoek beschreven in dit proefschrift is, dat de interactie tussen een veelvoud van genen, die de gevoeligheid voor schadelijke stoffen uit het milieu kunnen beïnvloeden, met biomarkers van blootstellingen en vroeg biologisch effect, de verschillen tussen mensen in hun respons op blootstelling aan schadelijke stoffen kan verklaren en daarbij nieuwe inzichten kan verschaffen voor milieugerelateerde gezondheidsrisico schatting.

De ontwikkeling en validatie van een snelle, rendabele methode voor de genotypering van multiële SNPs in grote populaties wordt beschreven in **Hoofdstuk 2**. Deze techniek, die is gebaseerd op 'Single Base Extension' (SBE) door middel van de SNaPSHOT techniek, werd allereerst opgezet voor het genotyperen van 9 SNPs in 5 biotransformatie enzymen. Ter validatie werden 67 DNA samples van gezonde vrijwilligers gegenotypeerd door middel van de beschreven assay, wat resulteerde in accurate en complete genotypes met frequenties die verwacht werden op basis van de literatuur. Twintig DNA samples werden in duplo gegenotypeerd om te testen voor inter-assay variatie, wat zeer consistente resultaten opleverde. Verdere validatie werd gedaan door

op z'n minst 10 samples te genotyperen met conventionele methoden ('Restriction Fragment Length Polymorphism' (RFLP) en Sequencing) wat resulteerde in 100% overeenstemming tussen de SBE assay in de conventionele methoden. Daarnaast blijkt uit de data dat de SBE assay voor *NAT2*5* genotypering nauwkeuriger is dan de traditionele RFLP analyse. Verder kan deze assay ook gebruikt worden voor de genotypering van *GSTM1* en *GSTT1* gen deleties, alhoewel er geen verschil gemaakt kan worden tussen heterozygote deleties en homozygoot positieve genotypes. Voor een optimaal gebruik van de high-throughput mogelijkheden van de SNaPShot techniek, werd een multiplex PCR amplificatie van, in eerste instantie, 9 SNPs in 5 verschillende genen ontwikkeld. Zowel genotype en fenotype informatie, in eerste instantie voor deze 9 SNPs, samen met PCR en SBE primer sequenties, werden verzameld in een SNP database. Deze database is inmiddels verder aangevuld tot 93 SNPs in 53 genen verdeeld over 10 series, welke makkelijk toepasbaar zijn in de multiplex genotyperings-assay (met een multiplex capaciteit tot 13 SNPs in 96 samples tegelijkertijd in één enkel experiment). Alle SNPs werden geselecteerd op hun effect op enzym activiteit en gezondheidseffecten, met de nadruk op carcinogenese (Hoofdstuk 1, Figuur 2).

Een eerste toepassingsstudie werd gedaan met 3 van deze sets polymorfismen. In deze studie, beschreven in **Hoofdstuk 3**, werden 63 gezonde Kaukasische rokende vrijwilligers gegenotypeerd voor 19 SNPs in 12 genen betrokken in carcinogeen metabolisme, 'DNA repair' en 'oxidatieve stress defence'. Hierbij werd de hypothese getest dat de analyse van multiële SNPs tot een betere voorspelling leidt van de inter-individuele variaties in DNA adduct niveaus als gevolg van blootstelling aan sigarettenrook. De totale som van risico allelen binnen deze selectie van SNPs was positief gecorreleerd met DNA adduct niveaus. Deze relatie werd nog beter wanneer werd gekeken naar de som van risico allelen voor de *GSTM1 null*, *mEH Y¹¹³H* en *GPX P¹⁹⁸L* genotypes, die als meest belangrijke polymorfismen voor de modulatie van bulky DNA adducten in respons op sigarettenrook waren geïdentificeerd. Rokers met deze zogenaamde 'at risk' genotypen zijn gevoeliger voor het ontwikkelen van DNA adducten als gevolg van blootstelling aan schadelijke stoffen in sigarettenrook, wat overeenkomt met de fenotypische effecten van de polymorfismen. Omdat de conventionele statistische analyses, waarin het effect van één enkele SNP op de variaties in DNA adduct niveaus wordt onderzocht, slechts één significant polymorfisme (*GSTM null*) opleverde, impliceren deze resultaten dat de simultane analyse van multiële polymorfismen belangrijk is om betere inzichten te verschaffen in biomonitoring studies die de inter-individuele variaties in blootstelling-effect relaties onderzoeken. Voor deze laatstgenoemde toepassing in biomonitoring studies, werd een nieuwe statistische benadering ontwikkeld die gebaseerd is op de formatie van subgroepen om de regressie lijn. In deze benadering wordt de geobserveerde waarde van het effect gedeeld door de verwachte waarde -berekend aan de hand van de regressie lijn. Een ratio boven de 1 (boven de regressie lijn) correspondeert met de personen die een hogere gevoeligheid hebben, ofwel de 'high responders', en een waarde beneden de 1 impliceert een meer beschermend effect en verwijst naar de low responders. Door middel van deze ratio wordt de studie populatie in

subgroepen verdeeld, waarbij het aantal gemaakte subgroepen de statistische methode bepaald voor discriminatie van de personen naar een van deze groepen, gebaseerd op hun verschillende genotypes. In deze studie werden 3 subgroepen gevormd ('low', 'medium' en 'high responders') waardoor classificatie van ieder individu naar een van deze groepen werd uitgevoerd door middel van discriminant analyse, met geslacht, leeftijd, het aantal sigaretten gerookt per dag en het genotype als onafhankelijke variabelen. Dit protocol kan makkelijk aangepast worden, bijvoorbeeld voor classificatie naar 2 groepen door middel van binaire logistische regressie, wat werd toegepast in de studies beschreven in **Hoofdstuk 4** en **5**.

Naast DNA adduct levels uit lymfocyten, worden eiwit adducten zoals 4-AminoBiPhenyl-hemoglobine (4-ABP-Hb) adducten ook regelmatig gebruikt als biologische (effect) marker voor roken. In de literatuur zijn goede correlaties beschreven tussen blootstelling aan sigarettenrook en 4-ABP-Hb adduct niveaus, met een saturatie effect bij hoge doses. Binnen deze relaties zijn grote inter-individuele variaties aanwezig. Het werd gesuggereerd dat beide observaties te verklaren zouden kunnen zijn aan de hand van genetische verschillen tussen personen in enzymen die betrokken zijn bij het metabolisme van 4-ABP. De grootste rol binnen de metabole transformatie van 4-ABP is weggelegd voor cytochroom P450 (CYP450) enzymen in de lever. Men is er altijd vanuit gegaan dat CYP1A2 het primaire enzym is voor deze oxidatie van 4-ABP, maar in meer recente studies is de rol van dit enzym in twijfel getrokken en wordt gesuggereerd dat andere enzymen betrokken zijn. Daarom hebben we een studie uitgevoerd om de belangrijkste genetische mediators te vinden in het metabolisme van 4-ABP en daarnaast hun mogelijk effect op de blootstelling-effect relatie tussen sigarettenrook en 4-ABP-Hb adduct niveaus aan het licht te brengen (**Hoofdstuk 4**). Hiervoor werd de statische benadering, beschreven in hoofdstuk 3, toegepast, ditmaal met 2 groepen voor classificatie: boven (meer gevoelig) en beneden (beschermend) de gevonden regressie lijn tussen sigarettenrook dosis en 4-ABP-Hb adduct niveaus. Het *CYP1B1* $L^{432}V$ polymorfisme had een cruciaal effect op de blootstelling-effect relatie, leidend tot een saturatie effect bij mensen die het polymorfisme dragen wat codeerd voor een hogere activiteit van het enzym (de homozygoot wild-typen en heterozygoten). In het geval van homozygote dragers van de *CYP1B1* $L^{432}V$ variant, leidend tot een enzym met een lagere activiteit, werd er geen plateau gevonden en waren de adduct levels in het algemeen lager dan bij de homozygoot wild-typen en heterozygoten. Deze resultaten bevestigen de eerdere hypothese dat het geobserveerde saturatie effect in de regressie lijn tussen sigarettenrook blootstelling en 4-ABP-Hb adduct niveaus, verklaard kan worden door verschillende metabole routes. Verder werd er ook in deze studie geen effect gevonden van CYP1A2 en fase 2 enzymen van de Glutathione-S-Transferase (GST) en *N*-AcetylTransferase (NAT) families, welke al eerder onderzocht zijn in dit verband, resulterend in tegenstrijdige bevindingen. Concluderend blijkt uit deze data dat CYP1B1 een belangrijke rol speelt in 4-ABP metabolisme. Daarnaast kan het simultaan analyseren van multiple polymorfismen belangrijke informatie opleveren over bepaalde enzymen

in toxicologisch relevante pathways en is het toe te passen als een biomarker voor gevoeligheid in biomonitoring studies.

Het genotyperen van SNPs toegepast in grootschalige biomonitoringstudies, met de ontwikkelde technieken en verkregen kennis uit de studies in de vorige hoofdstukken, wordt beschreven in **hoofdstuk 5**. Deze biomonitoring studie werd uitgevoerd binnen het framework van de Flanders Environment and Health Study (FLEHS), in opdracht van het Vlaams ministerie, met als doel betere inzichten te krijgen in de gezondheidsrisico's die gerelateerd zijn aan de blootstelling aan schadelijke stoffen in verschillende regio's in Vlaanderen. Omdat er grote verschillen bestaan in gevoeligheid voor een specifieke dosis van een toxische stof tussen personen binnen een bepaalde populatie, vragen beleidsmakers meer informatie over deze inter-individuele verschillen in humane repons op milieublootstellingen. Het doel van de studie beschreven in dit hoofdstuk was om deze verschillen in kaart te brengen, op basis van genetische gevoeligheid. Binnen de geobserveerde blootstelling-effect relaties waren 3 opmerkelijke resultaten, voor de eerste keer beschreven in deze studie. In de adolescenten populatie, werd een positieve relatie gevonden tussen blootstelling aan ethylbenzeen en oxidatieve DNA schade (8-OHdG). Meer gevoelige groepen binnen deze relatie konden worden gedefinieerd op basis van de *GSTT1 null* en *CAT C-²⁶²T* genotypes, die geassocieerd waren met respectievelijk 1.2 en 1.4 keer verhoogde niveaus van DNA schade vergeleken met de wild-types. In de volwassenen werd een positief verband gevonden tussen blootstelling aan Cadmium en serum levels van de tumor marker p53. De hoogste serum p53 niveaus werden gevonden in personen met het *GSTM1 wild-type* (1.1 keer zo hoog dan in personen met de *GSTM1 null* variant) en het *BRCA2 N³⁷²H* genotype (personen met dit genotype hadden respectievelijk 1.3 en 1.4 keer hogere niveaus dan wild-types en heterozygoten). Tenslotte werd er een relatie gevonden tussen lood blootstelling en een verhoging van een andere tumor merker, CEA. Groepen met een verhoogde gevoeligheid werden binnen deze relatie gevonden op basis van het *BRCA2 N³⁷²H* genotype en sigaretten rokers; respectievelijk bijna 2 en 3 keer hogere CEA levels in relatie met lood blootstelling. Concluderend kan worden gesteld dat het onderzoek gepresenteerd in dit hoofdstuk heeft geresulteerd in een innovatieve studie naar biomarkers voor genotoxisch risico, wat nieuwe inzichten verschaft in blootstelling-effect relaties in adolescenten en volwassenen blootgesteld aan verschillende schadelijke milieu polluenten. Daarnaast blijkt uit de resultaten dat er een belangrijke rol is voor verschillende genotypes binnen deze relaties, wat de identificatie van meer gevoelige subgroepen mogelijk maakt.

Bijkomend bewijs voor de invloed van een 'at risk' genotype op een verhoogde gevoeligheid voor gezondheidsrisico's door milieublootstellingen wordt gepresenteerd in **Hoofdstuk 6**. Hier werd de hypothese getest dat, als een onvoordelige combinatie van milieu (carcinogene) blootstellingen en een 'at risk' genotype, coderend voor enzymen betrokken bij chemische carcinogenese, daadwerkelijk een risico factor is voor het ontwikkelen van kanker en zelfs sterfte, dan zal het de levensverwachting van mensen beïnvloeden en kan worden gesteld dat de prevalentie van een carcinogenese rela-

teerd 'at risk' genotype verminderd met de leeftijd in een willekeurige selectie van gezonde personen. De totale som van risico allelen voor enzymen betrokken in biotransformatie, oxidatieve stress en DNA repair, was inderdaad significant lager in een willekeurige selectie ouderen (boven de 50 jaar) ten opzichte van een populatie jongere personen (tussen 14 en 16 jaar). Biotransformatie enzymen, en met name de enzymen die betrokken zijn in fase 2 detoxificatie reacties, droegen het meeste bij aan deze afname in de som van risico allelen. Gezien de hoge statistische power van de studie, wat de kans op vals positieven minimaliseert, en het feit dat de studiepopulatie geselecteerd was uit een identieke 'gene pool' (géén effect van geslacht en ras) binnen dezelfde regio's in Vlaanderen, kan worden geconcludeerd dat de resultaten van deze studie nieuw bewijs leveren dat een carcinogenese-gerelateerd 'at-risk' genotype een invloed heeft op een verhoogd risico voor chemische carcinogenese en een selectief nadeel heeft op de humane levensverwachting.

Discussie

Moleculair epidemiologische studies over de laatste 2 decennia hebben uitgewezen dat ziektes zoals kanker voortkomen uit een interactie tussen grotendeels te voorkomen schadelijke milieublootstellingen en genetische factoren. Daarom is het onderzoek naar kanker deels verschoven naar kanker preventie (1). Twee niveaus van preventie zijn voorgesteld, namelijk (i) het aanpassen van schadelijke leefgewoonten van mensen en (ii) de reductie van onvrijwillige blootstelling aan carcinogenen op een populatie niveau, door middel van beleidsvorming. Omdat niet alle mensen hetzelfde reageren op deze blootstellingen, bekijken beleidsmakers de mogelijkheden om rekening te houden voor deze inter-individuele variaties in gevoeligheid. In dit proefschrift werd de inter-individuele gevoeligheid bepaald, door het effect van genetische verschillen tussen personen te onderzoeken op verschillende blootstelling-effect relaties. Het schadelijke effect van roken, de meest bekende vorm van een gezondheidsbedreigende individuele leefgewoonte, werd onderzocht in hoofdstuk 3 en 4. In beide hoofdstukken werden goede positieve correlaties gevonden tussen blootstelling aan sigarettenrook en DNA (hoofdstuk 3) en eiwit (hoofdstuk 4) schade. Bovendien bleek het genotype een effect te hebben op beide relaties, zodat meer-gevoelige subgroepen konden worden geïdentificeerd. Meer gevoelige groepen met een hoger risico op nadelige gezondheidseffecten konden ook worden geïdentificeerd in milieublootstelling-effect relaties op populatie niveau, wat werd beschreven in hoofdstuk 5. Deze relaties waren echter niet zo sterk als de relaties met blootstelling aan sigarettenrook.

Een belangrijk verschil tussen beide resultaten is, dat de blootstelling-effect relaties die gepresenteerd werden in hoofdstuk 5 zich richtten op individuele carcinogene blootstellingen, terwijl de nadruk in hoofdstuk 3 en 4 lag op een mix van verschillende (lage dosis) carcinogenen, die voorkomen in sigarettenrook (2,3). Gekeken naar de carcinogene of genotoxische capaciteiten van deze blootstellingen, is gesuggereerd dat de interactie tussen verschillende lage dosis-carcinogenen, bijvoorbeeld het gezamenlijke

effect van genotoxische stoffen in sigarettenrook, sterker is dan individuele blootstellingen (4). Dit zou een verklaring kunnen zijn voor de sterkere stijging in effect gezien bij blootstelling aan sigarettenrook, in vergelijking met individuele stoffen. Een alternatieve verklaring is ook voorgesteld, waarbij de hypothese was dat genetische verschillen tussen mensen het lagere effect kunnen verklaren bij individuele carcinogeen blootstellingen (4). Er werd inderdaad een effect van verschillende polymorfismen gevonden op de blootstelling-effect relaties in hoofdstuk 5, waarbij een sterker effect werd gevonden bij personen met een “at risk” genotype in respons op blootstelling aan carcinogenen. De niveaus van deze effecten waren alleen nog steeds vrij laag (in de meeste gevallen slechts tussen 1 en 1.5 maal verschillend tussen de specifieke genotypes). Daarom wordt gesuggereerd dat het effect van deze polymorfismen groter zal zijn als men kijkt naar de response na blootstelling aan een mix van verschillende (lage doses) carcinogenen tegelijk. Dit komt ook dicht bij de buurt van de alledaagse milieublootstellingen (met de uitzondering van beroepsgebonden blootstelling aan specifieke hoge-doses carcinogenen).

Een ander belangrijk verschil is, dat de blootstelling aan sigarettenrook in hoofdstuk 3 en 4 relatief hoog is, terwijl de milieublootstellingen in hoofdstuk 5 vrij laag zijn. Daarom kunnen effecten op de gezondheid bij de laatste moeilijk te detecteren zijn, alhoewel de relaties die gevonden werden allemaal sterk statistisch significant zijn. In alle relaties werd een relatief lage R-kwadraat gevonden (data niet gepresenteerd), wat een grote variatie rondom de regressielijn suggereert. De hypothese dat in ieder geval een deel van deze variatie verklaard zou kunnen worden door het genotype werd bevestigd. Bovendien werd eerder gesuggereerd dat genetische gevoeligheid van een groter belang is bij lage doses van een bepaalde blootstelling (4). Vanuit een metabool oogpunt, kan het effect van het genotype op lage dosis niveau verklaard worden door de saturatie van de relevante enzymen bij hoge doses. Dit effect werd duidelijk geobserveerd in hoofdstuk 4, waar de impact van het *CYP1B1* $L^{432}V$ polymorfisme teniet gedaan werd bij een hoge dosis.

Een beperking van het onderzoek beschreven in dit proefschrift is de selectie van polymorfismen. Het doel was om de meest belangrijke genen in de meest relevante toxicologische pathways te selecteren, gebaseerd op literatuur data. Het mag duidelijk zijn dat dit wat arbitrair is. High-throughput genotyperingsmethoden zijn nog steeds volop in ontwikkeling, wat uiteindelijk complete genoom scans mogelijk zal maken. Een belangrijk hulpmiddel hierbij is de publicatie van de haplotype map, welke het aantal SNPs dat nodig is om het hele genoom te onderzoeken drastisch reduceert (5). Aan de andere kant, leidt dit tot een grote berg data, waarvan een groot deel biologisch irrelevant zal zijn. Het wordt gesteld dat, voor toxico-genetisch onderzoek, de meest relevante genen betrokken in de belangrijkste pathways voor chemische carcinogenese zijn geselecteerd voor het onderzoek in dit proefschrift. Bovendien bevestigen de resultaten van dit onderzoek, dat enzymen betrokken in het metabolisme van xenobiotica een belangrijke rol spelen in de gevoeligheid voor het schadelijke effect van milieublootstel-

lingen (6,7), omdat ze betrokken waren in bijna elke onderzochte blootstelling-effect relatie in de studies beschreven in dit proefschrift. Desalniettemin kan worden aanbevolen voor toekomstige studies, om polymorfismen in genen betrokken bij cel-cyclus en apoptose en genen betrokken bij de ontwikkeling en regulatie van het immuun systeem mee te nemen (8). Daarnaast zou men kunnen kijken naar haplotypes in plaats van individuele SNPs, om een beter beeld te krijgen van de functionele effecten van de gecombineerde varianten.

Naast de selectie van SNPs, is ook het aantal geanalyseerde samples in de bio-monitoring studies een aandachtspunt. Alhoewel de studie populaties in de studies beschreven in dit proefschrift groot genoeg waren gelet op het aantal onderzochte variabelen, zal een groter aantal samples de power van de analyses verhogen.

Wat betreft de beleidsondersteunende relevantie van het gepresenteerde onderzoek in dit proefschrift, is het belangrijk een verschil te maken tussen blootstelling aan carcinogene- en niet- carcinogene blootstellingen. Gelet op niet-carcinogene en niet-genotoxische blootstellingen is het concept van een zogeheten “zero-tolerance risk” model cruciaal. In dit model, wordt een threshold dosis (de dosis waaronder geen toxisch effect meer wordt gedetecteerd) gebruikt om de zogenaamde “Acceptable Daily Intake” (ADI) te berekenen. Binnen deze ADI, wordt rekening gehouden met een veiligheids marge van 10 om te corrigeren voor inter-individuele verschillen in gevoeligheid (9). Echter, alle blootstellingen beschreven in dit proefschrift zijn geclassificeerd als carcinogeen voor mensen door het IARC. Deze blootstellingen worden iets strikter gereguleerd. De extreem conservatieve benadering van een non-threshold hoog naar lage dosis extrapolatie wordt gehanteerd. In dit geval kan een threshold dosis niet worden toegepast, omdat elk individueel carcinogeen molecuul kanker kan veroorzaken en iedereen op een bepaald niveau is blootgesteld. Er wordt geen correctiefactor voor gevoeligheid toegepast, vanwege de algemene opvatting dat het model conservatief genoeg is en de limieten voor (genotoxische) carcinogenen laag genoeg zijn. Echter, deze benadering is in twijfel getrokken gebaseerd op resultaten van moleculair-epidemiologische studies die een interactie suggereren tussen polymorfismen, betrokken in chemische carcinogenese, met omgevingsfactoren, en daarbij mogelijk populaties identificeren met een hoger risico op het krijgen van kanker (4). De resultaten van het onderzoek beschreven in dit proefschrift laten in alle geobserveerde blootstelling-effect relaties meer gevoelige subgroepen zien met hogere respons niveaus van biomarkers voor genotoxische schade, en dus een hoger risico op het ontwikkelen van schadelijke effecten op de gezondheid naar aanleiding van genotoxische omgevingsblootstellingen. Alhoewel enige voorzichtigheid gebaat is bij de interpretatie van de resultaten in deze studie omdat biomarkers onderzocht zijn in plaats van de incidentie van kanker, wordt geconcludeerd dat het van sociaal belang is om het debat over inter-individuele verschillen in gevoeligheid voor omgevingsgerelateerd kankerrisico te continueren.

Conclusie

Het onderzoek gepresenteerd in dit proefschrift beschrijft een snelle, accurate en grondig gevalideerde methodologie voor het genotyperen van multiple SNPs in grote populaties, welke makkelijk toe te passen is als een biomarker voor genetische gevoeligheid in moleculair-epidemiologische studies. Een sterk punt van het gepresenteerde onderzoek is de grote aantallen genotypes, in vooraf geselecteerde pathways van interesse, die kunnen worden onderzocht (variërend van 684 in hoofdstuk 4 tot 24,986 in hoofdstuk 5), in tegenstelling tot de relatief lage capaciteit van conventionele methodes zoals RFLP. Een tweede aspect is dat de impact van multiple SNPs als geheel wordt onderzocht op de relatie tussen blootstelling- en effect markers, waar conventionele methoden zich richten op het effect van een enkel polymorfisme op een specifieke biomarker. Dit bleek uit de resultaten van de ontwikkelde multivariaat statistische methodologie, die meer informatie gaf dan de conventionele univariaat methoden. Voor toekomstige studies wordt aanbevolen, om de totale effecten van multiple polymorfismen op genomgevings interacties te benadrukken.

References

1. Perera, F. P. Molecular epidemiology: insights into cancer susceptibility, risk assessment, and prevention. *J Natl Cancer Inst*, 88: 496-509, 1996.
2. Smith, C. J., Perfetti, T. A., Garg, R., and Hansch, C. IARC carcinogens reported in cigarette mainstream smoke and their calculated log P values. *Food Chem Toxicol*, 41: 807-17, 2003.
3. Hecht, S. S. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst*, 91: 1194-210, 1999.
4. Vineis, P. Molecular epidemiology: low-dose carcinogens and genetic susceptibility. *Int J Cancer*, 71: 1-3, 1997.
5. The International HapMap Project. *Nature*, 426: 789-96, 2003.
6. Autrup, H. Genetic polymorphisms in human xenobiotica metabolizing enzymes as susceptibility factors in toxic response. *Mutat Res*, 464: 65-76, 2000.
7. Bartsch, H., and Hietanen, E. The role of individual susceptibility in cancer burden related to environmental exposure. *Environ Health Perspect*, 104 Suppl 3: 569-77, 1996.
8. Brennan, P. Gene-environment interaction and aetiology of cancer: what does it mean and how can we measure it? *Carcinogenesis*, 23: 381-7, 2002.
9. Knudsen, L. E., Loft, S. H., and Autrup, H. Risk assessment: the importance of genetic polymorphisms in man. *Mutat Res*, 482: 83-8, 2001.

Dankwoord

Chapter 7

Juli 2007, Rowwen Hèze is tourartiest geworden bij ‘radio tour’ en “Vechte, Valle en Opstaan” klinkt de kamer in. Toeval bestaat niet. Muziek gelukkig wel, en blijkt wederom een makkelijk middel om iets te zeggen. Want terugkijkend op 4,5 jaar AiO-schap lijken de woorden van Jack Poels precies op de plaats te vallen, juist op het moment dat ik eindelijk mijn dankwoord mag schrijven, voor de mensen die mij het ‘vechten’ en het ‘opstaan’ een stuk makkelijker hebben gemaakt.

Allereerst mijn begeleider en co-promotor, Dr. Joost van Delft. Beste Joost. Ik was natuurlijk een klein beetje een vreemde eend in de bijt, als enige van je AIOs die eigenlijk niets met micro-array’s te maken had. Dat neemt niet weg dat je me altijd heel enthousiast begeleidt hebt, met veel interesse in alles waar ik mee bezig was (ook de “minder” wetenschappelijke zaken!). Ook het feit dat ik ten alle tijden bij je terecht kon met vragen, frustraties (‘t “AIO-dipje”) of gewoon even “bijpraten” waardeert ik enorm. Daarnaast natuurlijk mijn promotor, Prof. Jos Kleinjans. Beste Jos, ik vind het echt ongelooflijk dat je altijd precies op de juiste momenten de juiste woorden of adviezen hebt, die me telkens weer een stap verder brachten (zodat ik die 2-stappen die je voor je gevoel *terug* maakt, weer makkelijk inhaalde), ook al week ik een enkele keer iets af van de voorgestelde route (overleving van de meest eigenwijze?). Jos en Joost, heel erg bedankt voor de goede samenwerking!

Alle collega’s van de vakgroep gezondheidsrisico analyse en toxicologie (GRAT), bedankt voor alle hulp en de leuke tijd! Een paar wil ik er graag speciaal noemen. Dit is in de eerste plaats Ralph. “Willi”, bedankt voor de grote hoeveelheden werk die je verzet hebt, de leuke tijd op (en buiten!) het lab en je eeuwige geduld als ik weer eens aan kwam zetten om vanalles van je te vragen. Zonder jou was dit proefschrift er niet (zo snel) geweest! Daarnaast Roger (“Dur Rodger”). Heel erg bedankt voor de vele, leuke discussies, en met name voor al je hulp bij het “statistisch geplaat” (door jou ben ik statistiek zelfs leuk gaan vinden!) en het schrijven. Ik heb veel van je geleerd! (Simpelvelds rules!). Ook Ad, “the dude”, bedankt voor alle hulp! (En natuurlijk voor het samen organiseren van een van de meest memorabele GRAT-uitstapjes ooit). Bukske, d’n Dulle, d’n Duffe, d’n Vette (Sodemieter op nao Venlo man), Luus, Willi bedankt voor de super gezellige tijd bij GRAT; de analisten (P^*K) brengen leven in de brouwerij! (wat ik in die jaren een aantal keer *letterlijk* heb mee mogen maken, met Val Dieu dit jaar als hoogtepunt). Natuurlijk mijn mede AIO’s, bedankt voor de gezelligheid en het delen van alle AIO hoogte- en dieptepunten! Danitsja, heel erg bedankt voor alle layout-tips! Tenslotte ‘n “ex-collega”, onze Ludwig-drummer, Harald, bedankt voor de gezellige jaren bij GRAT! (iek find mooi, puuh).

Gudrun en Elly van het Vlaams Instituut voor Technologisch Onderzoek (VI-TO), heel erg bedankt voor alle hulp!

Leden van de beoordelingscommissie, bedankt voor jullie tijd en de snelle beoordeling van mijn proefschrift.

Ook wil ik graag Antje bedanken voor de hulp bij het kaftontwerp, evenals de heer en mevrouw Van den Broek en Duco Straatman voor jullie steun!

Dankwoord

(Schoon-)familie en vrienden, bedankt voor de nodige ontspanning en interesse: dit heb ik dus 4 jaar uitgevoerd op het lab! De mafste vriendengroep van Venlo en omstreken, de Brandbeerkes, wil ik uiteraard even met name noemen. “Ouk al zuus-se mekaar jaore neet, de bis in 5 min weer beejgelölt, as of-se mekaar gister nog zoogs”. Funske, te gek dat je mijn paranimf wilt zijn! Highke (Storm in miène kop), ook al zien we mekaar niet veel meer, bedankt voor ’n te gekke vriendschap! (“Those were days of roses, poetry en proza”).

En net zoals een band z’n beste nummers tot het laatst bewaard, rest mij nog 4 personen te bedanken. Allereerst pap & mam en Immy. Heel erg bedankt voor de mogelijkheden die jullie me gegeven hebben, het vertrouwen en meeleven, en het feit dat jullie altijd voor me klaarstaan! P.S. Mam, de “samenvatting veur minse wie weej” zit achterin!

Leeve KiEkske, bedankt voor je onvoorwaardelijke steun, je interesse in alles wat ik doe en de onuitputtelijke bron van positieve energie die je hebt en op de juiste momenten op mij over weet te brengen! De woorden van Jack Poels hadden door jou geschreven kunnen zijn...

“neet zeure, begin, ut haet gènne zin, de wachs tevergaefs op ein wonder!”

...Samen kunnen we de wereld aan!

Bedankt!
Henk

Curriculum vitae

I was born on December 15th, 1978, in Leunen as Hans Bernardus Ketelslegers, and raised in Venlo. I have attended pre-university education from 1991-1997 at the “Collegium Marianum” in Venlo.

I started the study program “Biomedical Engineering” at the technical university of Eindhoven, and after 1,5 years I left Eindhoven to study “Biological Health Sciences” at Maastricht University. During this period, I did two internships, one at the Leiden University Medial Centre (LUMC) at the department of human genetics, on the subject “Reversed Transfection using cDNA Microarrays”, and one at the department of health risk analysis and toxicology (GRAT) at Maastricht University, on “Genotyping Single Nucleotide Polymorphisms in Biotransformation enzymes using Arrayed Primer EXtension”.

In December 2002 I obtained my MSc in biological health sciences, and in January 2003 I started the PhD research as described in this thesis.

I am currently working as a post-doc at the department of health risk analysis and toxicology at Maastricht University.

Op 15 december 1978 ben ik als Hans Bernardus Ketelslegers geboren te Leunen, en daarna opgegroeid in Venlo. In 1997 heb ik mijn VWO diploma gehaald, aan het “Collegium Marianum” in deze stad.

Na 1,5 jaar “Bio-Medische Technologie” aan de Technische Universiteit van Eindhoven, heb ik van 1998-2002 Biologische Gezondheidskunde gestudeerd aan de Universiteit Maastricht. Tijdens deze studie heb ik 2 stages gevolgd, de eerst in het Leids Universitair Medisch Centrum (LUMC) bij de vakgroep humane genetica, met als onderwerp “Reversed Transfection using cDNA Microarrays”, de tweede aan de vakgroep GezondheidsRisico Analyse en Toxicologie (GRAT) van de Universiteit Maastricht, over “Genotyping Single Nucleotide Polymorphisms in Biotransformation enzymes using Arrayed Primer EXtension”.

In december 2002 heb ik mijn Doctorandus diploma gehaald, en ben ik in januari 2003 bij GRAT begonnen als promovendus aan het onderzoek dat beschreven is in dit proefschrift.

Vanaf april 2007 ben ik werkzaam als post-doc bij de vakgroep GRAT aan de Universiteit Maastricht.

List of publications

Knaapen, A. M., Ketelslegers, H. B., Gottschalk, R. W., Janssen, R. G., Paulussen, A. D., Smeets, H. J., Godschalk, R. W., Van Schooten, F. J., Kleinjans, J. C., and Van Delft, J. H. Simultaneous genotyping of nine polymorphisms in xenobiotic-metabolizing enzymes by multiplex PCR amplification and single base extension. Clin Chem, 50: 1664-8, 2004.

Ketelslegers, H. B., Gottschalk, R. W., Godschalk, R. W., Knaapen, A. M., van Schooten, F. J., Vlietinck, R. F., Kleinjans, J. C., and van Delft, J. H. Interindividual variations in DNA adduct levels assessed by analysis of multiple genetic polymorphisms in smokers. Cancer Epidemiol Biomarkers Prev, 15: 624-9, 2006.

Hans B Ketelslegers, Roger W.L. Godschalk, Bart JM Eskens, Jan W Dallinga, Ralph W.H. Gottschalk, Frederik J. van Schooten, Jos C.S. Kleinjans, and Joost H.M. van Delft. Potential role of CYP1B1 in the metabolic activation of 4-AminoBiPhenyl in humans. (*In Preperation*)

Hans B Ketelslegers, Ralph W.H. Gottschalk, Gudrun. Koppen, Greet Schoeters, Willy Baeyens, Nicolas van Larebeke, Joost H.M. van Delft, and Jos C.S. Kleinjans. Multiplex genotyping as a biomarker for susceptibility to identify subgroups at increased risk in carcinogenic exposure-effect relationships in the FLEHS biomonitoring study. (*In Preperation*)

Hans B Ketelslegers, Roger W.L. Godschalk, Ralph W.H. Gottschalk, Ad M. Knaapen, Gudrun. Koppen, Greet Schoeters, Willy Baeyens, Nicolas van Larebeke, Joep P.M. Geraedts, Joost H.M. van Delft, Jos C.S. Kleinjans. The prevalence of an at-risk genotype for environmental carcinogenesis decreases with age; survival of the fittest? (*In Preperation*)

Wilms LC, Boots AW, de Boer VCJ, Maas LM, Pachen DMFA, Gottschalk RWH, Ketelslegers HB, Godschalk RWL, Haenen GRMM, van Schooten FJ, Kleinjans JCS. Impact of multiple genetic polymorphisms on effects of a four-week blueberry juice intervention on ex vivo induced lymphocyte DNA damage in human volunteers. Carcinogenesis. 2007 June 29; [Epub ahead of print]

Moons LMG, Kusters JG, van Delft JHM, Kuipers EJ, Gottschalk RWL, Geldof H, Bode WA, Stoof J, van Vliet AHM, Ketelslegers HB, Kleinjans JCS, Siersema PD. A pro-inflammatory genotype predisposes to the development of Barrett's esophagus. (*Accepted Carcinogenesis. August 2007*)

Abstracts

Ketelslegers HB, Gottschalk RWH, Godschalk RWL, Knaapen AM, Van Schooten FJ, Kleinjans JCS, Van Delft JHM. Prevalence of an at-risk genotype for environmental carcinogenesis decreases with age: Survival of the fittest? *The Annual meeting of the American Association for Cancer Research*, LA convention centre, Los Angeles, USA, April 2007.

Ketelslegers HB, Gottschalk RWH, Godschalk RWL, Knaapen AM, Van Schooten FJ, Kleinjans JCS, Van Delft JHM. Predicting Multiple Gene-Environment Interactions. *Gene-Environment Interactions Course*, Karolinska Institute, Stockholm, Sweden, June 2006.

Ketelslegers HB, Gottschalk RWH, Godschalk RWL, Knaapen AM, Van Schooten FJ, Vlietinck RFMH, Kleinjans JCS, Van Delft JHM. Inter-Individual variations in DNA adduct levels assessed by analysis of multiple polymorphisms in smokers. *SNP-NL Platform (Dutch SNP consortium) meeting*, Utrecht, The Netherlands, November 2005.

Ketelslegers HB, Gottschalk RWH, Godschalk RWL, Knaapen AM, Van Schooten FJ, Vlietinck RFMH, Kleinjans JCS, Van Delft JHM. Genotyping Single Nucleotide Polymorphisms as biomarker for genetic susceptibility. *Annual PhD Student Meeting of the Dutch Toxicology Society (NVT)*, Oss, The Netherlands, June 2005.

Ketelslegers HB, Gottschalk RWH, Godschalk RWL, Knaapen AM, Van Schooten FJ, Kleinjans JCS, Van Delft JHM. Predicting Multiple Gene-Environment Interactions Using Discriminant Function Analysis. *34th annual meeting of the European Environmental Mutagen Society (EEMS)*, Maastricht, The Netherlands, September 2004.

Ketelslegers HB, Gottschalk RWH, Godschalk RWL, Knaapen AM, Van Schooten FJ, Kleinjans JCS, Van Delft JHM. A Multiplex Correlation Study Between Genes Involved In Chemical Carcinogenesis And DNA Adducts. *1st International Conference "Molecular research in Environmental Medicine"*, Düsseldorf, Germany, March 2004.

Ketelslegers HB, Kleinjans JCS, Van Delft JHM. Toxicogenomic Applications in Biomonitoring studies: genotyping Single Nucleotide Polymorphisms. *Symposium of the "Flemish Environment and Health Survey" consortium*, University of Gent, Belgium, November 2003.

Ketelslegers HB, Van Leeuwen DM, Kleinjans JCS, Van Delft JHM. Toxicogenomic Applications in Biomonitoring studies. *10th anniversary Symposium on Environment and Health in Flanders*, Louvain, Belgium, October 2003.

Ketelslegers HB, Kleinjans JCS, Van Delft JHM. Toxicogenomic Applications in Biomonitoring studies. *Meeting of the "Flemish Environment and Health Survey" consortium*, Flemish Institute of Technological Research, Mol, Belgium, April 2003.

Ketelslegers HB, Van Delft JHM. Genotyping Single Nucleotide Polymorphisms in Biotransformation enzymes using Arrayed Primer Extension. *Course: "Apex on DNA microarrays: applications in SNP analysis, mutation detection and DNA resequencing"*, Tartu, Estonia, August 2002.

Een genetisch paspoort

“Samenvatting voor iedereen”

*Geschreven voor de cursus “ Schrijven over wetenschap”
21 april 2005*

Rokers sterven jonger. Hoewel, iedereen kent wel een roker die het levende bewijs is van het verhaal over iemand die zijn leven lang een pakje per dag heeft gerookt en nog steeds kerngezond is. Het is bekend dat mensen verschillend kunnen reageren op dezelfde blootstelling aan schadelijke stoffen. Een groot deel hiervan kan verklaard worden door genetische verschillen. We krijgen bij de geboorte een persoonlijk genetisch paspoort mee, bestaande uit ‘varianten van genen’: genetische polymorfismen.

De vraag is hoe bijzonder dit nu eigenlijk is, want bij het lezen van de genetische informatie van verschillende individuen lezen we voor 99.9% hetzelfde. Dan lijkt de conclusie dat niemand echt bijzonder is gauw getrokken. Maar als we die luttele 0.1% die het verschil maakt tussen mensen wat beter bestuderen, komen we er al gauw achter dat dit wat kortzichtig is.

Snips: Genetische polymorfismen

Dit minieme verschil in de genetische informatie tussen individuen wordt bepaald door genetische polymorfismen. De genetische informatie, het DNA, is opgebouwd uit 4 basen: Adenine (A), Guanine (G), Thymine (T) en Cytosine (C). Een lange reeks van deze basen achterelkaar is een *gen*, een code die als blauwdruk staat voor het maken van *eiwitten*. Er bestaan aangeboren afwijkingen in een gen, waarbij het DNA op één base verschilt. Deze ‘één base veranderingen’ heten Single Nucleotide Polymorphisms (SNPs, uitgesproken als ‘Snips’).

De invloed van één letter

De eiwitten waar een gen voor codeert kunnen functioneren als *enzymen*, kleine ‘katalysatoren’ die helpen bij het goede verloop van allerlei biologische reacties in het menselijk lichaam. Voorbeeld van zo’n biologisch proces is de opname en het onschadelijk maken van giftige stoffen van buitenaf. Is een gen op één base afwijkend van de ‘normale’ code, dan kan het zijn dat het enzym waar dit gen voor codeert een veranderde, of zelfs helemaal geen functie meer heeft. In het bovengenoemde biologisch proces zou dit betekenen dat toxische stoffen (langer) in het lichaam blijven en zo hun schadelijke werking kunnen uitoefenen op bijvoorbeeld het DNA. Zo kunnen veranderingen van één enkele letter in de genetische code, SNPs, een cruciale rol spelen in het ontstaan van ziekten, zoals kanker.

Genotypering

Het in kaart brengen van deze SNPs, heet genotypering. Vroeger was het mogelijk om met eenvoudige technieken één of meerdere variant(en) in een gen op te sporen. Maar de laatste jaren is de kennis op dit gebied sterk toegenomen, in het bijzonder door het afronden van het humaan genoom project. Dit werd in 1989 gestart onder de naam

Genetisch paspoort

‘Human Genome Organisation’ (HUGO). Wetenschappers over de hele wereld werkten hieraan mee met als doel de ontrafeling van het menselijk genoom: het totaal aan genetische informatie in een individu. Vijf jaar eerder dan verwacht, wordt in 2001 de eerste ruwe versie van de blauwdruk van het leven bekend: 3,5 biljoen letters, de basen, waarvan in 98% géén genen verborgen liggen. Één op de 1200 tot 1500 basen verschillen van elkaar, dit komt neer op zo’n 3 miljoen verschillen tussen het ene genoom en het andere. Door deze enorme groei in kennis, is het noodzakelijk veel meer varianten in meerdere genen tegelijkertijd te onderzoeken. Hiervoor zijn genotyperings-technieken nodig met een grotere capaciteit dan de klassieke methoden. Het ontwikkelen en toepassen hiervan is de doelstelling van het onderzoek dat werd uitgevoerd op de capaciteitsgroep ‘GezondheidsRisico Analyse en Toxicologie’ van de Universiteit Maastricht, en beschreven in dit proefschrift.

Een genetisch paspoort maakt iedereen bijzonder

Door het toepassen van technieken die onze genetische informatie kunnen lezen, worden steeds meer functies van (varianten van) genen bekend en krijgen we steeds meer inzicht in het ontstaan van ziektes zoals kanker. Ondanks, nee, eigenlijk juist doordat we genetisch maar heel weinig verschillend zijn, reageert iedereen anders op schadelijke stoffen van buitenaf. Daarom maakt het persoonlijke genotype, het genetisch paspoort verkregen bij de geboorte, ieder mens bijzonder!