

Cancer Evolution and Individual Susceptibility

Jesús Pérez-Losada ^{1*}, Andrés Castellanos-Martín ¹, and Jian-Hua Mao ^{2*}

¹ Instituto de Biología Molecular y Celular del Cáncer (IBMCC), Instituto Mixto Universidad de Salamanca / CSIC, Campus Miguel de Unamuno s/n, Salamanca, 37007, Spain.

² Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94127, USA.

*Correspondence Author: Jesus Perez-Losada or Jian-Hua Mao, Email:
jperezlosada@usal.es or JHMao@lbl.gov

Summary

Cancer susceptibility is due to interactions between inherited genetic factors and exposure to environmental carcinogens. The genetic component is constituted mainly by weakly acting low-penetrance genetic variants that interact among themselves, as well as with the environment. These low susceptibility genes can be categorized into two main groups: one includes those that control intrinsic tumor cell activities (i.e. apoptosis, proliferation or DNA repair), and the other contains those that modulate the function of extrinsic tumor cell compartments (i.e. stroma, angiogenesis, or endocrine and immune systems). Genome-Wide Association Studies (GWAS) of human populations have identified numerous genetic loci linked with cancer risk and behavior, but nevertheless the major component of cancer heritability remains to be explained. One reason may be that GWAS cannot readily capture gene-gene or gene-environment interactions. Mouse model approaches offer an alternative or complementary strategy, because of our ability to control both the genetic and environmental components of risk. Recently developed genetic tools, including high-throughput technologies such as SNP, CGH and gene expression microarrays, have led to more powerful strategies for refining quantitative trait loci (QTL) and identifying the critical genes. In particular, the cross-species approaches will help to refine locations of QTLs, and reveal their genetic and environmental interactions. The identification of human tumor susceptibility genes and discovery of their roles in carcinogenesis will ultimately be important for the development of methods for prediction of risk, diagnosis, prevention and therapy for human cancers.

Insight, Innovation, Integration

Mouse models are an excellent strategy to identify QTLs in a genetically and environmentally controlled-manner that could be extrapolated to human populations. Additionally, they are very useful to validate and refine candidate loci found in humans by GWAS. The integration of new technical innovations has improved QTL research: the employment of high resolution SNP, CGH and gene expression arrays speeds up the refinement of QTLs. Other technical advances such as whole genome sequencing are readily making the recognition of orthologous regions between both species straight forward, simplifying the refinement of QTLs found in human and mouse, and facilitating a cross-species strategy to identify QTL-driver genes.

Introduction

Cancer is a highly heterogeneous disease, both in terms of the time of individual cancer development, and the biological properties of each tumor. Only a subset of human individuals exposed to the same carcinogen, as for example after radiation exposure,¹ actually developed a tumor, and even among those who were susceptible, tumors did not appear at the same time. Individual tumors also vary enormously in terms of tumor evolution and behavior (i.e. local growth, distant dissemination, treatment response, and relapse or tumor dormancy).²⁻⁵ Thus, individuals that share the same apparent histopathological type of tumor and TNM (Tumor, Node, Metastasis) stage, and also receive the same treatment, could have tumors with completely different evolutionary histories.

One of the most important aspects that remain to be clarified in this field is the identification of the genetic and molecular components that determine the diverse tumor behaviors among different patients who seemingly have the same histopathological disease. It is assumed that most part of the genetic component that contributes to this variability is mainly constituted by the sum of actions of weakly acting low-penetrance genes, whose allelic forms interact among themselves and with the environment to determine the clinical variability among individuals. These genes (also called modifier genes), mostly follow a trend of quantitative inheritance⁶. Additionally, a major part of cancer growth is due to non-cell autonomous processes that consist in an aberrant tissue growing in an uncontrolled manner within the context of the physiology of a complex organism⁷. Therefore, it is a disease that not only depends on the properties of the tumor cells themselves, but also on other compartments like the immune and endocrine systems, stroma, vasculature, and others, all of which play key roles in the development and evolution of cancer. Consequently, differences in tumor behavior are not only determined by intrinsic factors to the tumor cells (such as proliferation, apoptosis, etc), but also by extrinsic factors outside the tumor compartment *per se*. Modifier genes could regulate both the molecular and cellular functions of these different compartments, and this fact could explain the differences in tumor behavior among patients who seemingly suffer the same disease. The identification of those modifier genes is one of the major challenges of the future in cancer research. In this review we use mainly breast cancer as paradigm to illustrate this issue.

Genetics analysis strategies are the only tools that allow us to consider the global scenario, because the susceptibility loci can contain risk genes controlling either intrinsic or extrinsic factors^{6,8}. Genome Wide Association Studies (GWAS) have permitted the identification of different susceptibility regions, genes and pathways, but

the reproducibility between studies is difficult, probably due to the high heterogeneity of human genetics and its complex interaction with environmental factors^{6, 8}. Crosses among inbred mouse strains of defined genetic backgrounds that exhibit strong differences in cancer susceptibility, have simplified the identification of Quantitative Trait Loci (QTL) and their interactions responsible for variable tumor behavior. Although the refinement of the QTLs to the gene level using mouse models was a very difficult task, the new high-throughput technologies recently developed in Genetics, Genomics and Bioinformatics help to tackle this complicate duty with success. All these new technologies are improving our understanding of the genetic component/networks that control the variability in tumor risk, development and clinical evolution. The final goal is to obtain a better understanding of the molecular factors that determine the variability of the disease, which will finally result in the development of more personalized clinical applications for the benefit of the patient.

Cancer Heritability

Cancer heritability is still nowadays very poorly understood. No more than a modest portion of cancers present an obvious trend of heredity (the considered “real” hereditary cancers). This is actually the situation of breast cancer where just a small portion of tumors could be identified by the inheritance of mutated variants of high penetrance genes like *BRCA1* and *BRCA2*. However, mutations of these genes only account for a small percentage of the human tumor predisposition, resulting in quite rare hereditary cancer syndromes. These kinds of uncommon, but severe alleles have been additionally implicated in most forms of hereditary cancer syndromes. Hereditary susceptibility to breast cancer has become connected with germline mutations in at least eighteen genes.⁹ A huge number of distinct loss-of-function mutations have been discovered within *BRCA1* as well as *BRCA2* genes; most of these variations are usually individually rare, and each one confers quite high susceptibility for breast and ovarian malignancies. Uncommon germline mutations of other genes are also connected with elevated risk of breast cancer, ranging from two-fold for *CHEK2* to ten-fold for *P53*. Interestingly, all of these genes function in networks that are crucial to DNA repair and preservation of genomic integrity. In most cases, the inherited mutation is followed by somatic loss of the corresponding wild type allele, resulting in the mistakes in DNA repair that finally lead to cancer development.¹⁰ However, it is important to emphasize that only a low percentage of global cancer risk can be attributed to hereditary mutations in the high-penetrance care-taker genes, and present obvious patterns of Mendelian inheritance.

The environmental component may play a more essential role in sporadic tumors than in hereditary cancer, and in the some cases may over-ride the genetics.¹¹⁻
¹³ Nevertheless, there is significant evidence from large scale epidemiological studies indicating that the chance of developing sporadic cancer also has a significant hereditary component. One of these studies, which was carried out on several thousand pairs of twins, demonstrated that when one twin developed cancer, the other had a significantly increased risk of generating the same type of disease, but without any obvious Mendelian inheritance pattern. In fact, many scientific studies have concluded that common cancers are polygenic diseases with a quantitative genetic pattern.¹¹⁻¹⁴

Even in families carrying specific mutant alleles of high-penetrance risk genes with potent effects such as *BRCA1* and *BRCA2*, phenotypes tend to be influenced by the hereditary background, becoming much more comparable between affected twins, but varying among more distant family members with the same gene alteration;^{6,14,15} this would indicate that low-penetrance genes could also modify the behavior of hereditary cancer. A deeper knowledge of the genetic component would be essential to estimate the individual genetic susceptibility to develop cancer, to improve early detection and diagnosis of the disease, and to understand the fundamental biochemical and physiological pathways governed by those low-susceptibility genes as a critical step for the development of new cancer treatments.¹⁶

Cancer has a Non-Cell Autonomous Disease Component

Cancer is in part a non-cell autonomous process; it is an aberrant tissue that grows in an uncontrolled manner in the context of the physiology and pathophysiology of a complex organism. Tumor growth, as that of any other tissue, depends not only on the intrinsic properties of the parenchymal component (tumor cells), but also on other organism compartments such as the immune and endocrine systems, stroma, vascular system, etcetera (Figure 1). All of them play key roles in the development and evolution of cancer. Consequently, tumor behaviour (i.e. susceptibility, development and clinical evolution) is not only going to be determined by factors intrinsic to the tumor cell, involved in processes such as proliferation, apoptosis, DNA repair etc; but it will also be influenced by those extrinsic factors from other compartments. Furthermore, these two main compartments are not independent, but rather they continuously crosstalk and interact with each other, so that the intrinsic factors are capable of recruiting the extrinsic ones, and the availability of the extrinsic factors determines the intrinsic cellular activity.

The fact is that connections among cancer cells, the stroma and the rest of the organism have their roots in the physiological responses that take part in regular tissue homeostasis.¹⁷⁻¹⁹ The equilibrium between cell-renewal and cell-reduction is tightly governed through connections between parenchyma stem cells and the microenvironment to carry out the tissue remodelling or respond to the stress caused by tissue injury. Cancer cells virtually do not react in response to normal physiological regulators of cell growth, and are constantly sending remodelling signals for the stroma to be reorganized in an activated form to permit tumour growth; to some degree tumors behave like a wound that does not heal.²⁰ Tumor modifier genes could play a role in controlling molecular and cellular factors of these two main compartments that would explain not only discrete physiological differences among individuals, but also differences in the susceptibility, development and the different clinical evolution among patients who seemingly suffer the same cancer disease.

The relevance of the microenvironment is highlighted by new studies that demonstrate how the apparently normal stromal cells can manipulate epithelial cancer cell activity in reconstitution experiments, and by recognition of particular somatic genetic alterations in the stromal element of the tumor.^{21,22} It has been proposed that the global microenvironment mostly functions as an epigenetic tumor modifier.²³ In fact, the genetic inactivation of *Pten* in stromal fibroblasts associated with mouse mammary glands speeds up the initiation, progression and malignant transformation of mammary epithelial tumors.²⁴ Furthermore, malignant cells can be reverted to a quiescent state simply by incorporation into an embryonic microenvironment.²⁵ This suggests the microenvironment is dominating over malignancy. Thus, for tumors to advance into a more malignant stage they must change their own microenvironment to a promoting one. The change in microenvironment probably originates from oncogenic mutations in proliferating tumor cells that send signals to the stroma, but possibly also mutations in the stroma itself.²¹⁻²⁴

The resulting tumors are complicated structures of malignant cancer cells surrounded by vasculature and associated with an active tumor stroma composed of several non-malignant cell types, such as fibroblasts and myeloid cells with an important role in global tumor behaviour. For example, evidence suggests that tumor initiation, progression, as well as metastasis are influenced by particular subpopulations of macrophages,^{26,27} and also other inflammatory cells, such as B and T- lymphocytes and mastocytes, have been shown to play a role in tumor promotion.²⁸ In fact, the milieu of the tumor microenvironment is similar to the one found in the inflammatory reactions within a restorative healing injury, which stimulates

angiogenesis, turnover of the extracellular matrix (ECM), as well as tumor cell motility.⁷ And, as occurs in inflammation, a growing body of data support the perspective that extracellular proteinases, like the matrix metalloproteinases (MMPs), mediate numerous modifications within the microenvironment in the course of tumor development.²⁹ Additionally, one of the most critical pathways controlling both inflammation and tumor microenvironment, is the TGF-beta signalling pathway, together with important cell-autonomous effects. GWAS have identified many SNPs close to genes that belong to the *TGF-beta* superfamily, such as *CREM1* and *SMAD7*. In addition, constitutively reduced *TGFBR1* expression is a powerful modifier of colorectal cancer susceptibility. All these data indicate that germline variations of the TGF-beta superfamily might account for a very important percentage of colorectal cancer susceptibility.³⁰

The modulation of stroma function by tumor susceptibility modifier genes is well-known. The first tumor-modifier gene identified was a modifier of the *Apc* (Adenomatous Polyposis Coli) gene function, located in the QTL named *Mom1* ("Modifier of Min1", which in turn means "Multiple Intestinal Neoplasia-1"). The gene responsible encodes a secretory phosphatase type II phospholipase A (*Pla2s*). *Pla2s* was proposed to modify polyp number by altering the cellular microenvironment within the intestinal crypt.³¹ Interestingly, this gene has been widely implicated in the inflammatory process,³² angiogenesis and has pro-atherogenic activity.³³ More importantly, later studies demonstrated that the *PLA2S* gene has a role in human cancer pathogenesis of the digestive tract,³⁴ supporting the importance of mouse tools to identify cancer modifier genes in human population. Thus, it is feasible for a number of these genetic determinants to be involved in the pathogenesis of different physiological and/or pathophysiological events at the same time; this effect is named "pleiotropy". This concept refers to those genes that concurrently have effects on different phenotypes. This has been demonstrated not only for the diverse subtypes of cancer, such as 8q12 abnormalities that are related with various types of tumors,³⁵ but also for autoimmune diseases³⁶ or very different pathologic conditions. Many parallels exist between different diseases and pathologic situations; for example, hypertension, hypercholesterolemia and obesity are included in known metabolic syndrome; or the existence of an association between those processes and particular types of cancer; or the relationship between certain autoimmune diseases and cancer.³⁷⁻³⁹ All of these data indicate that complex interactions take place among genes that simultaneously control different processes. It is also possible to relate all the disorders that share common conditions and the gene interactions that control them. This fact has

generated recently the interesting concept of diseasome,⁴⁰ which can be represented by two networks, first, by the human disease network, in which nodes represent disorders, and two disorders are connected to each other if they share at least one gene; and second, by a disease gene network where nodes represent disease genes, and two genes are connected if they are associated with the same disorder.⁴⁰

One of the most important challenges in cancer research is to understand the underlying basis of heterogeneity of tumor susceptibility, development, and evolution in the context of the physiology and pathophysiology of the organism. It would be desirable to tackle the cancer problem with tools that permit visualization of this global picture, integrating both intrinsic and extrinsic factors with the behaviour of the tumor cell. Genetic analysis offers a unique tool to embrace the global scenario, because each QTL regions could contain both intrinsic and/or extrinsic modifier genes, and can help to explain cancer as a complex disease.

Identification of Cancer Susceptibility Genes in Human Populations

Even though rare alleles with strong effects could be substantial contributors to sporadic cancer risk, the searching for tumor susceptibility has been mainly focused on the common disease-common variant model that presumes that cancer susceptibility originates from the additive effects of combinations of common low-penetrance variants.⁴¹ With this model, every susceptibility variant is assumed to contribute a small amount of risk, without any variant conferring enough by itself to result in tumor development. Cancer origin and evolution have been proposed to be the consequence of the merged effects of numerous of such alleles, which may control intrinsic and/or extrinsic functions. The search for tumor susceptibility genes has mainly been carried out by GWAS, in which allele frequencies at thousands of polymorphic sites (i.e. SNPs) are compared in a large number of cases versus a similar number of controls. As discussed in later sections, in spite of their limitations, these studies have successfully identified some of the common susceptibility variants for different common diseases and traits, including cancer¹⁰.

A. Identification of Susceptibility Genes in Breast Cancer

We will use breast cancer as a model for this discussion, as major efforts have been made to identify genetic components of both hereditary and sporadic versions of the disease. Studies of susceptibility genes in breast cancer initially focused on the detection of high-penetrance susceptibility genes through the analysis of linkage in family pedigrees comprising several affected members. These familial studies involve fewer patients, and need significantly reduced marker density, in comparison with

current GWAS, but the two approaches can be complementary. The results of pedigree evaluation can offer important and persuasive signs of genetic effects, because they are primarily based on genetic transmission of disease-causing alleles between affected family members. Inherited variations in the two main susceptibility genes already known for breast cancer, *BRCA1* and *BRCA2*, together account for only around 20% of hereditary breast cancer.³⁵ A few additional genes have been identified,^{42,43} but all these known mutations only can elucidate a small portion of familial breast cancers, and around less than 5% of the total breast cancer susceptibility. A number of linkage studies have described candidate loci that contain breast cancer susceptibility genes. However, these loci were not clearly statistically significant presumably due to the fact that the number of families affected by each locus was low. In a recent linkage study in Spanish breast cancer families, three more regions of interest came out, located on 3q25, 6q24 and 21q22;⁴³ it will be very important to further confirm these results in new populations.

The majority of the studies to identify susceptibility genes in breast cancer have been carried out by GWAS. These studies have identified several common variants that have an influence on breast cancer susceptibility, but only four were replicated in two or more GWAS (Table 1).⁴⁴⁻⁵¹ Meta-analysis of suggestive loci utilizing three published GWAS resulted in the detection of an extra locus on 5p12⁵² that appeared to be linked particularly with estrogen-receptor positive cancers of the breast. It must be taken into account that statements for associations with particular categories require much more cautious replication studies. For instance, the 2q35 locus was initially associated particularly with estrogen-receptor positive breast cancer, but a later study reported comparable results irrespective of the estrogen receptor status⁵³

Almost all of the individual low-penetrance variations discovered to date have weak effects (odds ratios per-allele are less than 1.41) and explain much less of the heritability of breast cancer, compared to the known *BRCA1* and *BRCA2* mutations; and maybe a few others such as a common variant within the *Fibroblast Growth Factor Receptor 2* gene (*FGFR2*). Nonetheless, a very significant portion of the breast cancer susceptibility presently continues to be uncharacterized and may be due to the sum of combinations or interactions of low-penetrance genes. These allele variants, together with the environmental exposure, may contribute to breast cancer susceptibility. The causative environmental exposures continue to be evasive, because many of the formerly suggestive environmental and life style risk factors (e.g. nutrition) for breast cancer have been recently refuted by large studies in the last decade.^{54,55}

In conclusion, around twenty different presumed breast cancer susceptibility loci have already been identified using GWAS studies, but few loci were replicated in

different studies.⁴⁴ In addition, almost all of these variants identified in breast cancer and other studies, have no demonstrated biological or mechanistic relevance to the disease, or medical utility for diagnosis or therapy. This could mean that causality within this framework can hardly ever be solved by large-scale association or case-control studies exclusively.¹⁰ A reason for this could be the genetic heterogeneity. A number of genes is presumed to play an important role in the susceptibility to breast cancer; but those genes could be only important exclusively within a limited amount of families, and could be absolutely lost as soon as they are diluted in the general population.¹⁶

B. The Challenging of Genome-Wide Association Studies

Within the last few years, over fifty GWAS have been performed to search for cancer susceptibility genes. As discussed by other authors, a few repetitive conclusions can be obtained from them: first, only few variants were found in every single GWAS; second, each locus has a tiny effect; and third, there is a relative deficit of replication of allele variants identified by diverse GWAS.¹⁶ One explanation for this last problem could be that, despite having large sample numbers, there is a restricted potential to identify modest genetic effects due to the strict levels of significance demanded in these studies. Therefore, variants that attain significant p-values, for instance $P > 10^{-8}$, are usually real,^{56,57} while those associated with significantly more modest P-values (e.g. 10^{-5} or 10^{-6}) might indicate false positives. Furthermore, a large number of those variants will not be replicated when screened in other samples. For instance, in GWAS the chance that a variant with a P-value of 10^{-5} shows a genuine association is actually lower than one percent.^{58,59}

A different situation that could play a role in clarifying the current incongruence of GWAS is the fact that cancer susceptibility is an extremely complicated phenotype and, together with the incomplete penetrance of the inherited tumor risk alleles, the interaction with environmental risk factors could substantially alter hereditary susceptibility. Based on environmental exposures, a person with high genetic susceptibility to develop malignancy may well never be affected, while a person at low cancer risk, but high exposure, might suffer the affliction⁶⁰ (Figure-2). This question is still widely debated for breast cancer, which has been linked to nutritional⁶¹ as well as reproductive factors,⁶² and alcohol,⁶³ along with other exposures. Inability to take into account this kind of variable in GWAS may well decrease the strength of analysis or even reduce our ability to discover genuine causative susceptibility loci.⁶⁴ This may also clarify the fact that, even though the hereditary element of developing prostate cancer was estimated to be around 40%¹¹, early studies including quite large high risk

families have not been confirmed these data.⁶⁵ This might be due to genetic heterogeneity (i.e. the causal polymorphisms that are the responsible for the phenotype vary among families).

Additionally, even though quite a few studies have discovered genes with important phenotypic effects,^{44,47,68-69} as we indicated above, the majority of cancer risk is most likely due to a selection of genes with additive effects. It is however also possible that many genes regulate cancer susceptibility mainly via non-additive interactions. These interactions could be multiplicative or conditional, in such a way that the principal effect of one gene would depend on the existence of a specific allele within a second locus, and so on, forming a network of gene interactions where the next interaction is only possible only if particular allelic form is already present. It might be also possible that those numerous weak-interacting loci would simply achieve suitable levels of significance within particular series of patient samples, depending on the hereditary background or environmental factors (Figure 2). Thus, on top of this currently complicated situation of the GWAS scenario would be the spectrum of hereditary interactions that depend on the genetic background, a fact that has been clearly demonstrated in animal models like the mouse and others.^{6,70-72} Moreover, genome-wide studies in mouse models of cancer have discovered loci that arise as a result of genetic interactions that are not viewed as individual QTL with major effects, utilizing common methods of analysis.⁷⁰⁻⁷³ These studies demonstrated, first, the power of mouse models to simplify the problem, and second, that more advanced statistical methods used to discover interactions among loci in linkage analyses might be required to discover the locations of multiple weak susceptibility alleles.^{6,74,75}

In summary, extrapolation of the final results obtained from GWAS to other human populations raises the uncomfortable possibility that a specific SNP discovered as a tumor modifier in a particular population, however lacks any effect (or even might work in the opposite direction) within another ethnic background. For that reason, even though hereditary background in individual patients is consequently capable of controlling illness development, as it has been evidently demonstrated in animal models^{76,77} very few of these human low susceptibility genes have been convincingly identified. Therefore, even though the present flood of GWAS show the strength of this strategy, there are natural restrictions of this whole-genome association analysis that circumvent the capture of most pertinent scientific data. Thus, GWAS are afflicted by implicit limitations and cannot provide us with an entire understanding of the intricate genetic and environmental interactions connected with common disease phenotypes. In fact, today no individual method is good enough to permit an extensive knowledge of cancer etiology and pathogenesis, in particular within the extremely complicated area

of human genetics. But there have been great improvements in QTL research within the last ten years, mainly by utilizing mouse cancer models. Mouse QTL analysis as well as GWAS will become complementary strategies that will improve the knowledge of the actual genetic basis of the human disease.⁸

The Mouse as a Complementary Approach

The use of mouse models can complement human GWAS by allowing a higher level of control over hereditary variance and environmental exposure. An intercross or a backcross carried out between two inbred mouse strains with divergent tumor phenotypes generates offspring where each mouse is genetically and phenotypically unique for different quantitative sub-phenotypes controlled by different interacting QTL. These strategies reproduce a simplified model of cancer as a polygenic complex disease.⁸ Furthermore, these approaches are facilitated by the production of a large number of inbred and outbred strains of various *Mus* species which have different evolutionary genealogies, together with recombinant inbred strains, congenics,⁶⁹ consomics,⁷⁸ and genetically engineered mice (GEM), all of them constituting a unique genetic resource among animal models that can greatly simplify the identification of susceptibility genes. Certainly, the enormous number of GEMs available, in particular through programs like the Knock-Out Mouse Project (KOMP) whose goal is to mutate all protein-encoding genes and make all these mice available to the scientific community⁷⁹ provide important tools to narrow down QTL candidate genes. Knockouts are used in this context to test the candidature of a driver gene at a QTL by what it has been named the QTL-knockout interaction test, by which the interaction between the null allele and the QTL is tested, and compared with the interaction with the wild type allele⁸⁰. The use of a GEM strain carrying a knock-out or a knock-in allele located in a QTL, can help to validate the participation of that gene in the QTL effect by linkage analysis⁸¹.

Additionally, there is increasing evidence showing that hereditary risk factors have a comparable role in complex disease pathogenesis within human and mouse models regardless of interspecies dissimilarities. Rodents develop cancer that appears to be amazingly similar in most cases to human tumors, and they accumulate mutations in a comparable spectrum involving the same genes and pathways.⁸² These facts suggest that, at least some of the numerous QTLs containing tumor risk genes that have been mapped in the mouse may be highly relevant to the human scenario and serve as an effective method of complementing observations within human populations.⁶ This has been demonstrated for example, for plasma levels of cholesterol.⁸³ Mouse QTLs have already been proven to be equivalent to human

disease susceptibility loci in a number of cases including cancer,⁸⁴⁻⁸⁷ and although mouse QTL analysis is not without drawbacks, there have been in the last few years important technical advances (see below) (Figure 3). Certainly, the introduction of novel techniques and resources can help to unravel the exceptionally complicated interwoven factors influencing cancer etiology. These data suggest that mouse studies, carried out in parallel with human sample analysis, may accelerate development of a deeper understanding of the hereditary risk component of complex diseases⁸ (Figure 4).

Technological Advances for Identification of QTLs in Mouse Models

Although mouse quantitative trait locus mapping has demonstrated to be an effective tool to identify trait risk, this strategy is not without drawbacks.⁸ These have prompted numerous researchers to examine substitute approaches for candidate gene detection.⁷⁷ Among the drawbacks are low mapping resolution, and secondly, the difficulty of identifying specific genes and nucleotides associated with complex genetic traits; thirdly, modelling multiple QTL, although easier than in GWAS carried out in human populations, requires complex statistical strategies. Technological improvements in the meantime have partially resolved several of these issues (Figure-3), including:

-Analysis of haplotype structure and in silico mapping. Understanding of those pieces of the genome that are the same by ancestry (i.e., have the same haplotype organization) among mouse strains is a useful approach to refining loci of interest. The strategy utilizes genetically more complex mice from natural origin such as partially inbred and outbred strains like *Mus spretus*⁸⁶ or heterogeneous stocks generated by combinations of various inbred mouse strains. In both cases this strategy may quickly minimize the number of candidate genes that have to be tested.^{86,88,89} The main idea is to identify frequent haplotype sections that segregate inside the genetically characterized candidate locus to restrict the quest for presumed genes of interest.⁸⁸⁻⁹¹

Heterogeneous stock mice like MF1⁹² have been developed through the arbitrary reproduction of progeny from normally four to eight inbred strains.^{93,94} QTL should be found in a region in which sequence divergence matches genetic origin. Therefore, whenever QTLs have been mapped in heterogeneous stock populations, the markers of strain distribution structure within the initial QTL can be joined with mapping data to refine the area which contains the functional variant.

This strategy has been utilized to carry out mapping at an exceptionally high resolution and to identify candidate genes,^{88-90,95} In summary, due to the fact that heterogeneous stocks are produced from known ancestral inbred strains, it is possible, by some Statistical Genetics and *in silico* analysis, to obtain the origin of every allele and to map QTL at sub-centimorgan resolution.

Similarly to the use of heterogeneous stocks of mice artificially generated, is the utilization of natural outbred stocks which accumulate recombinants with time, so they provide a substantial increase in mapping resolution, possibly sufficient enough to identify candidate genes. Specifically, outbred *Mus spretus* have been utilized with successful results to discover Aurora Kinase A (Aurka/Stk6)^{96,97} as a skin tumor susceptibility gene;⁸⁶ and an outbred population of CD1 mice has been utilized to chart a predisposing region for lung cancer.⁹⁸ These kinds of natural outbred stocks may well provide greater resolution than artificial versions; however they miss the benefits from parental information within the heterogeneous stocks. In addition, these approaches require many animals as well as high density genotyping; in fact, genome-wide mapping in heterogeneous stocks demands a minimum of 6000 genetic markers (i.e. SNPs). It is very important to consider that to reduce false-positive results to appropriate levels with such amount of markers it is necessary to utilize strict significance thresholds for the p-value.^{99,100}

Recently the Collaborative Cross project was launched to generate the largest panel of recombinant inbred (RI) strains with more than one thousand RI lines of mice. These strains originated from the crosses among five inbred and three wild-derived strains. This strategy will allow high resolution mapping equivalent to the heterogeneous stocks, together with the reproducibility of the inbred strains. The main aim is to reach a mapping resolution of about a megabase. Additionally, the genetic variation will be homogeneously distributed along all the genome without regions where there is no variation, so every single gene can potentially be tested for involvement in a particular phenotype.^{101,102}

-Analysis of *tumors using whole genome array comparative genomic hybridization (aCGH) and loss of heterozygosity (LOH) analysis by SNP arrays*: High penetrance germline susceptibility genes, are often linked to somatic loss of the wild type allele in tumors (the "two hit" Knudson hypothesis). The same could happen with at least some of the low-penetrance susceptibility genes that control intrinsic cellular activities. Cancer low-susceptibility genes could drive copy number gains in tumors in an allele-specific manner, while cancer resistance alleles may possibly be lost as a result of deletion or mitotic recombination leading to loss of heterozygosity (LOH). These types of allele-specific somatic losses and gains can be used to identify cancer

risk genes;^{87,103} this technique facilitated the recognition of Stk6 as a low-penetrance tumor susceptibility gene.⁸⁶

-Genome-wide expression arrays. The observation that most part of SNPs are located outside coding regions has led numerous investigators to hypothesize that many QTLs are probably attributed to delicate alterations in gene expression instead of to missense or nonsense mutations, as is the case for Kras2 in cancer induced by urethane.⁸¹ This idea has consequently resulted in the screening for genes in QTL regions that exhibit differential expression regarding the strains of interest.¹⁰⁴ This particular strategy, initially specified as genetical-genomics,¹⁰⁵ offered a good impartial method for quickly screening hundreds of possible candidate genes at the same time to reduce the list for additional evaluation to a workable quantity.^{90,106} Within this technology researchers could include *co-regulated networks of expression and QTL evaluation*: this allows identification of a group of genes that are operating collectively to impact a susceptibility phenotype based on the network of genes that are significantly correlated with each other, and their expression levels controlled by common genetic loci. In some informative circumstances, it would be possible to find the susceptibility locus, the candidate gene is affected in *cis* by that locus, and downstream genes which are influenced in *trans*. Thus, adding automatic finding and manual curation it is possible to define networks of genes with a common function and that are controlled by a common mechanism.^{72, 74, 75}

-The next generation of sequencing techniques together with the culmination of the human^{107,108} as well as mouse¹⁰⁹ genome sequencing projects: Thanks to the completion of the human genome project it is possible to identify most genes within a specified location. Next-generation sequencing will make it possible to investigate particular candidate genes without prior genomic screening. The power to discover and define candidate loci has continued to grow considerably since the whole genomes of many species have been sequenced.^{110,111} This has allowed recognition of evolutionary conserved sequence domains, and much more recently has allowed direct visualizations of SNPs among some of the widely used inbred mouse strains through the use of chip-based sequencing, as well as large-scale polymorphism screening.¹¹² Interestingly, sequence accessibility throughout species has allowed additional speeding of candidate gene recognition for all those traits which have already been mapped in several species. Recognition of orthologous chromosomal sections and their breakpoints inside genetically identified loci might help to refine QTL localization and candidate gene databases by restricting searches to those regions shared between the two species.^{8,106}

Future Perspectives for Human Cancer Risk

As we have discussed, the majority of studies in human families with higher cancer susceptibility continue to be centred on the chance that a single or even a small number of powerful genes could be the cause of the "missing" hereditary element of tumor risk. Even though there may well still be a number of high penetrance genes remaining to be discovered, mostly, we will be confronted with the difficulties of the existence of numerous low-penetrance modifier genes and their interconnections, as it has already been suggested using mouse models of cancer that combinations of these low-penetrance modifier genes may be responsible for the variable risk of both hereditary and sporadic cancer.⁶ In the last years, the efforts to identify this cancer genetic component in the human population have been mainly focused on the use of GWAS. Although these studies have been proven to be a very useful tool for the identification of some common genetic variants, how much this technique has contributed to clarification of the "missing heritability" of different complex diseases is a matter of controversy.^{10,113,114} There have been substantial attempts to recognize low penetrance cancer susceptibility genes by GWAS^{44,115,116} Even though this research has found a few allelic variants which influence cancer risk, the majority of them will probably be challenged by this method, because of the tiny impact that any single one confers on the total tumor risk. These low penetrance-genes are subject to strong interactions among themselves as well as with the natural environment, and the results can be quite inconsistent within different populations under the influence of diverse environmental elements. Therefore, the actual identification of low penetrance cancer risk genes within the human population is a challenging endeavour because of the huge heterogeneity within human genetics and the environment. This could explain why the majority of the heritable portion of tumor and complex traits has not yet been identified by GWAS. For the same reasons, even though some of the genetic loci discovered through GWAS initially possess robust statistical significance regarding association with specific tumors, the informative potential of these loci to predict individual tumor susceptibility is restricted by their small impact on global cancer risk, so the clinical importance of this kind of variant will be very limited. Therefore, with our current information, we can say that single SNPs will have limited utility in predicting if someone will suffer from cancer. But, although the diagnostic benefit of any genetic polymorphism alone is limited, we can anticipate that understanding of the combined interactions among those allelic variants that collectively possess considerably more potent consequences on risk would likely exert a significantly larger effect on the prognosis as well as on cancer therapy and its general clinical management.

Concluding Remarks

Even though mouse models have become an invaluable tool for QTL mapping, a refining of QTL locations remains problematic. This task is beginning to be tackled successfully with the help of newly developed technologies such as high throughput gene expression arrays together with systems genetics approaches,⁷⁴ whole genome SNP arrays and aCGH with allele specific analysis,⁸⁶ and high-coverage whole genome sequencing that will probably become the technique of reference as soon as it results cost-effective. In the following years, as a result of the application of these technical innovations we ought to start to see the refinement of several loci containing mouse cancer risk alleles and also the identification of clusters of them, jointly with their interactions, that may help selecting presumed genes and pathways to become analyzed in human populations.⁶ Moreover, considering the current speed of technical advancement, it is quite possible that in the near future, with the advent of new technologies such as whole genome sequencing, positional cloning may be unnecessary and fine mapping of significant loci may lead straight to their identification.

Mouse models not only are a good tool to identify QTL regions that can be extrapolated to human populations, but also offer a parallel system for immediate testing and verification of the results obtained from human epidemiology and GWAS. Also, moving back and forth between mouse and human systems will be a good strategy to recognize the causal genetic variant of a presumed candidate gene (Figure-4). Moreover, it is known that the environmental influences and way of life options have an important effect on tumor susceptibility in the humans. Gene-Environment interactions could also be investigated using mouse models, and will allow us to recognize how genes work together with particular environmental influences recognized by epidemiological studies. Enrolling together systems genetics and epidemiology ought to enable us to clarify the connections involving hereditary background and environmental factors that are the reason for part of the "obscure" cancer heritability.⁷⁵

The knowledge acquired by means of these genetics studies will have a significant effect on medical sciences, and should certainly lead to improved prognosis prediction and therapy of human cancer, leading to a more individualized clinical management of the disease.

Acknowledgments

We apologize for these contributions that could not be cited in the manuscript due to space restrictions. We thank Drs. Allan Balmain, Isidro Sánchez Garcia and Cesar Cobaleda for useful comments. J.H.Mao is supported by Office of Biological & Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, by Laboratory Directed Research & Development Program (LDRD), and by the National Institutes of Health, National Cancer Institute grant R01 CA116481. J. Pérez-Losada is partially supported by FEDER and MICINN (PLE2009-119), FIS (PI070057; PI10/00328), CSIC (2009201137), Junta de Castilla y León (SAN126/SA66/09 ;SA079A09). A. Castellanos-Martín is supported by FEDER and MICINN (PLE2009-119).

References

1. A. Carmichael, A. S. Sami, J. M. Dixon, Breast cancer risk among the survivors of atomic bomb and patients exposed to therapeutic ionising radiation. *Eur J Surg Oncol* 2003, *29*. 475-9.
2. J. A. Aguirre-Ghiso, Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer* 2007, *7*. 834-46.
3. D. J. Slamon, B. Leyland-Jones, S. Shak, H. Fuchs, V. Paton, A. Bajamonde, T. Fleming, W. Eiermann, J. Wolter, M. Pegram, J. Baselga, L. Norton, Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 2001, *344*. 783-92.
4. C. Smigal, A. Jemal, E. Ward, V. Cokkinides, R. Smith, H. L. Howe, M. Thun, Trends in breast cancer by race and ethnicity: update 2006. *CA Cancer J Clin* 2006, *56*. 168-83.
5. S. F. Winter, K. W. Hunter, Mouse modifier genes in mammary tumorigenesis and metastasis. *J Mammary Gland Biol Neoplasia* 2008, *13*. 337-42.
6. A. Balmain, Cancer as a complex genetic trait: tumor susceptibility in humans and mouse models. *Cell* 2002, *108*. 145-52.
7. L. M. Coussens, Z. Werb, Inflammation and cancer. *Nature* 2002, *420*. 860-7.
8. K. W. Hunter, N. P. Crawford, The future of mouse QTL mapping to diagnose disease in mice in the age of whole-genome association studies. *Annu Rev Genet* 2008, *42*. 131-41.
9. T. Walsh, M. C. King, Ten genes for inherited breast cancer. *Cancer Cell* 2007, *11*. 103-5.
10. J. McClellan, M. C. King, Genetic heterogeneity in human disease. *Cell* 2010, *141*. 210-7.
11. P. Lichtenstein, N. V. Holm, P. K. Verkasalo, A. Iliadou, J. Kaprio, M. Koskenvuo, E. Pukkala, A. Skytthe, K. Hemminki, Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000, *343*. 78-85.
12. K. Czene, P. Lichtenstein, K. Hemminki, Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer Database. *Int J Cancer* 2002, *99*. 260-6.

13. I. B. Borecki, M. A. Province, Linkage and association: basic concepts. *Adv Genet* 2008, *60*. 51-74.
14. F. H. Fodor, A. Weston, I. J. Bleiweiss, L. D. McCurdy, M. M. Walsh, P. I. Tartter, S. T. Brower, C. M. Eng, Frequency and carrier risk associated with common BRCA1 and BRCA2 mutations in Ashkenazi Jewish breast cancer patients. *Am J Hum Genet* 1998, *63*. 45-51.
15. K. L. Nathanson, R. Wooster, B. L. Weber, Breast cancer genetics: what we know and what we need. *Nat Med* 2001, *7*. 552-6.
16. A. Galvan, J. P. Ioannidis, T. A. Dragani, Beyond genome-wide association studies: genetic heterogeneity and individual predisposition to cancer. *Trends Genet* 2010, *26*. 132-41.
17. A. A. Donjacour, G. R. Cunha, Stromal regulation of epithelial function. *Cancer Treat Res* 1991, *53*. 335-64.
18. R. Paus, E. M. Peters, S. Eichmuller, V. A. Botchkarev, Neural mechanisms of hair growth control. *J Invest Dermatol Symp Proc* 1997, *2*. 61-8.
19. J. Perez-Losada, A. Balmain, Stem-cell hierarchy in skin cancer. *Nat Rev Cancer* 2003, *3*. 434-43.
20. H. F. Dvorak, Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. *N Engl J Med* 1986, *315*. 1650-9.
21. A. F. Olumi, G. D. Grossfeld, S. W. Hayward, P. R. Carroll, T. D. Tlsty, G. R. Cunha, Carcinoma-associated fibroblasts direct tumor progression of initiated human prostatic epithelium. *Cancer Res* 1999, *59*. 5002-11.
22. K. Kurose, K. Gilley, S. Matsumoto, P. H. Watson, X. P. Zhou, C. Eng, Frequent somatic mutations in PTEN and TP53 are mutually exclusive in the stroma of breast carcinomas. *Nat Genet* 2002, *32*. 355-7.
23. V. M. Weaver, P. Gilbert, Watch thy neighbor: cancer is a communal affair. *J Cell Sci* 2004, *117*. 1287-90.
24. A. J. Trimboli, C. Z. Cantemir-Stone, F. Li, J. A. Wallace, A. Merchant, N. Creasap, J. C. Thompson, E. Caserta, H. Wang, J. L. Chong, S. Naidu, G. Wei, S. M. Sharma, J. A. Stephens, S. A. Fernandez, M. N. Gurcan, M. B. Weinstein, S. H. Barsky, L. Yee, T. J. Rosol, P. C. Stromberg, M. L. Robinson, F. Pepin, M. Hallett, M. Park, M. C. Ostrowski, G. Leone, Pten in stromal fibroblasts suppresses mammary epithelial tumours. *Nature* 2009, *461*. 1084-91.
25. J. A. Joyce, J. W. Pollard, Microenvironmental regulation of metastasis. *Nat Rev Cancer* 2009, *9*. 239-52.
26. B. Qian, Y. Deng, J. H. Im, R. J. Muschel, Y. Zou, J. Li, R. A. Lang, J. W. Pollard, A distinct macrophage population mediates metastatic breast cancer cell extravasation, establishment and growth. *PLoS One* 2009, *4*. e6562.
27. B. Z. Qian, J. W. Pollard, Macrophage diversity enhances tumor progression and metastasis. *Cell* 2010, *141*. 39-51.
28. D. G. DeNardo, P. Andreu, L. M. Coussens, Interactions between lymphocytes and myeloid cells regulate pro- versus anti-tumor immunity. *Cancer Metastasis Rev* 2010, *29*. 309-16.
29. K. Kessenbrock, V. Plaks, Z. Werb, Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell* 2010, *141*. 52-67.
30. N. Bellam, B. Pasche, Tgf-beta signaling alterations and colon cancer. *Cancer Treat Res* 2010, *155*. 85-103.
31. M. MacPhee, K. P. Chepenik, R. A. Liddell, K. K. Nelson, L. D. Siracusa, A. M. Buchberg, The secretory phospholipase A2 gene is a candidate for the Mom1 locus, a major modifier of ApcMin-induced intestinal neoplasia. *Cell* 1995, *81*. 957-66.
32. W. Pruzanski, P. Vadas, Phospholipase A2--a mediator between proximal and distal effectors of inflammation. *Immunol Today* 1991, *12*. 143-6.
33. W. Jaross, R. Eckey, M. Menschikowski, Biological effects of secretory phospholipase A(2) group IIA on lipoproteins and in atherogenesis. *Eur J Clin Invest* 2002, *32*. 383-93.

34. S. Y. Leung, X. Chen, K. M. Chu, S. T. Yuen, J. Mathy, J. Ji, A. S. Chan, R. Li, S. Law, O. G. Troyanskaya, I. P. Tu, J. Wong, S. So, D. Botstein, P. O. Brown, Phospholipase A2 group IIA expression in gastric adenocarcinoma is associated with prolonged survival and less frequent metastasis. *Proc Natl Acad Sci U S A* 2002, **99**. 16203-8.
35. M. R. Stratton, N. Rahman, The emerging landscape of breast cancer susceptibility. *Nat Genet* 2008, **40**. 17-22.
36. J. C. Barrett, S. Hansoul, D. L. Nicolae, J. H. Cho, R. H. Duerr, J. D. Rioux, S. R. Brant, M. S. Silverberg, K. D. Taylor, M. M. Barmada, A. Bitton, T. Dassopoulos, L. W. Datta, T. Green, A. M. Griffiths, E. O. Kistner, M. T. Murtha, M. D. Regueiro, J. I. Rotter, L. P. Schumm, A. H. Steinhardt, S. R. Targan, R. J. Xavier, C. Libioulle, C. Sandor, M. Lathrop, J. Belaiche, O. Dewit, I. Gut, S. Heath, D. Laukens, M. Mni, P. Rutgeerts, A. Van Gossum, D. Zelenika, D. Franchimont, J. P. Hugot, M. de Vos, S. Vermeire, E. Louis, L. R. Cardon, C. A. Anderson, H. Drummond, E. Nimmo, T. Ahmad, N. J. Prescott, C. M. Onnie, S. A. Fisher, J. Marchini, J. Ghorri, S. Bumpstead, R. Gwilliam, M. Tremelling, P. Deloukas, J. Mansfield, D. Jewell, J. Satsangi, C. G. Mathew, M. Parkes, M. Georges, M. J. Daly, Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008, **40**. 955-62.
37. P. J. Wysocki, B. Wierusz-Wysocka, Obesity, hyperinsulinemia and breast cancer: novel targets and a novel role for metformin. *Expert Rev Mol Diagn* 2010, **10**. 509-19.
38. P. Tai, E. Yu, K. Joseph, T. Miale, A review of autoimmune diseases associated with cancer. *Front Biosci (Elite Ed)* 2010, **2**. 122-6.
39. A. A. Siddiqui, Metabolic Syndrome and Its Association With Colorectal Cancer: A Review. *Am J Med Sci* 2010.
40. K. I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, A. L. Barabasi, The human disease network. *Proc Natl Acad Sci U S A* 2007, **104**. 8685-90.
41. N. Risch, K. Merikangas, The future of genetic studies of complex human diseases. *Science* 1996, **273**. 1516-7.
42. H. Meijers-Heijboer, A. van den Ouweland, J. Klijn, M. Wasielewski, A. de Snoo, R. Oldenburg, A. Hollestelle, M. Houben, E. Crepin, M. van Veghel-Plandsoen, F. Elstrodt, C. van Duijn, C. Bartels, C. Meijers, M. Schutte, L. McGuffog, D. Thompson, D. Easton, N. Sodha, S. Seal, R. Barfoot, J. Mangion, J. Chang-Claude, D. Eccles, R. Eeles, D. G. Evans, R. Houlston, V. Murday, S. Narod, T. Peretz, J. Peto, C. Phelan, H. X. Zhang, C. Szabo, P. Devilee, D. Goldgar, P. A. Futreal, K. L. Nathanson, B. Weber, N. Rahman, M. R. Stratton, Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet* 2002, **31**. 55-9.
43. J. M. Rosa-Rosa, G. Pita, M. Urioste, G. Llort, J. Brunet, C. Lazaro, I. Blanco, T. Ramon y Cajal, O. Diez, M. de la Hoya, T. Caldes, M. I. Tejada, A. Gonzalez-Neira, J. Benitez, Genome-wide linkage scan reveals three putative breast-cancer-susceptibility loci. *Am J Hum Genet* 2009, **84**. 115-22.
44. D. F. Easton, K. A. Pooley, A. M. Dunning, P. D. Pharoah, D. Thompson, D. G. Ballinger, J. P. Struewing, J. Morrison, *et al.*, Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007, **447**. 1087-93.
45. S. N. Stacey, A. Manolescu, P. Sulem, T. Rafnar, J. Gudmundsson, S. A. Gudjonsson, G. Masson, M. Jakobsdottir, S. Thorlacius, A. Helgason, K. K. Aben, L. J. Strobbe, M. T. Albers-Akkers, D. W. Swinkels, B. E. Henderson, L. N. Kolonel, L. Le Marchand, E. Millastre, R. Andres, J. Godino, M. D. Garcia-Prats, E. Polo, A. Tres, M. Mouy, J. Saemundsdottir, V. M. Backman, L. Gudmundsson, K. Kristjansson, J. T. Bergthorsson, J. Kostic, M. L. Frigge, F. Geller, D. Gudbjartsson, H. Sigurdsson, T. Jonsson, J. Hrafnkelsson, J. Johannsson, T. Sveinsson, G. Myrdal, H. N. Grimsson, T. Jonsson, S. von Holst, B. Werelius, S. Margolin, A. Lindblom, J. I. Mayordomo, C. A. Haiman, L. A. Kiemeny, O. T. Johannsson, J. R. Gulcher, U. Thorsteinsdottir, A.

- Kong, K. Stefansson, Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2007, 39. 865-9.
46. J. M. Murabito, C. L. Rosenberg, D. Finger, B. E. Kreger, D. Levy, G. L. Splansky, K. Antman, S. J. Hwang, A genome-wide association study of breast and prostate cancer in the NHLBI's Framingham Heart Study. *BMC Med Genet* 2007, 8 Suppl 1. S6.
47. D. J. Hunter, P. Kraft, K. B. Jacobs, D. G. Cox, M. Yeager, S. E. Hankinson, S. Wacholder, Z. Wang, R. Welch, A. Hutchinson, J. Wang, K. Yu, N. Chatterjee, N. Orr, W. C. Willett, G. A. Colditz, R. G. Ziegler, C. D. Berg, S. S. Buys, C. A. McCarty, H. S. Feigelson, E. E. Calle, M. J. Thun, R. B. Hayes, M. Tucker, D. S. Gerhard, J. F. Fraumeni, Jr., R. N. Hoover, G. Thomas, S. J. Chanock, A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007, 39. 870-4.
48. B. Gold, T. Kirchhoff, S. Stefanov, J. Lautenberger, A. Viale, J. Garber, E. Friedman, S. Narod, A. B. Olshen, P. Gregersen, K. Kosarin, A. Olsh, J. Bergeron, N. A. Ellis, R. J. Klein, A. G. Clark, L. Norton, M. Dean, J. Boyd, K. Offit, Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. *Proc Natl Acad Sci U S A* 2008, 105. 4340-5.
49. W. Zheng, J. Long, Y. T. Gao, C. Li, Y. Zheng, Y. B. Xiang, W. Wen, S. Levy, S. L. Deming, J. L. Haines, K. Gu, A. M. Fair, Q. Cai, W. Lu, X. O. Shu, Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat Genet* 2009, 41. 324-8.
50. G. Thomas, K. B. Jacobs, P. Kraft, M. Yeager, S. Wacholder, D. G. Cox, S. E. Hankinson, A. Hutchinson, Z. Wang, K. Yu, N. Chatterjee, M. Garcia-Closas, J. Gonzalez-Bosquet, L. Prokunina-Olsson, N. Orr, W. C. Willett, G. A. Colditz, R. G. Ziegler, C. D. Berg, S. S. Buys, C. A. McCarty, H. S. Feigelson, E. E. Calle, M. J. Thun, R. Diver, R. Prentice, R. Jackson, C. Kooperberg, R. Chlebowski, J. Lissowska, B. Peplonska, L. A. Brinton, A. Sigurdson, M. Doody, P. Bhatti, B. H. Alexander, J. Buring, I. M. Lee, L. J. Vatten, K. Hveem, M. Kumle, R. B. Hayes, M. Tucker, D. S. Gerhard, J. F. Fraumeni, Jr., R. N. Hoover, S. J. Chanock, D. J. Hunter, A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat Genet* 2009, 41. 579-84.
51. S. Ahmed, G. Thomas, M. Ghousaini, C. S. Healey, M. K. Humphreys, R. Platte, J. Morrison, M. Maranian, *et al.*, Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat Genet* 2009, 41. 585-90.
52. S. N. Stacey, A. Manolescu, P. Sulem, S. Thorlacius, S. A. Gudjonsson, G. F. Jonsson, M. Jakobsdottir, J. T. Bergthorsson, J. Gudmundsson, K. K. Aben, L. J. Strobbe, D. W. Swinkels, K. C. van Engelenburg, B. E. Henderson, L. N. Kolonel, L. Le Marchand, E. Millastre, R. Andres, B. Saez, J. Lambea, J. Godino, E. Polo, A. Tres, S. Picelli, J. Rantala, S. Margolin, T. Jonsson, H. Sigurdsson, T. Jonsdottir, J. Hrafnelsson, J. Johannsson, T. Sveinsson, G. Myrdal, H. N. Grimsson, S. G. Sveinsdottir, K. Alexiusdottir, J. Saemundsdottir, A. Sigurdsson, J. Kostic, L. Gudmundsson, K. Kristjansson, G. Masson, J. D. Fackenthal, C. Adebamowo, T. Ogundiran, O. I. Olopade, C. A. Haiman, A. Lindblom, J. I. Mayordomo, L. A. Kiemeny, J. R. Gulcher, T. Rafnar, U. Thorsteinsdottir, O. T. Johannsson, A. Kong, K. Stefansson, Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2008, 40. 703-6.
53. R. L. Milne, J. Benitez, H. Nevanlinna, T. Heikkinen, K. Aittomaki, C. Blomqvist, J. I. Arias, M. P. Zamora, B. Burwinkel, C. R. Bartram, A. Meindl, R. K. Schmutzler, A. Cox, I. Brock, G. Elliott, M. W. Reed, M. C. Southey, L. Smith, A. B. Spurdle, J. L. Hopper, F. J. Couch, J. E. Olson, X. Wang, Z. Fredericksen, P. Schurmann, M. Bremer, P. Hillemanns, T. Dork, P. Devilee, C. J. van Asperen, R. A. Tollenaar, C. Seynaeve, P. Hall, K. Czene, J. Liu, Y. Li, S. Ahmed, A. M. Dunning, M. Maranian, P. D. Pharoah, G. Chenevix-Trench, J. Beesley, N. V. Bogdanova, N. N. Antonenkova, I. V. Zalutsky, H. Anton-Culver, A. Ziogas, H. Brauch, C. Justenhoven, Y. D. Ko, S. Haas,

- P. A. Fasching, R. Strick, A. B. Ekici, M. W. Beckmann, G. G. Giles, G. Severi, L. Baglietto, D. R. English, O. Fletcher, N. Johnson, I. dos Santos Silva, J. Peto, C. Turnbull, S. Hines, A. Renwick, N. Rahman, B. G. Nordestgaard, S. E. Bojesen, H. Flyger, D. Kang, K. Y. Yoo, D. Y. Noh, A. Mannermaa, V. Kataja, V. M. Kosma, M. Garcia-Closas, S. Chanock, J. Lissowska, L. A. Brinton, J. Chang-Claude, S. Wang-Gohrke, C. Y. Shen, H. C. Wang, J. C. Yu, S. T. Chen, M. Bermisheva, T. Nikolaeva, E. Khusnutdinova, M. K. Humphreys, J. Morrison, R. Platte, D. F. Easton, Risk of estrogen receptor-positive and -negative breast cancer and single-nucleotide polymorphism 2q35-rs13387042. *J Natl Cancer Inst* 2009, *101*. 1012-8.
54. P. Boyle, P. Boffetta, P. Autier, Diet, nutrition and cancer: public, media and scientific confusion. *Ann Oncol* 2008, *19*. 1665-7.
55. R. C. Travis, G. K. Reeves, J. Green, D. Bull, S. J. Tipper, K. Baker, V. Beral, R. Peto, J. Bell, D. Zelenika, M. Lathrop, Gene-environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study. *Lancet* 2010, *375*. 2143-51.
56. C. J. Hoggart, T. G. Clark, M. De Iorio, J. C. Whittaker, D. J. Balding, Genome-wide significance for dense SNP and resequencing data. *Genet Epidemiol* 2008, *32*. 179-85.
57. J. P. Ioannidis, E. E. Ntzani, T. A. Trikalinos, D. G. Contopoulos-Ioannidis, Replication validity of genetic association studies. *Nat Genet* 2001, *29*. 306-9.
58. M. J. Khoury, L. Bertram, P. Boffetta, A. S. Butterworth, S. J. Chanock, S. M. Dolan, I. Fortier, M. Garcia-Closas, M. Gwinn, J. P. Higgins, A. C. Janssens, J. Ostell, R. P. Owen, R. A. Pagon, T. R. Rebbeck, N. Rothman, J. L. Bernstein, P. R. Burton, H. Campbell, A. Chockalingam, H. Furberg, J. Little, T. R. O'Brien, D. Seminara, P. Vineis, D. M. Winn, W. Yu, J. P. Ioannidis, Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases. *Am J Epidemiol* 2009, *170*. 269-79.
59. J. P. Ioannidis, Calibration of credibility of agnostic genome-wide associations. *Am J Med Genet B Neuropsychiatr Genet* 2008, *147B*. 964-72.
60. J. P. Ioannidis, E. Y. Loy, R. Poulton, K. S. Chia, Researching genetic versus nongenetic determinants of disease: a comparison and proposed unification. *Sci Transl Med* 2009, *1*. 7ps8.
61. E. Linos, W. C. Willett, Diet and breast cancer risk reduction. *J Natl Compr Canc Netw* 2007, *5*. 711-718.
62. K. Britt, A. Ashworth, M. Smalley, Pregnancy and the risk of breast cancer. *Endocr Relat Cancer* 2007, *14*. 907-33.
63. A. Schatzkin, M. P. Longnecker, Alcohol and breast cancer. Where are we now and where do we go from here? *Cancer* 1994, *74*. 1101-10.
64. C. B. Ambrosone, The promise and limitations of genome-wide association studies to elucidate the causes of breast cancer. *Breast Cancer Res* 2007, *9*. 114.
65. E. A. Ostrander, J. L. Stanford, Genetics of prostate cancer: too many loci, too few genes. *Am J Hum Genet* 2000, *67*. 1367-75.
66. P. A. Futreal, Q. Liu, D. Shattuck-Eidens, C. Cochran, K. Harshman, S. Tavtigian, L. M. Bennett, A. Haugen-Strano, J. Swensen, Y. Miki, et al., BRCA1 mutations in primary breast and ovarian carcinomas. *Science* 1994, *266*. 120-2.
67. R. A. Oldenburg, K. Kroeze-Jansema, J. Kraan, H. Morreau, J. G. Klijn, N. Hoogerbrugge, M. J. Ligtenberg, C. J. van Asperen, H. F. Vasen, C. Meijers, H. Meijers-Heijboer, T. H. de Bock, C. J. Cornelisse, P. Devilee, The CHEK2*1100delC variant acts as a breast cancer risk modifier in non-BRCA1/BRCA2 multiple-case families. *Cancer Res* 2003, *63*. 8153-7.
68. S. Wang-Gohrke, T. R. Rebbeck, W. Besenfelder, R. Kreienberg, I. B. Runnebaum, p53 germline polymorphisms are associated with an increased risk for breast cancer in German women. *Anticancer Res* 1998, *18*. 2095-9.

69. R. Wooster, G. Bignell, J. Lancaster, S. Swift, S. Seal, J. Mangion, N. Collins, S. Gregory, C. Gumbs, G. Micklem, Identification of the breast cancer susceptibility gene BRCA2. *Nature* 1995, 378. 789-92.
70. T. van Wezel, A. P. Stassen, C. J. Moen, A. A. Hart, M. A. van der Valk, P. Demant, Gene interaction and single gene effects in colon tumour susceptibility in mice. *Nat Genet* 1996, 14. 468-70.
71. R. J. Fijneman, R. C. Jansen, M. A. van der Valk, P. Demant, High frequency of interactions between lung cancer susceptibility genes in the mouse: mapping of Sluc5 to Sluc14. *Cancer Res* 1998, 58. 4794-8.
72. T. F. Mackay, Quantitative trait loci in Drosophila. *Nat Rev Genet* 2001, 2. 11-20.
73. H. Nagase, J. H. Mao, J. P. de Koning, T. Minami, A. Balmain, Epistatic interactions between skin tumor modifier loci in interspecific (spretus/musculus) backcross mice. *Cancer Res* 2001, 61. 1305-8.
74. D. A. Quigley, M. D. To, J. Perez-Losada, F. G. Pelorosso, J. H. Mao, H. Nagase, D. G. Ginzinger, A. Balmain, Genetic architecture of mouse skin inflammation and tumour susceptibility. *Nature* 2009, 458. 505-8.
75. D. Quigley, A. Balmain, Systems genetics analysis of cancer susceptibility: from mouse models to humans. *Nat Rev Genet* 2009, 10. 651-7.
76. A. Balmain, H. Nagase, Cancer resistance genes in mice: models for the study of tumour modifiers. *Trends Genet* 1998, 14. 139-44.
77. J. H. Nadeau, W. N. Frankel, The roads from phenotypic variation to gene discovery: mutagenesis versus QTLs. *Nat Genet* 2000, 25. 381-4.
78. A. Matin, G. B. Collin, Y. Asada, D. Varnum, J. H. Nadeau, Susceptibility to testicular germ-cell tumours in a 129.MOLF-Chr 19 chromosome substitution strain. *Nat Genet* 1999, 23. 237-40.
79. F. S. Collins, R. H. Finnell, J. Rossant, W. Wurst, A new partner for the international knockout mouse consortium. *Cell* 2007, 129. 235.
80. A. Darvasi, Dissecting complex traits: the geneticists' "Around the world in 80 days". *Trends Genet* 2005, 21. 373-6.
81. M. D. To, J. Perez-Losada, J. H. Mao, J. Hsu, T. Jacks, A. Balmain, A functional switch from lung cancer resistance to susceptibility at the Pas1 locus in Kras2LA2 mice. *Nat Genet* 2006, 38. 926-30.
82. A. Balmain, C. C. Harris, Carcinogenesis in mouse and human cells: parallels and paradoxes. *Carcinogenesis* 2000, 21. 371-7.
83. X. Wang, B. Paigen, Genetics of variation in HDL cholesterol in humans and mice. *Circ Res* 2005, 96. 27-42.
84. N. P. Crawford, X. Qian, A. Ziogas, A. G. Papageorge, B. J. Boersma, R. C. Walker, L. Lukes, W. L. Rowe, J. Zhang, S. Ambs, D. R. Lowy, H. Anton-Culver, K. W. Hunter, Rrp1b, a new candidate susceptibility gene for breast cancer progression and metastasis. *PLoS Genet* 2007, 3. e214.
85. N. P. Crawford, A. Ziogas, D. J. Peel, J. Hess, H. Anton-Culver, K. W. Hunter, Germline polymorphisms in SIPA1 are associated with metastasis and other indicators of poor prognosis in breast cancer. *Breast Cancer Res* 2006, 8. R16.
86. A. Ewart-Toland, P. Briassouli, J. P. de Koning, J. H. Mao, J. Yuan, F. Chan, L. MacCarthy-Morrogh, B. A. Ponder, H. Nagase, J. Burn, S. Ball, M. Almeida, S. Linardopoulos, A. Balmain, Identification of Stk6/STK15 as a candidate low-penetrance tumor-susceptibility gene in mouse and human. *Nat Genet* 2003, 34. 403-12.
87. A. Ewart-Toland, Q. Dai, Y. T. Gao, H. Nagase, M. G. Dunlop, S. M. Farrington, R. A. Barnetson, H. Anton-Culver, D. Peel, A. Ziogas, D. Lin, X. Miao, T. Sun, E. A. Ostrander, J. L. Stanford, M. Langlois, J. M. Chan, J. Yuan, C. C. Harris, E. D. Bowman, G. L. Clayman, S. M. Lippman, J. J. Lee, W. Zheng, A. Balmain, Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. *Carcinogenesis* 2005, 26. 1368-73.

88. Y. G. Park, R. Clifford, K. H. Buetow, K. W. Hunter, Multiple cross and inbred strain haplotype mapping of complex-trait candidate genes. *Genome Res* 2003, *13*. 118-21.
89. X. Wang, R. Korstanje, D. Higgins, B. Paigen, Haplotype analysis in multiple crosses to identify a QTL gene. *Genome Res* 2004, *14*. 1767-72.
90. R. Hitzemann, B. Malmanger, S. Cooper, S. Coulombe, C. Reed, K. Demarest, J. Koyner, L. Cipp, J. Flint, C. Talbot, B. Rademacher, K. Buck, J. McCaughran, Jr., Multiple cross mapping (MCM) markedly improves the localization of a QTL for ethanol-induced activation. *Genes Brain Behav* 2002, *1*. 214-22.
91. R. Hitzemann, B. Malmanger, C. Reed, M. Lawler, B. Hitzemann, S. Coulombe, K. Buck, B. Rademacher, N. Walter, Y. Polyakov, J. Sikela, B. Gensler, S. Burgers, R. W. Williams, K. Manly, J. Flint, C. Talbot, A strategy for the integration of QTL, gene expression, and sequence analyses. *Mamm Genome* 2003, *14*. 733-47.
92. B. Yalcin, S. A. Willis-Owen, J. Fullerton, A. Meesaq, R. M. Deacon, J. N. Rawlins, R. R. Copley, A. P. Morris, J. Flint, R. Mott, Genetic dissection of a behavioral quantitative trait locus shows that *Rgs2* modulates anxiety in mice. *Nat Genet* 2004, *36*. 1197-202.
93. R. Mott, J. Flint, Simultaneous detection and fine mapping of quantitative trait loci in mice using heterogeneous stocks. *Genetics* 2002, *160*. 1609-18.
94. R. Chia, F. Achilli, M. F. Festing, E. M. Fisher, The origins and uses of mouse outbred stocks. *Nat Genet* 2005, *37*. 1181-6.
95. G. Manenti, F. Galbiati, R. Gianni-Barrera, A. Pettinicchio, A. Acevedo, T. A. Dragani, Haplotype sharing suggests that a genomic segment containing six genes accounts for the pulmonary adenoma susceptibility 1 (*Pas1*) locus activity in mice. *Oncogene* 2004, *23*. 4495-504.
96. H. Nagase, S. Bryson, H. Cordell, C. J. Kemp, F. Fee, A. Balmain, Distinct genetic loci control development of benign and malignant skin tumours in mice. *Nat Genet* 1995, *10*. 424-9.
97. H. Nagase, J. H. Mao, A. Balmain, A subset of skin tumor modifier loci determines survival time of tumor-bearing mice. *Proc Natl Acad Sci U S A* 1999, *96*. 15032-7.
98. G. Manenti, F. Galbiati, S. Noci, T. A. Dragani, Outbred CD-1 mice carry the susceptibility allele at the pulmonary adenoma susceptibility 1 (*Pas1*) locus. *Carcinogenesis* 2003, *24*. 1143-8.
99. J. Flint, W. Valdar, S. Shifman, R. Mott, Strategies for mapping and cloning quantitative trait genes in rodents. *Nat Rev Genet* 2005, *6*. 271-86.
100. J. Flint, Mapping quantitative traits and strategies to find quantitative trait genes. *Methods* 2010.
101. G. A. Churchill, D. C. Airey, H. Allayee, J. M. Angel, A. D. Attie, J. Beatty, W. D. Beavis, J. K. Belknap, et al., The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat Genet* 2004, *36*. 1133-7.
102. F. A. Iraqi, G. Churchill, R. Mott, The Collaborative Cross, developing a resource for mammalian systems genetics: a status report of the Wellcome Trust cohort. *Mamm Genome* 2008, *19*. 379-81.
103. A. Balmain, J. Gray, B. Ponder, The genetics and genomics of cancer. *Nat Genet* 2003, *33 Suppl*. 238-44.
104. M. Arbilly, A. Pisante, M. Devor, A. Darvasi, An integrative approach for the identification of quantitative trait loci. *Anim Genet* 2006, *37 Suppl 1*. 7-9.
105. R. C. Jansen, J. P. Nap, Genetical genomics: the added value from segregation. *Trends Genet* 2001, *17*. 388-91.
106. K. DiPetrillo, X. Wang, I. M. Stylianou, B. Paigen, Bioinformatics toolbox for narrowing rodent quantitative trait loci. *Trends Genet* 2005, *21*. 683-92.
107. E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, et al., Initial sequencing and analysis of the human genome. *Nature* 2001, *409*. 860-921.

108. J. C. Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, *et al.*, The sequence of the human genome. *Science* 2001, **291**. 1304-51.
109. R. H. Waterston, K. Lindblad-Toh, E. Birney, J. Rogers, J. F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, *et al.*, Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002, **420**. 520-62.
110. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 2005, **437**. 69-87.
111. K. Lindblad-Toh, C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe, M. Kamal, M. Clamp, J. L. Chang, *et al.*, Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 2005, **438**. 803-19.
112. K. A. Frazer, E. Eskin, H. M. Kang, M. A. Bogue, D. A. Hinds, E. J. Beilharz, R. V. Gupta, J. Montgomery, M. M. Morenzoni, G. B. Nilsen, C. L. Pethiyagoda, L. L. Stuve, F. M. Johnson, M. J. Daly, C. M. Wade, D. R. Cox, A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* 2007, **448**. 1050-3.
113. R. J. Klein, X. Xu, S. Mukherjee, J. Willis, J. Hayes, Successes of genome-wide association studies. *Cell* 2010, **142**. 350-1; author reply 353-5.
114. K. Wang, M. Bucan, S. F. Grant, G. Schellenberg, H. Hakonarson, Strategies for genetic studies of complex diseases. *Cell* 2010, **142**. 351-3; author reply 353-5.
115. B. W. Zanke, C. M. Greenwood, J. Rangrej, R. Kustra, A. Tenesa, S. M. Farrington, J. Prendergast, S. Olschwang, T. Chiang, E. Crowdy, V. Ferretti, P. Laflamme, S. Sundararajan, S. Roumy, J. F. Olivier, F. Robidoux, R. Sladek, A. Montpetit, P. Campbell, S. Bezieau, A. M. O'Shea, G. Zogopoulos, M. Cotterchio, P. Newcomb, J. McLaughlin, B. Younghusband, R. Green, J. Green, M. E. Porteous, H. Campbell, H. Blanche, M. Sahbatou, E. Tubacher, C. Bonaiti-Pellie, B. Buecher, E. Riboli, S. Kury, S. J. Chanock, J. Potter, G. Thomas, S. Gallinger, T. J. Hudson, M. G. Dunlop, Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet* 2007, **39**. 989-94.
116. I. Tomlinson, E. Webb, L. Carvajal-Carmona, P. Broderick, Z. Kemp, S. Spain, S. Penegar, I. Chandler, M. Gorman, W. Wood, E. Barclay, S. Lubbe, L. Martin, G. Sellick, E. Jaeger, R. Hubner, R. Wild, A. Rowan, S. Fielding, K. Howarth, A. Silver, W. Atkin, K. Muir, R. Logan, D. Kerr, E. Johnstone, O. Sieber, R. Gray, H. Thomas, J. Peto, J. B. Cazier, R. Houlston, A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet* 2007, **39**. 984-8.
117. K. A. Frazer, S. S. Murray, N. J. Schork, E. J. Topol, Human genetic variation and its contribution to complex traits. *Nat Rev Genet* 2009, **10**. 241-51.

Figure Legends

Fig. 1. Cancer is not purely a cell-autonomous disease: There are connections between cancer cells and immune and endocrine systems, vasculature and stroma surrounding them that modify the tumor behaviour and susceptibility.

Fig. 2. Tumor risk is the consequence of the interaction between constitutional genetics and environmental exposures. The combination between the genetic background (modifier genes, mainly low susceptibility cancer genes) and the environmental factors varies among individuals and might explain the different tumor susceptibility and behaviour observed in patients.

Fig. 3. Recent technical advances have improved QTL research: The use of genetically more complex mice (outbred and artificial strains) in combination with high resolution SNPs arrays and new techniques of statistical genetics greatly improve the definition of new QTLs. In tumor cells, the analysis of changes in copy number and expression by whole genome comparative hybridization and expression arrays allow the further refinement of QTL.

Fig. 4. Mouse models are a good tool to identify QTLs in an environmentally controlled way that could be extrapolated to human population. At the same time they are very useful to verify and refine candidate loci found in humans by GWAS. Technical advances such as whole genome sequencing are readily making the recognition of orthologous chromosomal regions between species straight forward simplifying the refinement of QTLs found in both kind of studies.

Table-1: Main regions found by human GWAS for breast cancer susceptibility

Susceptibility Region	Reference
5q11.2; 8q24; 10q26; 11p15.5; 16q12.1	Easton et al., 2007 ⁴⁴
2q35; 16q12	Stacey et al., 2007 ⁴⁵
Three ERBB4 SNPs	Murabito et al., 2007 ⁴⁶
10q26 (intron 2 of FGFR2)	Hunter et al., 2007 ⁴⁷
6q22.33	Gold et al., 2008 ⁴⁸
6q25.1	Zheng et al., 2009 ⁴⁹
1p11.2; 14q24.1	Thomas et al., 2009 ⁵⁰
3p24; 17q23	Ahmed, et al., 2009 ⁵¹
5p12	Stacey et al., 2008 ⁵²
2q35	Milne et al., 2009 ⁵³
