

**COMMITTEE ON CARCINOGENICITY OF CHEMICALS IN FOOD,  
CONSUMER PRODUCTS AND THE ENVIRONMENT.****A REVIEW OF POTENTIAL TARGET GENES FOR SUSCEPTIBILITY TO  
CARCINOGENESIS****Introduction**

1. A genetic component within the complex, multistage process of carcinogenesis has been acknowledged for many years and recent advances in biotechnology have facilitated the identification and investigation of numerous genes that are involved. Susceptibility to carcinogenesis should therefore be subject to heritable variation, and this is clearly demonstrated by the strong, familial nature of many rare, dominantly-inherited human cancer syndromes (Fearon, 1997). In such cases, genetic analyses have established the existence of inherited defects in individual, “high-penetrance” cancer genes as the underlying basis for a strong predisposition towards the development of a specific cancer type(s) within individuals carrying the affected genotype. This, in turn, has led to the identification and characterisation of genes that play key roles in oncogenesis. Examples of such rare, high-penetrance familial cancer syndromes, and their associated cancer genes, are shown in Table 1. In such cases the inherited single gene-type is the fundamental basis for the development of the associated cancer, typically at an early age in comparison to that for average sporadic cancer incidence. Environmental factors are presumed to play a minimal role in the occurrence of these cancers (see accompanying paper – CC/01/3).

2. High-penetrance hereditary cancer syndromes, although devastating within affected families, are nonetheless responsible for only a small percentage of the overall cancer burden within society. Rather, the vast majority of cancers are presumed to result from an interaction of genetics and the environment (gene-environment interactions)– i.e. from the interaction of susceptibility alleles at a series of loci and moderate or low level exposure to carcinogenic agents (Perera, 1997). Ottman (1996) has described 5 models, based on different biological scenarios, for the possible relationships that can exist between a susceptibility genotype and environmental exposure in terms of their effects on cancer outcome (Figure 1). Most studies to date have considered gene-environment interactions that conform to model B. In this situation, a single susceptibility locus variant is postulated to confer a modest increase in risk of tumour incidence (for example, less than two-fold), in combination with the associated exposure – i.e. to modulate the effect of the exposure. Thus, given the small increase in risk, characterisation with the susceptible genotype is unlikely to be of particular significance in the ascertainment of risk for an individual subject (in the absence of consideration of the entire “genetic background” of the subject). There may, however, be a measurable effect on overall cancer burden within a population if the variant allele is present at a sufficiently high population frequency (see accompanying paper – CC/01/5).

3. Variant alleles that are associated with an altered risk for disease susceptibility have been termed “low penetrance susceptibility alleles”. The majority of such

sequences within the human genome probably exist as simple mutations, the total number of which may be very large (see accompanying report – CC/01/3). Variation may also exist as genomic deletions/insertions, variable repeat polymorphisms, etc.. Altered susceptibility would be assumed to occur as a result of the presence of a gene<sup>1</sup> variant (i.e. within a gene or locus control region) conferring functional variation in a biological pathway (for example, by specifying a product with altered function, expression or stability). Functional alterations in biological pathways may be hypothesised to affect susceptibility to carcinogenesis in various ways (Figure 2), for example;

- 1] By encoding a phenotype with increased likelihood of exposure to carcinogenic insult (e.g. increased likelihood of exposure through behavioural differences or distribution of the agent to a specific body compartment; increased capacity for metabolic activation of a precarcinogen or reduced capacity to detoxify or eliminate an active carcinogen).
- 2] Encoding a phenotype with reduced capacity to neutralise the effects of the insult at the level of the individual cell, tissue or organism (e.g. decreased effectiveness of DNA repair, defective immune function).
- 3] Encoding a phenotype with increased potential for cell proliferation/survival (e.g. alterations in proliferative signalling pathways and/or those that affect control of the cell cycle, apoptosis or senescence).

4. Kinsler & Vogelstein (1997) have recently proposed the concept of high penetrance cancer susceptibility genes as “caretakers” or “gatekeepers”, based on their roles in the maintenance of genomic integrity (e.g. DNA repair), and cellular proliferation (e.g. cell cycle control, DNA replication), respectively (although there will obviously be overlap between these two categories). Dysfunctional caretaker genes increase the probability of mutations in gatekeeper genes, which are necessary to initiate the molecular pathogenesis of cancer. Shields & Harris (2000) propose that this model can also be applied to low-penetrance genes (Figure 3).

5. The focus of this paper is to review current and potential low penetrance gene targets for molecular epidemiology studies in carcinogenesis. It is clear that this may include a large number (probably many thousands) of potential variant loci, most of which have not yet been identified. In this respect, an ideal study would involve the sequencing of the entire genome of a large number of individuals. An epidemiologic study associating genome-wide scan with causes of mortality would demonstrate whether these low penetrance loci exist and can be associated with risk of cancer at the population level (see CC/01/3). However, current technologies do not allow this type of systematic study. Rather, the approach that has been most widely used is to assess specific candidate loci. An in-depth discussion of all potential candidates is clearly beyond the scope of this review. Rather the aim is to give an overview of the major relevant pathways, with representative examples where appropriate. The vast majority of research efforts to date have focussed on the evaluation of candidate genes encoding enzymes involved in the metabolism (activation, detoxification and elimination) of xenobiotic chemicals (and endogenous compounds). This field has been reviewed in detail by many authors and only the major findings to date will be summarised here. There is currently a growing interest in the likelihood that variation

---

<sup>1</sup> Within this report the term “gene” may be used in a generic sense to refer to any (coding or non-coding) genomic sequence.

in other caretaker gene functions (for example, DNA repair, immune surveillance and response) may have an impact on carcinogenic risk, and these aspects will be considered in more detail. Studies are also underway to assess the role of genetic factors in association with diet and nutrition in carcinogenesis (and cancer prevention). Finally, it is reasonable to predict that altered susceptibility will be associated with genetic variation affecting pathways controlling DNA replication, cellular proliferation and survival. This area is likely to be a major focus for future investigations.

### **Metabolic Polymorphisms**

6. Many chemical carcinogens are biotransformed within the body, either to compounds that are more toxic than the original compound, or to less toxic metabolites that can be excreted. The activities of these pathways should therefore be important in determining the level of DNA damage elicited by a carcinogen. Historically, it is well established that genetic differences occur in the expression of xenobiotic metabolising enzymes, originally identified due to individual differences in sensitivity to therapeutically used drugs – the field of “pharmacogenetics”. Recent molecular studies have also shown numerous mutant or variant genes encoding these enzymes (see IARC, 1999; Autrup, 2000; <http://www.imm.ki.se/CYPalleles/>; <http://www.louisville.edu/medschool/pharmacology/NAT.html>) (**Table 2**), and it is hypothesised that such variants underlie the basis for some of the interindividual variation in bioactivation/detoxication of carcinogenic compounds – i.e. influence individual cancer susceptibility (see accompanying paper – CC/01/3 – for a discussion of the *a priori* selection of candidate genes based on known or likely function which are causally related to cancer). Many studies have been carried out to assess the potential association of these variants with risks of environmentally induced cancers. The majority of studies have focussed on three major groups; the P450 family of enzymes (CYP), glutathione *S*-transferases (GST) and *N*-acetyltransferases (NAT). Findings have been extensively reviewed by many authors (e.g. Gonzalez, 1997; IARC, 1999; Taningher et al., 1999) and will not be described in detail here. The International Agency for Cancer Research (IARC) has published a detailed analysis of studies to 1999. What follows (paragraphs 8-19) is a brief outline of the main conclusions of the IARC report. These are summarised in **Table 3**.

### **CYP**

7. The CYP-dependent monooxygenases are a supergene family of enzymes that are responsible for the phase I metabolism of many xenobiotic compounds. These enzymes have differing but overlapping substrate specificities. In most cases, CYP metabolism is associated with detoxication, however the enzymes can also be responsible for the metabolic activation of certain chemicals to genotoxic intermediates. Many CYPs are constitutively expressed within the liver, but tissue specificity and distribution vary widely amongst the different forms. Although some CYP enzymes are constitutively expressed at relatively low levels, they may be induced in a tissue-specific manner in response to hormonal and/or environmental stimuli. Polymorphisms have been identified for a number of CYPs.

8. **CYP1A1** –This is one of the principal enzymes capable of activating polycyclic aromatic hydrocarbons (PAHs), several of which are considered to be

important carcinogens. It is present in several extrahepatic tissues, particularly after induction, but appears to be absent from the liver. *CYP1A1* gene expression is regulated transcriptionally by the aromatic hydrocarbon (Ah) receptor. A number of allelic variants (polymorphisms) have been described.

*MspI polymorphism* - Data from studies in Asian (mainly Japanese) populations were not very consistent, but did suggest the existence of an association of the polymorphism with increased risk of lung cancer, specifically squamous and small-cell carcinomas. The risk was modified by smoking (risk relative to the non-susceptibility genotype was almost double amongst light smokers, but was close to 1.00 amongst heavy smokers). Data did not support an association with lung cancer in non-Asian populations. A possible association with breast and endometrial cancers (particularly in African-Americans) was noted.

*Exon 7 polymorphism* - There was no consistent evidence for an association with lung cancer, although it was noted that two out of three studies conducted amongst Asians did find a significant association. This polymorphism was associated with a (non-significant) relative risk of  $\approx 1.5$  for breast cancer. Possible associations with endometrial and basal cell skin cancers were also suggested.

*Aryl hydrocarbon hydroxylase (AHH) activity* - Data suggested a positive association of the high inducibility phenotype (genetic basis unknown) with lung cancer. One study showed a strong association of the high inducibility phenotype with larynx cancer, but another study showed, conversely, association of the low-inducibility phenotype with larynx cancer.

9. CYP1A2 - This is a liver-specific enzyme with broad substrate specificity including structurally diverse compounds such as nitrosamines, heterocyclic arylamines (some of which are important food mutagens) and aflatoxin B<sub>1</sub>. *CYP1A2* gene expression is also regulated by the Ah receptor. No genetic polymorphism has been identified for *CYP1A2*, but substantial interindividual variability in the *in vivo* activity of the enzyme (i.e. phenotypic variability) has been shown, and several allelic variants have been reported.

Some studies suggested a possible association of the extensive metaboliser phenotype with colorectal polyps, colorectal cancer and bladder cancer.

10. CYP2D6 - This enzyme metabolises several important clinically used drugs, but there is little evidence for a role in carcinogen activation. Large interindividual variability in activity has been observed *in vivo* and the molecular bases for these phenotypic variations have now been established. Studies showed an increased risk of lung cancer associated with the extensive metaboliser (EM) phenotype, although no such associations were found in genotyping studies. Smoking did not modify risk associated with the EM phenotype. Limited data suggested an association of the homozygous EM genotype with liver cancer and cervical intraepithelial neoplasia (CIN), whilst the poor metaboliser phenotype/genotype seemed to be associated with increased risk for brain cancer.

11. CYP2E1 - This CYP is expressed at high levels in the liver and lower levels in several extrahepatic tissues. It is involved in the metabolism of many low-molecular-weight toxins and carcinogens, such as nitrosamines, ethanol, benzene, carbon tetrachloride, vinyl chloride and acrylonitriles. The enzyme is induced, for example

during fasting or diabetes. Several allelic variants (polymorphisms) have been described.

*5'-flanking region polymorphism* – Evidence suggested an association of the “wild type” genotype with lung cancer, in particular squamous cell and small-cell carcinomas, although the results of a meta-analysis were not statistically significant. Two studies also suggested an association of the “wild type” genotype with bladder cancer. Limited data suggested associations of the “wild type” and mutant genotypes, respectively, with liver and nasopharyngeal cancers.

## **GST**

12. The glutathione S-transferases (GSTs) are a supergene family of enzymes that exhibit differing, but overlapping, substrate specificities. The enzymes are involved in the phase II detoxication of toxins including certain carcinogenic compounds and also play a role in neutralising damaging electrophiles and products of oxidative stress.

13. GSTM1 – This enzyme is found mainly in the liver, but is unexpressed in only  $\approx$  50% of Caucasians due to homozygous deletion, the so-called null genotype. Studies suggested an association of the null genotype with lung cancer, with higher risk of squamous and small-cell carcinomas. Results were stronger and more consistent for Asians, showing an interaction with heavy smoking. There was an association with bladder cancer in all ethnic groups (overall relative risk  $\approx$  1.5), although results were more heterogeneous for Caucasians, unless analysis was specifically focussed upon transitional cell carcinomas and smokers. Data from a few studies suggested an association of the null genotype with increased risk of breast cancer in younger women and of cancers of the distal colon. Some studies showed a significant association between *GSTM1* null subjects and gastric cancer (particularly in studies dealing with adenocarcinoma) in Asians. Data suggested an increased risk of the null genotype with laryngeal cancer (significant for squamous cell carcinoma). There were also suggested associations of the null genotype with pituitary adenoma, endometrial cancer and acute lymphoblastic leukaemia.

14. GSTT1 – This enzyme participates in the metabolism of monohalomethanes and metabolites of butadiene, mainly in the liver. A deletion polymorphism also exists for *GSTT1*. The prevalence of null individuals varies among ethnic populations, and is reportedly around 10-20% in Caucasians. A few studies suggested an association of *GSTT1* null carriers with larynx cancer. Meta-analysis showed a significant association of null carriers with colorectal cancer, although it was noted that the result was strongly influenced by one study. Data also suggested associations of the *GSTT1* null genotype with basocellular skin cancer of the trunk, myelodysplastic syndrome, and cancers of the brain and cervix.

## **NAT**

15. Polymorphisms in the genes encoding the acetyltransferases, NAT1 and NAT2, are associated with interindividual variation in the biotransformation (both activation and deactivation) of many primary aromatic amines or hydrazines, and give rise to the slow and fast acetylator phenotypes (NAT2).

16. NAT1 – Data supported a possible association of the rapid genotype with increased risk of colorectal cancer and adenoma.

17. NAT2 – Caucasians carrying the *NAT2* slow genotype showed increased risk of bladder cancer (relative risk  $\approx$  1.4). Risks were higher amongst smokers and subjects exposed to occupational carcinogens. Phenotype, but not genotype, based studies showed an increased risk of breast cancer in rapid carriers. Phenotype based studies also suggested an association of colorectal cancer with the slow phenotype, but this was not confirmed in genotype studies. Other studies suggested possible associations of the slow phenotype with colorectal cancer, liver cancer and, amongst subjects heavily exposed to asbestos, mesothelioma.

### **Interactions between metabolic polymorphisms**

18. Some studies have attempted to evaluate the combined effect of two or more metabolic polymorphisms on cancer risk. The available evidence suggested;

- 1- an increased risk of lung cancer for carriers of the *GSTM1* null genotype and either the *CYP1A1 Msp I* or the *CYP1A1 exon7* high-risk alleles
- 2- an increased risk of bladder cancer and a decreased risk of colorectal cancer among carriers of the *GSTM1* null and the *NAT2* slow genotypes
- 3- a decreased risk of bladder cancer among carriers of *NAT1* rapid and *NAT2* slow genotypes
- 4- an increased risk of liver cancer among *NAT2* slow acetylators and *CYP2D6* extensive metabolisers.

These interactions are, however, largely unconfirmed to date and of uncertain biological significance.

19. In summary, many studies have been carried out to assess the potential association of metabolic gene variants with risks of environmentally induced cancers. The majority of investigations have focussed on three main groups; the P450 family of enzymes (CYP), glutathione *S*-transferases (GST) and *N*-acetyltransferases (NAT). Although some studies have suggested the existence of genotypes associated with increased susceptibility to chemical carcinogenesis, overall the data for most metabolic polymorphisms are not convincing. There are numerous other xenobiotic metabolising enzymes that are candidates for a role in increased susceptibility to environmental carcinogens. In addition, the relatively recent recognition of the importance of transporters such as P-glycoprotein in determining systemic and specific tissue exposure to a range of xenobiotic compounds also makes them potential susceptibility targets. Indeed, functional polymorphisms of P-glycoprotein have been reported, although their relevance to cancer susceptibility has yet to be established.

### **Variation in DNA repair pathways**

#### *DNA repair and human cancer syndromes*

20. DNA damage may occur as a consequence of normal cellular processes (for example, replication or attack from the products of oxidative metabolism) or as a result of extrinsic exposures (such as xenobiotic chemicals or radiation). DNA repair comprises a complex network of interacting mechanisms, which function to prevent

the detrimental consequences of such damage. By inference from this role of DNA repair as the guardian of cellular integrity, inherited variations affecting the function of one or more of the multitude of components involved in these pathways provide a likely target for altered susceptibility to cancer. Indeed, that DNA repair plays a significant role in cancer biology is demonstrated by the existence of high-penetrance cancer syndromes associated with mutations in genes encoding DNA repair enzymes (Guilford, 2000). Xeroderma pigmentosum (XP), for example, is a rare, autosomal recessive disorder attributed to mutations in genes (*XPA* – *XPG*) encoding enzymes of the DNA nucleotide excision repair pathway (see below), in which affected individuals have a > 1000-fold increased risk of skin cancer in association with exposure to sunlight. Hereditary nonpolyposis colon cancer (HNPCC) is another familial cancer syndrome associated with mutations in DNA repair genes, in this case attributed to inactivating mutations of (mut S and mut L-related) genes encoding proteins involved in DNA mismatch repair (see below).

### *DNA repair pathways*

21. DNA repair pathways are classified with respect to the type of DNA lesion to which they are specific (Friedberg et al., 1995), although it is apparent that there is actually substantial overlap in substrate specificity, and that some proteins function in more than one pathway. The four major classes of DNA repair are represented in Figure 4 and summarised, briefly, below;

- 1] Mismatches or structural abnormalities at replication forks are repaired by the mismatch repair (MMR) pathway, which is currently thought to include six dedicated mismatch repair proteins (hMSH2/3/6, hMLH1/3, hPMS2) as well as members of the replication machinery (Jiricny & Nystrom-Lahti, 2000).
- 2] Double-strand breaks are repaired by homologous-recombination repair or in an end-joining reaction. Analysis of mammalian double strand break repair-deficient cell lines, in combination with studies in yeast, have identified 15-20 genes in this group, where loss of function is associated with impaired ability to repair DNA damage induced by ionising radiation (Jeggo et al., 1995; Lehmann, 1996).
- 3] The majority of small base modifications are removed by base excision repair (BER), the predominant pathway for removing changes arising from spontaneous hydrolysis, alkylation and reactive oxygen attack. A group of 15-20 genes has been identified in this pathway (Memisoglu & Samson, 2000).
- 4] Bulky, helix-distorting DNA lesions are repaired by the nucleotide excision repair (NER) pathway (de Boer & Hoeijmakers, 2000). This pathway involves > 20 proteins, many of which form a multiprotein complex. The majority of physical and chemical carcinogens (except ionising radiation and most alkylating agents) produce bulky lesions which are usually repaired by NER (Benhamou & Sarasin, 2000), and thus this pathway has been a particular focus of many studies regarding DNA damage due to extrinsic carcinogens.

### *Phenotypic variation in DNA repair capacity*

22. Interindividual variation in the activities of specific DNA repair pathways has been demonstrated using *in vitro* assays for functional DNA repair capacity, for example lymphocyte-based assays which measure the efficiency of repair of a plasmid vector carrying a marker gene, or which assess mutagen sensitivity (chromatid breaks

or chromosomal aberrations, etc., in response to a genotoxic agent). The nature of these studies is such that, rather than revealing information regarding specific variant alleles involved in DNA repair mechanisms, they provide information about the capacity of a pathway, or interacting pathways, as a whole – i.e. on the overall phenotype conferred by a set of alleles, whose products may act independently, or as members of multi-protein complexes within one or more specific repair pathways.

23. For example, a number of studies investigating NER or double-strand break repair in association with breast cancer have shown that subjects with reduced repair capacity (65-80% in comparison to the population mean) are over-represented in some cancer cohorts (in general these studies have assessed familial clusters or first-degree relatives of cancer patients), as compared with control cohorts (with reported odds ratios for increased cancer risk in the range of approximately 1.5-10) (see Mohrenweiser & Jones, 1998 *and refs therein*). For comparison, patients with XP generally show 1-2% normal NER capacity and a cancer risk approaching unity. Although these studies have not assessed the contributions of individual gene variants to cancer risk, they do suggest a genetic basis for individual differences in repair capacity; intraindividual variation is much lower than interindividual variation, distinct pathways show consistent and independent repair capacity phenotypes within individual subjects, and there is evidence for an increased likelihood of reduced DNA repair capacity amongst relatives of cancer patients.

24. With regard to the general population, limited data from a relatively small number of epidemiological studies have indicated that reduced DNA repair capacity (as a phenotype) may be a contributory factor in skin and smoking-related cancers (see Benhamou & Sarasin, 2000 *and refs therein*), although larger-scale studies are required to further evaluate these associations. Phenotypic assays have also shown decreasing DNA repair capacity with age (Wei et al., 1993; Liu et al., 1994), which has been suggested as indicative of reduced DNA repair as an acquired risk factor contributing to increasing cancer risk with age (Simpson, 1997; Shields & Harris, 2000).

#### *Variations in DNA repair genes*

25. Although altered DNA repair capacity has been demonstrated to exhibit polymorphic variation within the population, suggesting the existence of relatively common variant alleles within these pathways, there have been few reports to date of studies of the associations of specific alleles and cancer susceptibility. Researchers have begun to characterise polymorphic variants in several known DNA repair genes, but very few functional biochemical or epidemiological studies have been carried out. For example, Broughten et al. (1996) reported 5 polymorphic variants within the coding sequence of the *XPD* gene. More-recently, Mohrenweiser and colleagues have reported preliminary data from a study to characterise interindividual variation in several known DNA repair genes. Resequencing of the exons of five NER genes, two BER genes, a gene involved in double strand break repair/recombination and a gene functioning in both BER and the repair of radiation induced damage revealed variants encoding 15 different amino acid substitution variants (Mohrenweiser & Jones, 1998; Shen et al., 1998). In an initial sample of 12 healthy individuals, the frequencies for the 15 different variant alleles ranged from 0.02-0.42, with an average allele

frequency of 0.14 – sufficiently prevalent to be of potential significance at the population level. The functional consequences of most of these variants have yet to be ascertained; although none is identical to alleles previously associated with a disease state, several encode non-conservative substitutions at residues which are normally conserved between human and mouse, suggesting a potential to alter functional DNA repair capacity.

26. A small number of case-control, epidemiological studies have assessed the potential association of one *XPD* gene variant (an exon 23 A → C polymorphism encoding substitution of Lys → Gln at amino acid residue 751; average population frequency 0.32) with (lung and skin) cancer (Dybdahl et al., 1999; Escobar et al., 1999; Wu et al., 1999). The findings of these preliminary studies were, however, inconclusive, whilst the functional significance of the polymorphism is not known.

27. In summary, from the paradigm of inactivating DNA repair gene mutations in high penetrance cancer syndromes such as XP, it would be logical to assume that a genotype encoding functionally reduced DNA repair capacity may also predispose to increased cancer risk, although data to support this hypothesis are currently limited. Very few studies have addressed the issue of the importance of genetic variation in DNA repair pathways with respect to specific chemical susceptibility.

### **Variability in immune function**

#### *Cancer and the immune system*

28. In the process of malignant transformation, if a cell has failed to repair or to apoptose, it may accumulate mutations resulting in the expression of aberrant protein products which could be recognised and eliminated by the host immune system (Bremers & Parmiani, 2000). However, if the immune system is compromised, or if a tumour cell can avoid immune surveillance or response, then that cell may survive to accumulate further detrimental mutations. Hence, tumour immunology consists of two essential concepts: immune surveillance and immune escape. The relevance of the immune response in protecting from tumour development may be indicated by the high incidence of certain cancers (for example Kaposi's sarcoma and non-Hodgkin's lymphoma) associated with immunosuppression disorders and in patients given long-term immunosuppression therapy (Beral & Newton, 1998). By implication, any genetic variation resulting in a functionally reduced ability of the host immune system to recognise and/or destroy tumour cells, or an increased potential for such cells to avoid an immune response, would be a potential target for conferring increased susceptibility to the development of a clinically-relevant tumour.

#### *Mechanisms of tumour immune response – the importance of the MHC Class I/ cytotoxic T lymphocyte response*

29. Antigens can be processed to induce cell-mediated immunity, essentially mediated by cytotoxic T lymphocytes (CTL) or to produce a humoral immune response mediated by antibodies. There is now strong evidence to regard CTL-mediated immunity as the major anti-tumour immune mechanism. The ability to trigger a CTL response is an exclusive property of a specialised type of cell referred to

as professional antigen presenting cells (APC). By presenting antigens through the MHC class I pathway on APCs, naïve lymphocytes are activated into CTLs (Figure 5). The proteins that constitute the MHC are encoded by genes clustered in the MHC locus which is, in humans, highly-polymorphic. This locus contains the human leukocyte antigen (HLA) complex, as well as genes encoding other molecules required for antigen processing and presentation. In addition to CTL-mediated immunity, an anti-tumour response may also involve factors, which are normally suppressed as part of the innate immune system (e.g. natural killer - NK - cells) and/or a specific antibody response (antibody dependent cellular cytotoxicity - ADCC) against oncogene or tumour suppressor gene products (e.g. K-Ras, Her2/neu, p53).

### *Mechanisms of tumour escape*

30. Several mechanisms have been recognised which allow malignant cells to escape from a normally functioning immune system (see review by Bremers & Parmiani, 2000). These may be summarised, generally, as follows;

1. Recognition and selection – for example;

- a] loss of previously recognised tumour antigens – “antigen-loss variants”.
- b] defective antigen processing due to alteration in gene(s) encoding molecule(s) crucial for antigen presentation and peptide transport (e.g. TAP [transporter associated with antigen presentation]), resulting in lack of antigen presentation at the cell surface
- c] down-regulation of MHC molecules, inhibiting recognition of the cell by T lymphocytes.

2. Down-regulation of the immune response – a process which occurs in some essential organs (e.g. the liver) under normal physiological conditions. The effect is obtained *via* the local release of inhibitory molecules (e.g. TGF $\beta$ , Fas ligand [FasL]), which interact with receptors or Fas on the T-lymphocyte surface, resulting in apoptosis of the T cell. Aberrant signalling by such factors has been reported for a range of malignancies, including lung and colon carcinoma, ovarian and oesophageal cancer, melanoma, astrocytoma and glioblastoma.

3. Tolerance induction – the activation of mechanisms which cause immune recognition to result in (T-cell) tolerance (lack of activation in response to a specific HLA/antigen complex), for example;

- a] absence of co-stimulatory molecules.
- b] failure to provide the optimal “danger” signalling microenvironment and associated cytokines.

### *Genetic variation and tumour escape?*

31. Genetic polymorphism which may favour a functional increase in the potential for malignant cells to escape from immune detection or response is, logically, likely to affect individual susceptibility to the development of a clinically relevant cancer. A major target for the investigation of such variation is the highly polymorphic MHC region (located on human chromosome 6) which contains > 200 genes, including the HLA genes and many others with immunological functions in antigen processing and presentation (Bateman & Howell, 1999; Little & Stern, 1999). An individual can express one of numerous different HLA allotypes arising from the potential contribution of 12 different HLA Class I (*HLA-A, B* and *C*) and Class II (*HLA-DR, DQ* and *DP*) loci, for which > 800 different alleles have been defined to date

(variation between alleles is restricted predominantly to the exons encoding the peptide-binding domain of the protein) (see the IMGT/HLA database at <http://www.ebi.ac.uk/imgt/hla/>). From the concept of tumour immune surveillance, it follows that an individual's HLA type could influence the risk of developing a particular cancer(s). The function of the HLA molecule encoded by a particular allele may differ significantly from the molecules encoded by alternative alleles in its ability to present antigens to T-lymphocytes. For example, the observation that the *HLA-DQB1\*0301* allele may be associated with colorectal carcinomas of less advanced stage has been suggested to be due to the ability of the encoded molecule to recognise and present processed peptides containing a particular K-ras mutation more efficiently than other HLA molecules (Fossum et al., 1994). There is also evidence that certain mutations in the *TP53* gene which could generate epitopes presentable by HLA-A2 are under-represented in lung cancer patients with this genotype (Weidenfeld et al., 1994). Alternatively, sequence polymorphisms within the regulatory regions of *HLA* genes may differentially affect the expression of various alleles for each locus (Thorsby, 1995), which could also affect the efficiency of antigen presentation to T lymphocytes.

32. Because of the potential contribution of 12 different, polymorphic HLA allotypes, studies of the correlations of specific HLA types with particular cancers are clearly not straightforward. In fact, the majority of studies to date have focussed on investigating associations of HLA types with cancers which show strong associations with specific viral infections (e.g. HPV, HBV, HTLV-I/II, EBV) (Table 4). However, although studies have clearly demonstrated associations between HLA genotype and susceptibility to virus infection *per se*, strong associations between HLA types and the emergence of malignant disease, independent of viral susceptibility, have not been identified.

33. Apart from the classical HLA complex, the MHC region on chromosome 6 also encodes a number of other non-classical HLA, and non-HLA genes, variation in which may, at least hypothetically, alter the effectiveness of the anti-tumour immune response. Candidates include genes encoding proteins involved with intracellular synthesis and processing of HLA molecules, which may affect antigen recognition and selection (for example, TAP, LMP [large multifunctional protease] and HLA-DM, encoded within the MHC Class II region) (see Bateman & Howell, 1999). Alternatively, polymorphisms leading to aberrant expression of certain cytokines by tumour cells or tumour-infiltrating lymphocytes or macrophages may favour suppression of the anti-tumour response, facilitating T-cell tolerance induction. For example, some studies have reported an association of increased TNF- $\alpha$  expression (encoded within the MHC Class III region) with tumour promotion *in vivo* and poor clinical prognosis of some cancers (Gadducci et al., 1995; Partanen et al., 1995; Salles et al., 1996). Bi-allelic polymorphisms in the promoter regions of the closely linked genes encoding TNF- $\alpha$  (TNF) and LT- $\alpha$  (TNF- $\beta$ ) have been shown to affect *TNF* transcription. The "high expressor" haplotype at these two loci (which occurs in strong linkage disequilibrium with the HLA-A1, B8, DR3 haplotype), would thus be a candidate for negatively affecting the tumour immune response, and has been implicated in susceptibility to breast cancer, non-Hodgkin's lymphoma and myeloma (Chouchane et al., 1997; Davies et al., 2000) and to the progression of some haematological malignancies (Demeter et al., 1997; Warzocha et al., 1998).

Polymorphisms in hsp70 (heat shock protein 70)-encoding genes, again located within the MHC Class III region, have also been reported to be associated with susceptibility to breast cancer and non-Hodgkin's lymphoma (Chouchane et al., 1997).

34. In summary, any genetic polymorphism which may favour a functional increase in the potential for malignant cells to escape from immune detection or response is, at least hypothetically, likely to affect individual susceptibility to the development of a clinically relevant cancer. The extent to which such polymorphisms may interact with environmental chemicals has not been evaluated.

### **Variability at loci involved in the control of DNA replication, cell growth and survival**

35. The cell cycle is controlled by multiple mechanisms on which exogenous and endogenous stimuli converge. Pathways governing the different phases of the cell cycle are central for the decision of the cell to commit to DNA synthesis and proliferation versus growth arrest, DNA repair or apoptosis. These pathways incorporate various oncogenes and tumour suppressor genes and are a central target for genetic alterations in cancers. Such alterations may ultimately lead to aberrant cell proliferation and increased genetic instability (Funk, 1999). Inherited or acquired alterations in genes which are key to pathways which regulate cell growth and survival have been identified in many cancers. For example mutation/loss of *TP53* (the product of which is central to several cellular processes including gene transcription, cell cycle arrest, senescence and apoptosis) has been demonstrated in  $\approx$  50% of sporadic tumours (Hussain & Harris, 2000), whilst inherited mutations in this gene underlie the familial Li-Fraumini cancer syndrome (Evans & Lozano, 1997). Other examples include alterations at loci encoding or regulating the expression of products involved in cell-cycle control (e.g. cyclin D1, retinoblastoma [RB]), apoptosis/cell death (e.g. Bcl-2), growth factor or hormone signalling (e.g. EGF-receptor family and ligands; Ras proteins; androgen [AR] and oestrogen [ER] receptors and their ligands, and the (re)activation of anti-senescence genes (e.g. telomerase) (see, for example, review by Michalides, 1999). It is also hypothetically possible that such loci could show subtle mutations conferring less dramatic variation, which would not be apparent in sporadic or high penetrance tumours, but may affect risk in association with exposure to a carcinogen, for example by reducing time to tumour appearance.

36. Such variants would provide likely targets for future evaluation as low penetrance cancer susceptibility alleles. Indeed, some studies have already shown positive associations of variant alleles with cancer risk. For example, an association has been reported between rare alleles at the highly polymorphic *HRAS* (*HRAS1*) VNTR minisatellite locus and some cancers (colorectal, breast, lung, bladder, leukaemia). It is proposed that the variations at this minisatellite (located  $\approx$  1 kb downstream of the gene's coding sequence, and composed of 30-100 units of a 28 bp consensus sequence) interfere with regulatory mechanisms governing the control of nearby genes, including *HRAS* gene expression (Krontiris et al., 1993). Another example is the inverse association reported by some authors of a variable number CAG repeat polymorphism within the amino-terminal-encoding region of the *AR* (androgen receptor) gene and susceptibility to the development or progression of

prostate cancer (Giovannuci et al., 1997; Stanford et al., 1997, Kantoff et al., 1998). In this case, the amino-terminal portion of the androgen receptor protein contains a transcriptional activation domain, which regulates the level of expression of target genes, and it is postulated that CAG length affects androgen receptor transactivation function, conferring variable androgenicity upon the prostate over a lifetime.

37. There is now a growing number of reports describing the identification and (to a small extent) evaluation of polymorphisms at other important cancer loci. Taking breast cancer as an example, recent reports have suggested polymorphic variants of genes such as *ERBB2* (Ameyaw et al., 2000; Xie et al., 2000), *PTEN* (Carroll et al., 1999), *BRCA2* (Healey et al., 2000), *TP53* (Sjalander et al., 1996; Wang-Gohrke et al., 1998) and *LMYC* (Togo et al., 2000) as candidate low penetrance, or modifier genes for this disease. However, the majority of such studies to date have been preliminary and, perhaps in some cases, based on the identification of a variant allele in a known cancer gene *per se*, rather than a biologically plausible hypothesis. The extent to which any of these variants may interact with environmental factors has yet to be determined.

### **Genetic variation and nutritional factors**

38. There is considerable evidence that dietary factors contribute to the aetiology of many cancers (Department of Health, 1998), although the mechanisms underlying such associations and the specific food components responsible for the observed effects are mostly not well elucidated. In addition, some dietary components have been implicated in cancer chemoprevention or chemoprotection. A recent area of study in this field involves the investigation of the effects of genetic factors in modifying individual response to dietary carcinogens or anticarcinogens (Sinha & Caporaso, 1999; Rock et al., 2000) (Table 5). As discussed earlier in this report, polymorphisms exist in genes encoding enzymes involved in the activation or elimination of dietary carcinogens. For example activation of heterocyclic amines (a group of procarcinogens found in meats cooked at high temperature [Sugimura, 2000]), is hypothesised to be mediated by CYP1A2 and NAT2, both of which show variation within the population. Some authors have suggested that “rapid-metabolising” alcohol dehydrogenase (ADH) variants may predispose to increased cancer risk in association with alcohol intake, *via* a proposed mechanism of increased activation of ethanol to the reactive metabolite, acetaldehyde (Harty et al., 1997). Diet-gene interactions may also involve variation in genes involved in vitamin metabolism. For example, polymorphisms in non-coding regions of the gene encoding the vitamin D receptor (involved in vitamin D and calcium metabolism) have been linked to prostate cancer risk (Taylor et al., 1996; Ingles et al., 1997) and have been suggested as possible candidates for modulating susceptibility to colorectal cancer (see Sinha & Caporaso, 1999). Recent studies have suggested that individuals who have a methyl-deficient diet (low in folate and methionine and high in alcohol) are at higher risk of colon cancer if they carry a defective polymorphic form of the enzyme methylene tetrahydrofolate reductase (MTHFR) (Chen et al., 1999). Future studies for diet-gene interactions and cancer susceptibility may include investigations of polymorphic variations in genes involved in energy metabolism or lipid metabolism. A variety of mechanisms have been implicated in chemoprevention and chemoprotection, including modification of stress responses and cell growth. Hence,

polymorphisms in any of the genes involved could adversely influence the response of an individual to such factors and thereby increase susceptibility.

### **Future studies for the identification of low penetrance cancer susceptibility genes**

40. The majority of high penetrance cancer genes identified to date appear to encode products which act in a cell autonomous fashion (for example, tumour suppressor genes such as *APC*, *RBI* and *TP53*). Such genes have been identified by classical genetic linkage analysis of Mendelian inheritance in affected “cancer families”. This approach is not applicable to the identification of low penetrance cancer susceptibility loci, which are associated with minimally increased relative risks (see CC/01/3). To date, the major approach used for the identification of low penetrance alleles has been to conduct association studies using candidate loci. Such loci may include genes with cell autonomous effects, but are also likely to include those with broader effects (for example, carcinogen metabolism polymorphisms, behavioural factors, anti-tumour immune response) or with effects on the local tumour environment. Modifier loci for high penetrance cancer genes may also function as low penetrance susceptibility loci. For example, substantial phenotypic variability is seen amongst carriers of identical *APC* mutations and also in (Min) mouse models of FAP (Fodde et al., 1999). It is likely that a gene which modifies the number of adenomas in FAP may also influence the probability of an individual developing a (sporadic) colorectal adenoma, and hence a carcinoma, and such a gene may act as a low penetrance susceptibility gene if it is sufficiently variable within the population. Similarly, modifier traits such as those which modulate the progression of a benign lesion to malignancy or the propensity to metastasise may also act as low penetrance susceptibility genes (see Houlston & Tomlinson, 2000). Future studies for the identification of such cancer modifier loci may be facilitated by the use of animal models.

41. To date, the majority of candidate low penetrance cancer susceptibility gene studies have focussed on genes associated with the activation, detoxication and elimination of carcinogens. Researchers are beginning to investigate functional variation in DNA repair capacity, although there are currently few data regarding the involvement of individual variant alleles. There is also a growing body of literature describing the identification of polymorphic alleles at a variety of other cancer gene loci (for example, loci which have been shown to be somatically mutated in cancer, or genes involved in critical cell regulatory pathways), although such variants have mostly not yet been evaluated in large-scale association or functional studies.

42. In addition to the evaluation of individual candidate gene loci, the availability of data from large scale human genome mapping projects (such as high density marker maps of single nucleotide polymorphisms [SNPs]) will facilitate the performance of systematic, genome wide association studies for the identification of low penetrance cancer susceptibility loci (see Houlston & Tomlinson, 2000). Large-scale projects which aim to evaluate polymorphic variation at many important candidate gene loci are also being undertaken. For example, the Genetic Annotation Initiative (GAI) of the NCI's Cancer Genome Anatomy Project (CGAP) is a research programme aimed at the identification and characterisation of genetic variation in cancer-related genes (see <http://lpg.nci.nih.gov/>). The US National Institute of

Environmental Health Sciences (NIEHS) Environmental Genome Project (Kaiser, 1997; see <http://www.niehs.nih.gov/envgenom/home.htm>) aims to identify and evaluate functionally important polymorphic variations in genes that may play an important role in environmentally associated diseases (“environment response genes”) (Table 6). Similar research projects, with the aim to identify, characterise and perform large-scale association studies of candidate low penetrance cancer genes, are also being carried out within other major cancer research institutes, such as the International Agency for Cancer Research (IARC) (see <http://www.iarc.fr/>) (see CC/01/5 for further discussion).

### **Overview**

43. There is currently a great interest in low-penetrance cancer susceptibility genes, both as targets of interest for cancer researchers in investigating genes involved in the causation of common sporadic tumours (the field of cancer molecular epidemiology), and for those involved in the assessment of risk and the determination and implementation of public health policies. There is clearly an enormous range of target loci, the vast majority of which have not yet been identified and/or evaluated. Future studies should aid in increasing knowledge in the field of cancer research and will hopefully lead to the development of targets for chemoprevention and cancer therapy. The application of this knowledge to public health policy is addressed in the accompanying report (CC/01/5).

### **Conclusions and recommendations**

44. To date, most studies of gene-environment interactions in cancer causation have focussed on enzymes of xenobiotic metabolism. Whilst, in principal, many such interactions are plausible, the relative risks identified have been low to negligible. Consideration should be given to the value of further studies of this nature.

45. As understanding develops of the genes involved in the biological processes critical to cancer development, the number of pathways and of candidate genes within those pathways, that may be relevant to study for interaction with environment is increasing rapidly. It is not difficult to envisage biological plausibility for any one of these processes, or for any one of the genes involved. Hence, it is not possible to prioritise targets for candidate selection. The committee may wish to consider whether further work may involve generating some hierarchy of effects, that might help prioritise candidate gene selection.

46. Whilst much of the research on gene-environment interactions to date has concentrated on structural mutations, there is an increasing recognition of the importance of regulatory mutations. Hence, candidate genes will include not only those encoding products involved directly in cancer development, but also those involved in regulating the expression of these genes, and of their transcription products (mRNA stability, post-translational modification). Again, further work might include consideration of likelihood of an important role for a given gene in a regulatory network. However, selection is likely to be pragmatic, on the basis of polymorphisms that influence activity or expression.

## **References**

- Ameyaw, M.M., Thornton, N., McLeod, H.L. (2000). Re: population-based, case-control study of HER2 genetic polymorphism and breast cancer risk. *J. Natl. Cancer Inst.*, 92, 1947.
- Autrup, H. (2000). Genetic polymorphisms in human xenobiotica metabolizing enzymes as susceptibility factors in toxic response. *Mutat. Res.*, 464, 65-76.
- Bateman, A.C., Howell, W.M. (1999). Human leukocyte antigens and cancer: is it in our genes? *J. Pathol.*, 188, 231-236.
- Benhamou, S., Sarasin, A. (2000). Variability in nucleotide excision repair and cancer risk: a review. *Mutat. Res.*, 462, 149-158.
- Beral, V., Newton, R. (1998). Overview of the epidemiology of immunodeficiency-associated cancers. *Monogr. Natl. Cancer Inst.*, 23, 1-6.
- de Boer, J., Hoeijmakers, J.H.J. (2000). Nucleotide excision repair and human syndromes. *Carcinogenesis*, 21, 453-460.
- Bremers, A.J.A., Parmiani, G. (2000). Immunology and immunotherapy of human cancer: present concepts and clinical developments. *Crit. Rev. Oncol. Haematol.*, 34, 1-25.
- Broughton, B.C., Steingrimsdottir, H., Lehmann, A.R. (1996). Five polymorphisms in the coding sequence of the xeroderma pigmentosum group D gene. *Mutat. Res.*, 362, 209-211.
- Carroll, B.T., Couch, F.J., Rebbeck, T.R., Weber, B.L. (1999). Polymorphisms in PTEN in breast cancer families. *J. Med. Genet.*, 36, 94-96.
- Chen, J., Giovannucci, E.L., Hunter, D.J. (1999). MTHFR polymorphism, methyl-replete diets and the risk of colorectal carcinoma and adenoma among U.S. men and women: an example of gene-environment interactions in colorectal tumorigenesis. *J. Nutr.*, 129, 560S-564S.
- Chouchane, L., Ben Ahmed, S., Baccouche, S., Remadi, S. (1997). Polymorphism in the tumor necrosis factor- $\alpha$  promoter region and in the heat shock protein 70 genes associated with malignant tumors. *Cancer*, 80, 1489-1496.
- CC/01/03 (2001) Criteria for the design of gene-environment epidemiology studies. COC meeting, 22 March 2001.
- CC/01/05 (2001) How gene-environment studies should be used in risk assessment. COC meeting, 22 March 2001.
- Davies, F.E., Rollinson, S.J., Rawstron, A.C. et al. (2000). High-producer haplotypes of tumor necrosis factor alpha and lymphotoxin alpha are associated with an increased

risk of myeloma and have an improved progression-free survival after treatment. *J. Clin. Oncol.*, 18, 1843-1851.

Demeter, J., Porzolt, F., Ramisch, S. et al. (1997). Polymorphism of the tumour necrosis-alpha and lymphotoxin genes in chronic lymphocytic leukaemia. *Mr. J. Haematol.*, 97, 107-112.

Department of Health (1998). Nutritional aspects of the development of cancer. Report on Health and Social Subjects 48. TSO, London.

Dybdahl, M., Vogel, U., Frentz, G. et al. (1999). Polymorphisms in the DNA repair gene *XPD*: correlations with risk and age at onset of basal cell carcinoma. *Cancer Epidemiol. Biomarkers Prev.*, 8, 77-81.

Escobar, P., Modugno, F., Kanbour-Shakir, A. et al. (1999). Analysis of the DNA repair enzyme *XPD* exon 23 polymorphism and cancer risk. *Proc. Am. Assoc. Cancer Res.*, 40, 213.

Evans, S.C., Lozano, G. (1997). The Li-Fraumeni syndrome: an inherited susceptibility to cancer. *Mol. Med. Today*, 3, 390-395.

Fearon, E.R. (1997). Human cancer syndromes: clues to the origin and nature of cancer. *Science*, 278, 1043-1050.

Fodde, R., Smits, R., Hofland, N. et al. (1999). Mechanisms of APC-driven tumorigenesis: lessons from mouse models. *Cytogenet. Cell Genet.*, 86, 105-111.

Fossum, B., Breivick, J.H., Meling, G.I. et al. (1994). A K-ras 13 Gly-Asp mutation is recognised by HLA-DQ7 restricted T cells in a patient with colorectal cancer. Modifying effect of DQ7 on established cancer harbouring this mutation? *Int. J. Cancer*, 58, 506-511.

Friedberg, E.C., Walker, G.C., Siede, W. (1995). DNA repair and mutagenesis. ASM Press, Washington, DC.

Funk, J.O. (1999). Cancer cell cycle control. *Anticancer Res.*, 19, 4772-4780.

Gadducci, A., Ferdeghini, M., Castellani, M. et al. (1995). Serum levels of tumour necrosis factor (TNF), soluble receptors for TNF (55 and 75 kDa sTNFr) and soluble CD14(sCD14) in epithelial ovarian cancer. *Gynecol. Oncol.*, 58, 184-188.

Giovannucci, E., Stampfer, M.J., Krithivas, K. et al. (1997). The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc. Natl. Acad. Sci. USA*, 94, 3320-3323.

Gonzalez, F.J. (1997). The role of carcinogen-metabolizing enzyme polymorphisms in cancer susceptibility. *Reproductive Toxicology*, 11, 397-412.

- Guilford, P. (2000). The inherited susceptibility to cancer. *Cell. Mol. Life Sci.*, 57, 589-603.
- Harty, L.C., Caporaso, N.E., Hayes, R.B. et al. (1997). Alcohol dehydrogenase 3 genotype and risk of oral cavity and pharyngeal cancers. *J. Natl. Cancer Inst.*, 89, 1698-1705.
- 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, 362-364.
- Houlston, R.S., Tomlinson, I.P.M. (2000). Detecting low penetrance genes in cancer: the way ahead. *J. Med. Genet.*, 37, 161-167.
- Hussain, S.P., Harris, C.C. (2000). Molecular epidemiology and carcinogenesis: endogenous and exogenous carcinogens. *Mutat. Res.*, 462, 311-322.
- International Agency for Research on Cancer (1999). Metabolic polymorphisms and susceptibility to cancer (eds. Vineis, P. et al.). IARC Scientific Publication No. 148, IARC Press, Lyon, France.
- Ingles, S.A., Ross, R.K., Yu, M.C. et al. (1997). Association of prostate cancer risk with genetic polymorphisms in vitamin D receptor and androgen receptor. *J. Natl. Cancer Inst.*, 89, 166-170.
- Jeggo, P.A., Taccioli, G.E., Jackson, S.P. (1995). Menage 'a trois: double strand break repair, V9D)J recombination and DNA-PK. *Bioessays*, 17, 949-957.
- Jiricny, J., Nystrom-Lahti, M. (2000). Mismatch repair defects in cancer. *Curr. Opin. Genet. Dev.*, 10, 157-161.
- Kaiser, J. (1997). Environment institute lays plans for gene hunt. *Science*, 278, 569-570.
- Kantoff, P., Giovannucci, E., Brown, M. (1998). The androgen receptor CAG repeat polymorphism and its relationship to prostate cancer. *Biochim. Biophys. Acta*, 1378, C1-C5.
- Kinzler, K.W., Vogelstein, B. (1997). Cancer susceptibility genes: gatekeepers and caretakers. *Nature*, 386, 761-763.
- Krontiris, T.G., Devlin, B., Karp, D.D. et al. (1993). An association between the risk of cancer and mutations in the *HRAS1* minisatellite locus. *N. Engl. J. Med.*, 329, 517-523.
- Lehmann, A.R. (1996). Molecular biology of DNA repair in the fission yeast *Schizosaccharomyces pombe*. *Mutat. Res.*, 363, 147-161.

- Little, A.-M., Stern, P.L. (1999). Does HLA type predispose some individuals to cancer? *Mol. Med. Today*, 5, 337-342.
- Liu, Y., Hernandez, A.M., Shibata, D. et al. (1994). BCL2 translocation frequency rises with age in humans. *Proc. Natl. Acad. Sci. USA*, 91, 8910-8914.
- Memisoglu, A., Samson, L. (2000). Base excision repair in yeast and mammals. *Mutat. Res.*, 451, 39-51.
- Michalides, R.J.A.M. (1999). Cell cycle regulators: mechanisms and their role in aetiology, prognosis, and treatment of cancer. *J. Clin. Pathol.*, 52, 555-568.
- Mohrenweiser, H.W., Jones, I.M. (1998). Variation in DNA repair is a factor in cancer susceptibility: a paradigm for the promises and perils of individual and population risk estimation? *Mutat. Res.*, 400, 15-24.
- Olden, K., Wilson, S. (2000). Environmental health and genomics: visions and implications. *Nature Reviews*, 1, 149-153.
- Ottman, R. (1996). Gene-environment interaction: definitions and study designs. *Preventive Medicine*, 25, 764-770.
- Partanen, R., Koskinen, H., Hemminki, K. (1995). Tumor necrosis factor-alpha (TNF-alpha) in patients who have asbestosis and develop cancer. *Occup. Environ. Med.*, 52, 316-319.
- Perera, F.P. (1997). Environment and cancer: who are susceptible? *Science*, 278, 1068-1073.
- Rock, C.L., Lampe, J.W., Patterson, R.E. (2000). Nutrition, genetics, and risks of cancer. *Annu. Rev. Public Health*, 21, 47-64.
- Salles, G., Bienvenu, G., Bastion, Y. et al. (1996). Elevated circulating levels of TNF $\alpha$  and its p55 soluble receptor are associated with an adverse prognosis in lymphoma patients. *Br. J. Haematol.*, 93, 352-359.
- Shen, M.R., Jones, I.M., Mohrenweiser, H. (1998). Non-conservative amino acid substitution variants exist at polymorphic frequency in DNA repair genes in healthy humans. *Cancer Res.*, 58, 604-608.
- Shields, P.G., Harris, C.C. (2000). Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. *J. Clin. Oncol.*, 18, 2309-2315.
- Simpson, A.J. (1997). The natural somatic mutation frequency and human carcinogenesis. *Adv. Cancer Res.*, 71, 209-240.
- Sinha, R., Caparaso, N. (1999). Diet, genetic susceptibility and human cancer etiology. *J. Nutr.*, 129, 556S-559S.

- Sjalander, A., Birgander, R., Hallmans, G. et al. (1996). p53 polymorphisms and haplotypes in breast cancer. *Carcinogenesis*, 17, 1313-1316.
- Stanford, J.L., Just, J.J., Gibbs, M. et al. (1997). Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Res.*, 57, 1194-1198.
- Sugimura, T. (2000). Nutrition and dietary carcinogens. *Carcinogenesis*, 21, 387-395.
- Taningher, M., Malacarne, D., Izzotti, A. et al. (1999). Drug metabolism polymorphisms as modulators of cancer susceptibility. *Mutat. Res.*, 436, 227-261.
- Taylor, J.A., Hirvonen, A., Watson, M. et al. (1996). Association of prostate cancer with vitamin D receptor gene polymorphism. *Cancer Res.*, 56, 4108-4110.
- Thorsby, E. (1995). HLA-associated disease susceptibility – which genes are primarily involved? *Immunologist*, 3, 51-58.
- Togo, A.V., Suspitsin, E.N., Grigoriev, M.Y. et al. (2000). *L-myc* polymorphism in cancer patients, healthy and elderly, tumor-free individuals in Russia. *Int. J. Cancer*, 85, 747-750.
- Wang-Gohrke, S., Rebbeck, T.R., Besenfelder, W. et al. (1998). p53 germline polymorphisms are associated with an increased risk for breast cancer in German women. *Anticancer Res.*, 18, 2095-2099.
- Warzocha, K., Ribeiro, P., Bienvenu, J. et al. (1998). Genetic polymorphisms in the tumor necrosis factor locus influence non-Hodgkin's lymphoma outcome. *Blood*, 92, 3574-3581.
- Wei, Q., Matanoski, G.M., Farmer, E.R. et al. (1993). DNA repair and aging in basal cell carcinoma: A molecular epidemiology study. *Proc. Natl. Acad. Sci. USA*, 90, 1614-1618.
- Weidenfeld, E.A., Fernandez-Vinas, M., Berzofsky, J.A., Carbone, D.P. (1994). Evidence for selection against human lung cancers bearing p53 missense mutations which occur with HLA-A\*0201 peptide consensus motif. *Cancer Res.*, 54, 1175-1177.
- Wu, X., Amos, C.I., Wang, M. et al. (1999). Association of a polymorphism in exon 23 of the DNA repair gene, XPD, and lung cancer risk. *Proc. Am. Assoc. Cancer Res.*, 40, 89.
- Xie, D., Shu, X.-O., Deng, Z. et al. (2000). Population-based, case-control study of HER2 genetic polymorphism and breast cancer risk.

**Table 1.** Summary of genes responsible for inherited cancer predisposition, their chromosomal location, the syndromes they cause and the functions of the gene products (*mode of inheritance shown in italics*) (from Guilford, 2000).

Gene	Syndrome	Location	Principal function	Principal malignancies
<i>RB1</i>	familial retinoblastoma; <i>dominant</i>	13q14	transcriptional/ cell cycle regulator	retinoblastoma
<i>P16<sup>INK4a</sup></i>	familial melanoma; <i>dominant</i>	9p21	CDK inhibitor	melanoma
<i>CDK4</i>	familial melanoma; <i>dominant</i>	12q13	CDK	melanoma
<i>P53</i>	Li-Fraumeni; <i>dominant</i>	17p13.1	transcription factor	sarcomas, breast cancer
<i>APC</i>	familial adenomatous polyposis; <i>dominant</i>	5q21	growth factor signalling	colorectal cancer
<i>CDH1</i>	hereditary diffuse gastric cancer; <i>dominant</i>	16q22.1	cell-to-cell adhesion	diffuse gastric cancer
<i>LKB1</i>	Peutz-Jeghers; <i>dominant</i>	19p13.3	serine threonine kinase	gastrointestinal cancer
<i>PTEN</i>	Cowden syndrome; juvenile polyposis coli; <i>dominant</i>	10q23.3	phosphatase, cytoskeletal protein?	breast cancer, gastrointestinal cancer
<i>SMAD4</i>	juvenile polyposis coli; <i>dominant</i>	18q21.2	growth factor signalling	gastrointestinal cancer
<i>MEN1</i>	multiple endocrine neoplasia type 1; <i>dominant</i>	11q13	transcription co-factor	endocrine
<i>RET</i>	multiple endocrine neoplasia type 2; <i>dominant</i>	10q11.2	receptor tyrosine kinase	endocrine
<i>MET</i>	Hereditary papillary renal cancer; <i>dominant</i>	7q31	receptor tyrosine kinase	papillary renal cancer
<i>KIT</i>	familial gastrointestinal stromal tumours; <i>dominant</i>	4q12	receptor tyrosine kinase	gastrointestinal cancer (stromal)
<i>PTCH</i>	basal cell nevus syndrome; <i>dominant</i>	9q22.3	membrane receptor	basal cell (skin)
<i>NF1</i>	neurofibromatosis type 1; <i>dominant</i>	17q11.2	GTPase-activating protein	neurofibrosarcomas
<i>NF2</i>	neurofibromatosis type 2; <i>dominant</i>	22q12.2	cytoskeletal protein?	central nervous system tumours
<i>VHL</i>	von Hippel-Lindau <i>dominant</i>	3p25	protein maturation? RNA elongation?	renal clear cell carcinomas, pheochromocytomas
<i>WT1</i>	Wilms tumour; <i>dominant</i>	11p13	transcription factor	nephroblastoma
<i>BLM</i>	Bloom syndrome; <i>recessive</i>	15q26.1	dsDNA repair?	leukaemia, lymphoma
<i>FANCA; FANCC; others</i>	Fanconi anaemia; <i>recessive</i>	16q24.3; 9q22.3; ?	dsDNA repair?	leukaemia
<i>XPB; XPD others</i>	xeroderma pigmentosum; <i>recessive</i>	2q21; 19q13; ?	helicases, nucleotide excision repair	basal cell and squamous cell carcinomas
<i>ATM</i>	ataxia telangiectasia; <i>recessive</i>	11q22.3	serine-threonine protein kinase	lymphoma, leukaemia
<i>NBS1</i>	Nijmegen breakage syndrome; <i>recessive</i>	8q21	transcription factor?	lymphoma
<i>BRCA1</i>	familial breast/ovarian cancer; <i>dominant</i>	17q21	dsDNA repair? transcription factor?	breast, ovarian cancer
<i>BRCA2</i>	familial breast/ovarian cancer; <i>dominant</i>	13q12	dsDNA repair? transcription factor?	breast, ovarian cancer
<i>MLH1; MSH2; PMS1; PMS2; MSH6</i>	hereditary non-polyposis colorectal cancer; <i>dominant</i>	3p21; 2p16; 2q32; 7p22; 2p16	dsDNA repair? DNA mismatch repair	colorectal, endometrial cancer

CDK, cyclin-dependent kinase.

**Table 2. Evaluation of the available evidence on the associations between selected metabolic polymorphisms and human cancers. (from IARC, 1999).**

	<i>CYP1A1</i>	<i>CYP1A1</i>	<i>CYP1A1</i>	<i>CYP1A2</i>	<i>CYP2D6</i>	<i>CYP2E1</i>	<i>CYP2E1</i>	<i>GSTM1</i>	<i>GSTT1</i>	<i>NAT1</i>	<i>NAT2</i>
<i>Polymorphism</i>	<i>Msp I</i>	<i>Exon 7</i>	<i>AHH ind</i>	<i>Rapid</i>	<i>EM</i>	<i>5'FR</i>	<i>Dra I</i>	<i>Null</i>	<i>Null</i>	<i>Slow</i>	<i>Slow</i>
Lung	+ A (=) C	(+) A (=) C	+	NA	+ P = G	(-)	(=)	+	(-)	NA	=
Bladder	(=)	(=)	NA	(+)	=	NA	(=)	(+)	(-)	NA	= A + C
Breast	(+)	(+)	NA	NA	(=)	(=)	(=)	(+)	NA	NA	(-) P (=) G
Colorectum	(+)	NA	NA	(+)	NA	NA	NA	(+) D (=) O	(+)	(-)	(+) P (=) G
Larynx	(=)	NA	(=)	NA	(=)	(=)	(=)	(+)	(+)	NA	NA
Stomach	NA	NA	NA	NA	NA	(=)	NA	NA	(-)	NA	NA
Liver	NA	NA	NA	NA	(+)	(-)	(=)	NA	NA	NA	(+)
Endometrium	(+)	(+)	NA	NA	NA	NA	NA	(+)	NA	NA	NA
BCC	(=)	(+)	NA	NA	(=)	NA	NA	NA	(+)	NA	NA
Brain	NA	NA	NA	NA	(-)	NA	NA	NA	(+)	NA	NA

+ Increased risk  
 = No effect  
 NA Available data do not allow any conclusion  
 A Asians  
 D Distal colon  
 P Phenotype-based studies  
 (+) Suggested increased risk  
 (-) Suggested reduced risk  
 (=) Suggested no effect  
 BCC Basal cell carcinoma of the skin  
 C Caucasians  
 O Other parts of colon  
 G Genotype-based studies

**Table 3. Viruses associated with malignant disease and potential association with human leukocyte antigen (HLA) type.** (from Little & Stern, 1999; references contained therein).

Virus	Malignant disease risk	Viral-HLA associations; high risk areas	Example of relative risk (RR)
Human T-cell leukemia virus type I (HTLV-I)	Acts as early and critical etiological factor in pathogenesis of adult T-cell leukemia (ATL); maternal or sexual transmission; lifetime cumulative risk of 1–5% for HTLV-I carriers (worldwide) of developing disease <sup>26</sup>	The highest-risk populations are Negroid and Mongoloid <sup>24–26</sup> ; significant increased frequencies for A26 and B61 in Mongoloid ATL patients (Japanese); Negroid patients are more heterogeneous and no significant correlations are seen	Japanese: A26, B61 (RR = 2.1)
Hepatitis B virus (HBV)	Persistent infection with HBV is associated with 85–90% of hepatocellular carcinoma (HCC) with long latency; HBV carriers have 200× the risk for HCC over 30–40 years	No significant HLA associations found within high-risk regions, including China, south east Asia and sub-Saharan Africa; increased HLA-DRB1*0701 and decreased DQB1*02 associated with lack of response to hepatitis B surface antigen in vaccination and chronic persistent infection <sup>27,28</sup> ; HLA-DRB1*1302 associated with protection against persistent HBV infection <sup>29</sup>	No reported associations with HCC
Epstein-Barr virus (EBV)	Burkitt lymphoma	Africa	HLA-DR7 (RR = 3.4) <sup>30</sup>
	Nasopharyngeal carcinoma	South east Asia HLA-B58 and B46 haplotypes are found at increased frequency in affected individuals	B58 (RR = 3.1) and B46 (RR = 2) in southern Chinese <sup>26</sup> ; HLA-A2-positive US Caucasians have a lower risk (RR = 0.46) <sup>31</sup>
	Hodgkin's lymphoma (40% EBV positive)	Worldwide; increase in HLA-DPB1*0301 in Caucasians <sup>31</sup>	DPB1*0301: RR = 1.95 in large multicentre Caucasian study, with RR of 6.19 in French population <sup>31</sup>
Human papilloma virus type 16 (HPV16)	Cervical neoplasia: HPV infection in cervical intraepithelial neoplasia (CIN) has 116× higher risk of developing cancer	Worldwide; HLA associations reviewed in Ref. 34	DRB1*03 (RR range 1.0–7.1 in various studies); DRB1*07 / DQB1*0201 (RR 4.0 or 2.3); DRB1*1502/DQB1*0602 (RR 1.5 or 3.0) <sup>34</sup> in UK or USA-Hispanic studies, respectively

**Table 4. Polymorphic genes, dietary components and cancer: possible candidates.** (from Sinha & Caporaso, 1999).

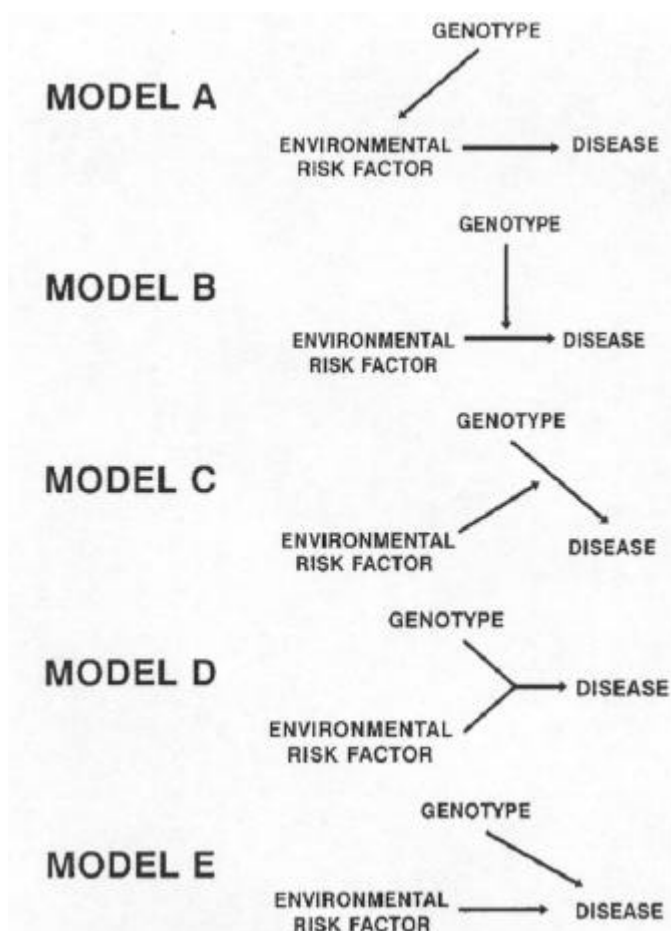
Dietary component	Polymorphic gene/phenotype <sup>1</sup>	Cancer site
<b>Carcinogens</b>		
Heterocyclic amines	NAT2, (NAT1), CYP1A2 (CYP1A1)	Colorectal, breast, other sites
Polycyclic hydrocarbons	CYP1A1, GSTM1	Gastrointestinal tract
Nitrosamines	CYP2E1	Nasopharyngeal, stomach
Aflatoxins	GSTM1, EPHX	Liver
Alcohol	ADH (ALDH, CYP2E1)	Colorectal, oral
<b>Anticarcinogens</b>		
Cruciferous vegetables	CYP1A2, GST	Colorectal, other sites
Fruits and vegetables	CYP1A2, GST	Many sites
Calcium/vitamin D	Vitamin D receptor	Colorectal, prostate
Retinoids	Retinoic acid receptor variant	Acute promyelocytic Leukemia, skin, Head and neck, breast
Folate, methionine	MTHFR, Methionine Synthase	Colorectal, cervix

<sup>1</sup> Abbreviations used: NAT, N-acetyltransferase; CYP, cytochrome p450; GST, glutathione-S-transferase; EPHX, epoxide hydrolase; ADH, alcohol dehydrogenase; MTHFR, methylenetetrahydrofolate reductase.

**Table 5. Categories of environmental-response genes for the first phase of the Environmental Genome Project.** (*from Olden & Wilson, 2000*).

- Xenobiotic metabolism and detoxification
- Cell-surface receptors
- DNA repair
- Cell cycle
- Cell death
- Immune and inflammatory response
- Hormone metabolism
- Nutrition
- Oxidative metabolism and stress
- Membrane pumps and/or drug resistance
- Signal transduction

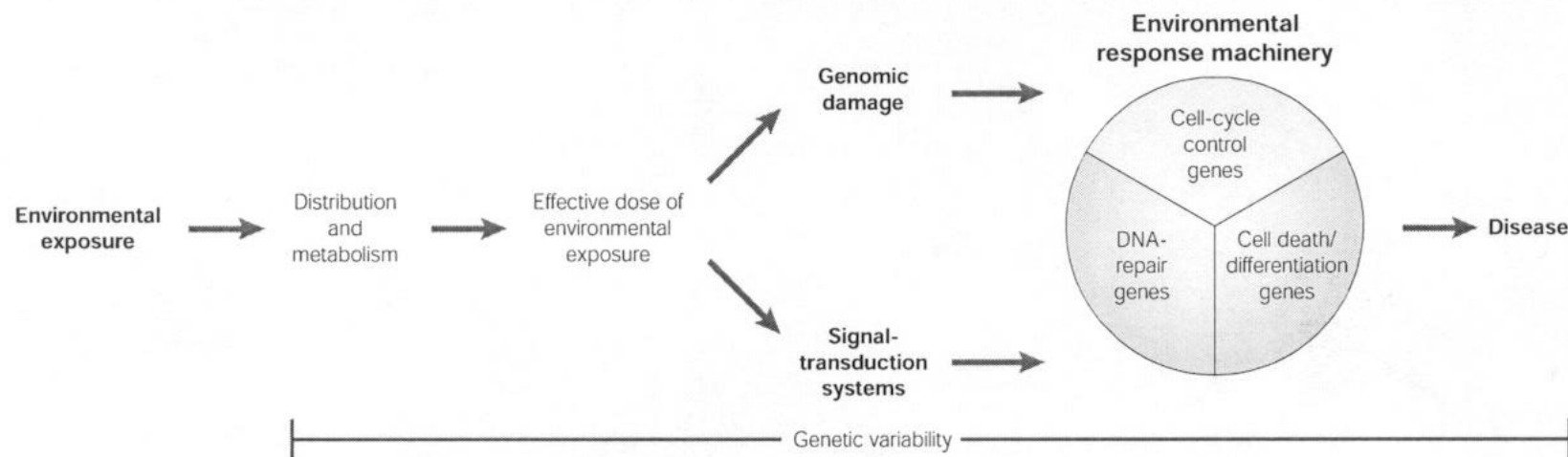
**Figure 1. Five models of the relation between a high-risk genotype and an environmental exposure, in terms of their effect on disease risk. (from Ottman, 1996).**



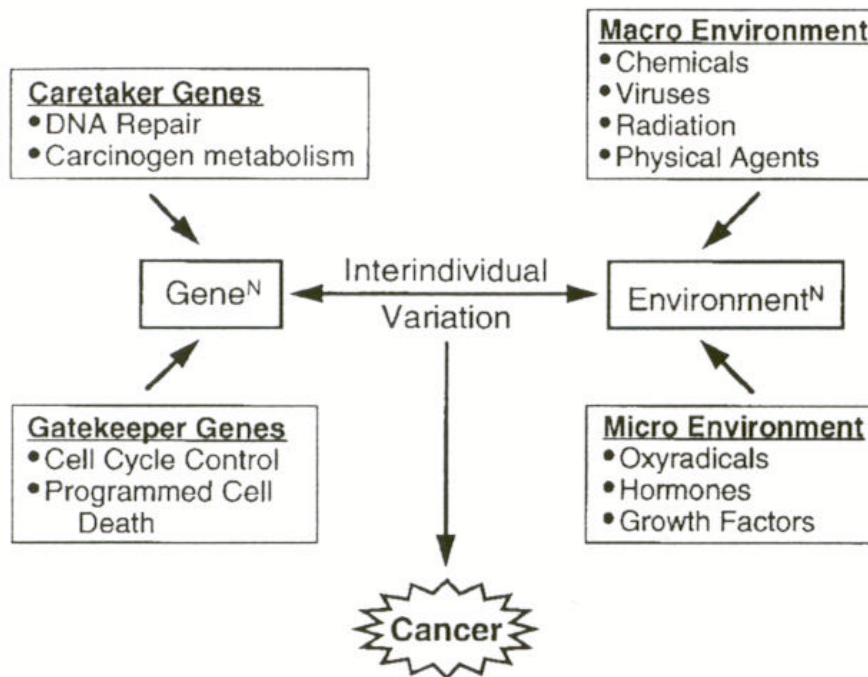
In model A, the effect of the genotype is to produce or increase expression of a “risk factor” that can also be produced non-genetically. This model is not viewed as a gene-environment interaction<sup>2</sup>. In model B, the genotype exacerbates the effect of the environmental risk factor, but there is no effect of the genotype in the unexposed. In model C, the exposure exacerbates the effect of the genotype but there is no effect of exposure in the low-risk genotype. In model D, both exposure and genotype are required to increase risk. In models B - D it should be noted that the presence or absence of the genetic risk factor is irrelevant for disease causation if there is no exposure to an environmental agent. Finally, in model E, exposure and genotype each have some effect on disease risk, and when they occur together, the risk is higher or lower than when they occur alone. Unlike models B-D, an increased risk from the presence of the susceptible genotype is not due solely to gene-environment interaction

<sup>2</sup> A gene-environment interaction is defined as “a different effect of an environmental exposure on disease risk in persons with different genotypes” or “a different effect of a genotype on disease risk in persons with different environmental exposures”  
Ottman 1996.

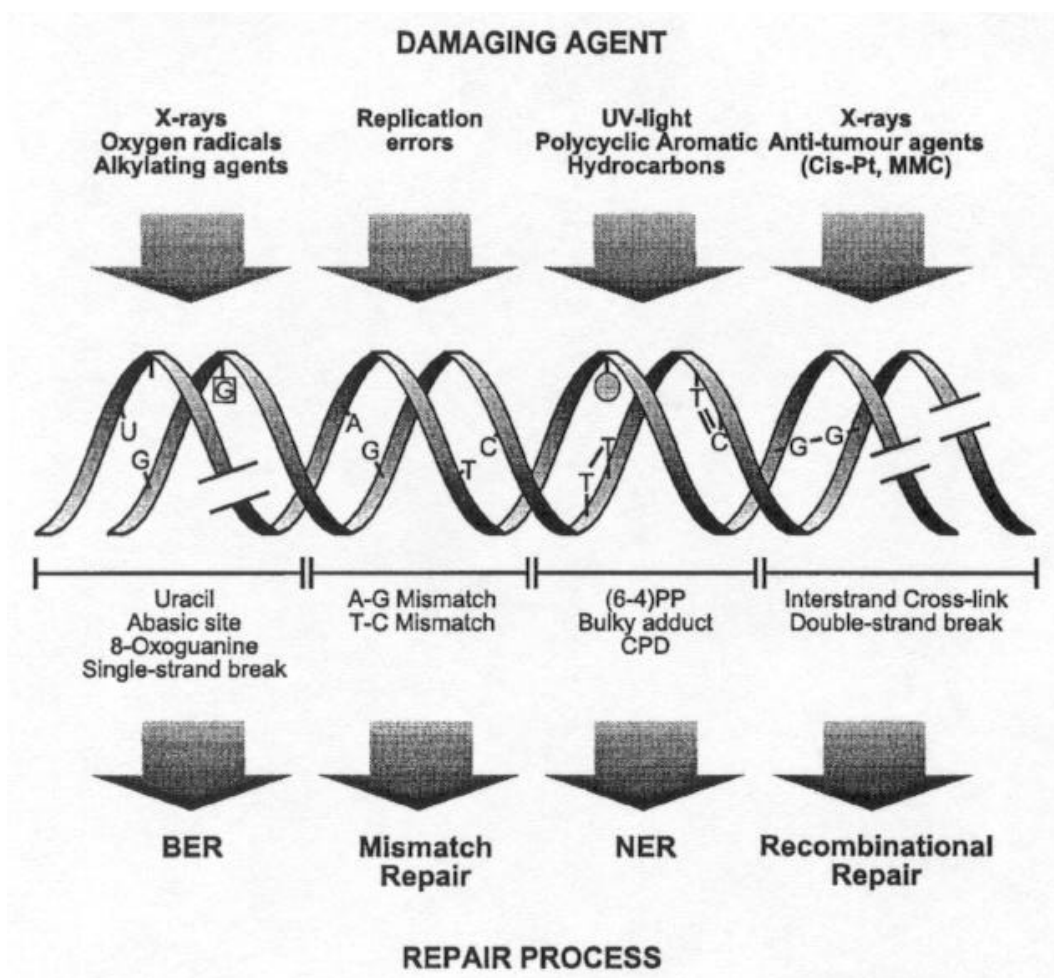
**Figure 2. The environmental exposure-disease model.** Polymorphisms in environmental response genes can modify a person's risk for disease. (from Olden & Wilson, 2000).



**Figure 3.** Low penetrance cancer susceptibility genes as “caretakers” or “gatekeepers”. (from Shields & Harris, 2000).

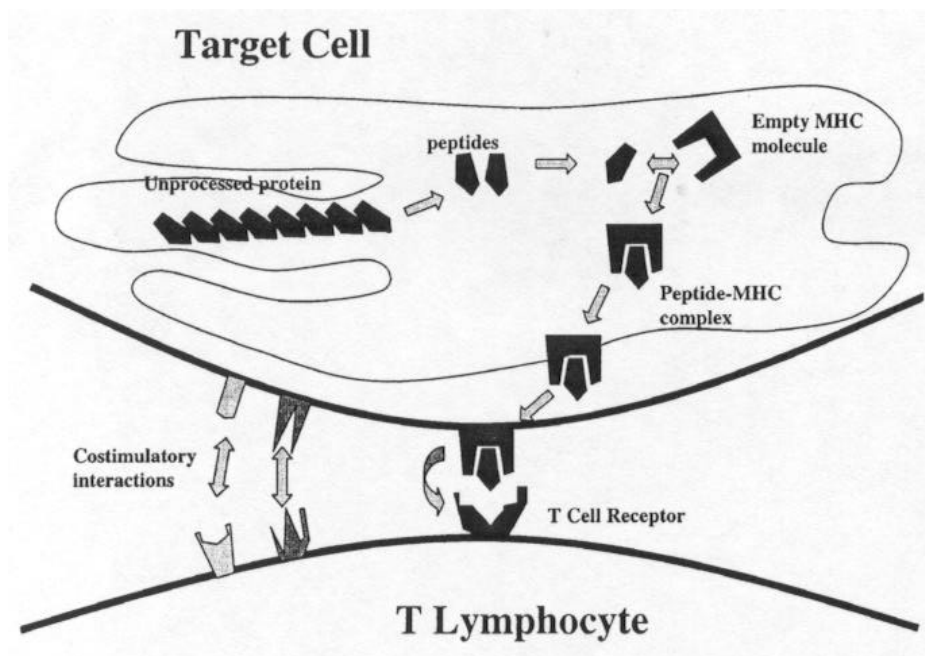


**Figure 4.** DNA lesions and repair mechanisms. (from de Boer & Hoeijmakers, 2000).



(Top) Common DNA damaging agents. (Middle) Examples of lesions that can be introduced by these agents into the DNA double helix. (Bottom) The most frequently used repair mechanisms for such lesions. [BER – Base excision repair, NER – Nucleotide excision repair].

**Figure 5. Endogenous processing of antigen and interactions between the cytotoxic T lymphocyte and target cells.** (from Bremers & Parmiani, 2000).



Within a target or antigen presenting cell (APC) antigenic peptides derived from a tumour cell are proteolytically degraded and transported to the endoplasmic reticulum (processed), where they associate with an MHC (HLA in humans) Class I and  $\beta_2$ -microglobulin chain. This complex is transported to the cell surface, where it can be recognised by the appropriate T-cell receptor. Antigen presentation by HLA Class II molecules may also be important for the induction of an anti-tumour CTL response, activated *via* a  $CD4^+$ -mediated T-helper response. Costimulatory signals (eg, cytokine-mediated signals, specific costimulatory molecules) are also important for adequate activation of T-lymphocytes. In the absence of these signals, T-cell tolerance may occur.