

# Single nucleotide polymorphisms and cancer susceptibility

Na Deng<sup>1,2</sup>, Heng Zhou<sup>1</sup>, Hua Fan<sup>2</sup> and Yuan Yuan<sup>1,3</sup>

<sup>1</sup>Tumor Etiology and Screening Department of Cancer Institute and General Surgery, The First Affiliated Hospital of China Medical University, and Key Laboratory of Cancer Etiology and Prevention (China Medical University), Liaoning Provincial Education Department, Shenyang 110001, China

<sup>2</sup>Department of Hematology, The Fourth Affiliated Hospital of China Medical University, Shenyang 110001, China

<sup>3</sup>National Clinical Research Center for Digestive Diseases, Xi'an 110001, China

**Correspondence to:** Yuan Yuan, **email:** yuanyuan@cmu.edu.cn

**Keywords:** *single nucleotide polymorphism, genetic, epigenetic, susceptibility, cancer*

**Received:** August 26, 2016

**Accepted:** October 03, 2017

**Published:** November 07, 2017

**Copyright:** Deng et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

**A large number of genes associated with various cancer types contain single nucleotide polymorphisms (SNPs). SNPs are located in gene promoters, exons, introns as well as 5'- and 3'- untranslated regions (UTRs) and affect gene expression by different mechanisms. These mechanisms depend on the role of the genetic elements in which the individual SNPs are located. Moreover, alterations in epigenetic regulation due to gene polymorphisms add to the complexity underlying cancer susceptibility related to SNPs. In this systematic review, we discuss the various genetic and epigenetic mechanisms involved in determining cancer susceptibility related to various SNPs located in different genetic elements. We also discuss the diagnostic potential of these SNPs and the focus for future studies.**

## INTRODUCTION

Single nucleotide polymorphisms (SNPs) are one of the most common types of genetic variations in the human genome. SNPs in genes that regulate DNA mismatch repair, cell cycle regulation, metabolism and immunity are associated with genetic susceptibility to cancer [1–12]. Understanding the mechanisms underlying the effects of SNPs that result in cancer susceptibility is critical to understanding the molecular pathogenesis of various cancers. From a clinical perspective, SNPs are potential diagnostic and therapeutic biomarkers in many cancer types.

SNPs are located in different regions of genes such as promoters, exons, introns as well as 5'- and 3' UTRs. Therefore, alterations in gene expression and their effect on cancer susceptibility vary depending on the location of the SNPs. The promoter region SNPs affect gene expression by altering promoter activity, transcription-factor binding, DNA methylation and histone modifications [13–20]. The exonal SNPs affect cancer susceptibility by suppressing gene transcription and translation [21–23]. SNPs in intron regions generate splice variants of transcripts and promote or disrupt binding and function of long non-coding RNAs (lncRNAs) [24–26]. SNPs in the 5'-UTR affect translation, whereas SNPs in the 3'-UTR affect microRNA (miRNA) binding [9, 27–29]. SNPs in regions that are located far

from the actual genes reduce or enhance gene transcription through long-range cis effects [30]. In this systematic review, we discuss the genetic and epigenetic mechanisms that underlie the SNP-related cancer susceptibility and the potential utility of SNPs as cancer biomarkers.

## PROMOTER REGION SNPs AND CANCER SUSCEPTIBILITY

The promoter region regulates the initiation and rate of gene transcription through cis-acting elements and trans-acting factors. Promoter-related polymorphisms affect transcription factor binding that alter promoter activity, gene transcription, mRNA stability and translation. Subsequently, these effects alter protein levels that potentially determine the individual's susceptibility to diseases including cancer. Moreover, polymorphisms in the promoter regions also affect cancer susceptibility by altering epigenetic mechanisms such as DNA methylation and histone modifications (Figure 1).

### Promoter region SNPs affect promoter activity

#### Effect of promoter region SNPs on TATA box function

Polymorphisms in the TATA box inhibit promoter activity and decrease genetic transcription because it is

part of the core promoter. For example, a point mutation (A > C) at position -27 in the putative TATA box of the *EDH17B2* gene decreases promoter activity [31].

### **Effect of promoter region SNPs on transcription factor binding**

The promoter region contains numerous binding sites for transcription factors that regulate gene transcription. Polymorphisms in the promoter region decrease the transcription of target genes by inhibiting the binding of the transcription factors to the promoter. The rs16260 (-160C > A) and rs5030625 (-347G-> GA) in the *CDH1* promoter reduce transcriptional activity to different degrees [32, 33]. The rs16260(-160C > A) SNP decreases gene transcription by inhibiting transcription factor binding at the *CDH1* promoter and promotes tumorigenesis, including prostate, breast, colon and pancreatic cancers, etc [32]. The *CDH1* rs5030625 (-347G->GA) decreases CDH1 expression by 10-times (P < 0.001) via inhibiting transcription factor binding and correlates with increased familial gastric cancer susceptibility [33]. Conversely, *APE1* rs1760944(-656T > G) is associated with increased transcription factor binding and higher cervical cancer risk [34].

Promoter region SNPs modulate binding between transcription factor and promoter based on the number of binding sites for transcription factors such as SP1, c-Myb, E2F1, Ets, and GATA-1. For example, the rs2596538 in the *MICA* promoter region increases the risk of hepatitis C virus-associated liver cancer (OR = 1.34) with the transcription factors binding to the G allele and not the A allele [35]. The rs689466 (-1195G > A) SNP in the promoter region of the *COX-2* gene generates a new c-Myb binding site, thereby enhancing *COX-2* expression and increasing the risk of oesophageal cancer by 1.72-fold (95% CI, 1.35–2.20) [36]. A T→G substitution rs2279744 in the promoter region of *MDM2* increases transcriptional activity by enhancing the binding between the promoter and MDM2 transcripts; MDM2 overexpression mediated by cyclin D1 promotes tumorigenesis [37]. The rs1799750 -16071G→2G polymorphism in the *MMP-1* promoter generates a Ets binding site that increases *MMP-1* expression [38]. The C→T transitions at -1306 and -735 in the *MMP-2* promoter eliminate a Sp1-binding site (CCACC box), thereby decreasing *MMP-2* transcription and increasing the risk of esophageal squamous cell carcinoma (OR = 6.53; 95% CI = 2.78–15.33) [39]. The rs8179090 (-418G > C) in the *TIMP-2* promoter eliminates the Sp1-binding site and decreases *TIMP-2* expression [40].

SNPs in cis-acting elements such as GATA-1 transcription factor binding sites enhance the promoter activity of the survivin gene in breast cancer patients. For example, The G→A transitions at -235 SNP in the promoter of survivin gene generates a second GATA-1

binding site in its promoter region, thereby increasing survivin expression in breast cancer tissues [41].

### **Promoter region SNPs impact epigenetic mechanisms**

#### **Effect of promoter region SNPs on DNA methylation**

DNA methylation occurs primarily in the CpG islands of the promoter region. Therefore, SNPs in the promoter region can alter DNA methylation status and profoundly impact gene expression. The SNPs associated with DNA methylation of CpG loci are referred to as methylation SNPs [15, 42–44]. SNPs in genes rich in G and C bases are frequently mutated in many human diseases. SNPs in the promoter region alter the number of CpG dinucleotides in a CpG island, leading to changes in methylation, histone acetylation, chromatin modification, and gene silencing [45].

#### **Promoter region SNPs alter the number of methylation loci, thus changing gene expression and increasing the risk of cancer**

Some promoter region SNPs alter methylation in an allele-specific manner. Genome wide association studies (GWAS) showed that 38 SNPs in 12 CpG loci correlated with changes in methylation and expression of 10 genes (*IRF6*, *TSPYL5*, *CRIM1*, *CHL1*, *DDT*, *PIGC*, *TMOD1*, *ZNF266*, *BDKRB2*, *GSTT1*) [46]. Zhang *et al.* reported that the *CHEK2* rs2236141 (-48 G > A) variant was associated with lower lung cancer risk (adjusted OR = 0.73) because it eliminated a methylation locus, thereby relieving transcriptional repression [47]. The *EZH2* rs6950683 C allele showed reduced OSCC risk compared to the wildtype T allele because it was methylated and resulted in lower *EZH2* expression [48].

Furthermore, CpG SNPs affect many non-imprinted autosomal genes in normal human tissues by allele-specific DNA methylation (ASM), allele-specific gene expression (ASE) and allele-specific transcription factor binding (ASTF) [49]. SNPs in the promoter region modulate cytosine methylation of adjacent non-polymorphic CpG sites by increasing or decreasing the methylated loci. Methylation is more effective in regions of high CpG density or regions with abundant CpG SNPs since they promote effective binding of methylating enzymes and related factors.

#### **Promoter region SNPs affect transcription factor binding that protects CpG islands**

DNA-protein interactions in the promoter region can influence the methylation status and gene transcription. For example, transcription factors such as SP1 and CTCF influence methylation on CpG islands [49]. In sporadic breast tumors, methylation status was altered by mutations in the SP1 and CTCF binding sites in the promoter, which correlated with *BRCA1* downregulation [50]. SNPs in the Sp1-binding domain of RIL which is a frequent

methylation target in cancer prevent SP-1 binding and increase CpG methylation. Boumber *et al.* showed that SNPs influence time-dependent gene silencing by altering methylation levels [51].

### Promoter region SNPs alter DNA methylation by activating methylation-related enzymes

Polymorphisms in the promoter region of DNA methyltransferase (*DNMT*), methylene tetrahydrofolate reductase (*MTHFR*) and methionine synthase (*MS*) inhibit DNA synthesis and promote abnormal DNA methylation. Ogino *et al.* identified a common MGMT promoter SNP, rs16906252 (-56C > T), which recruited DNA methyltransferases resulting in loss of MGMT expression in colorectal tumors [52].

### Promoter region SNPs affects histone modification

Polymorphisms in non-coding regulatory sequences alter histone modifications such as acetylation, methylation, phosphorylation, ubiquitination, and glycosylation, which affect transcriptional rates. Promoter SNPs located in transcription factor binding sites are associated with regional histone modifications. SNPs in distant regulatory regions also regulate gene expression [53]. For example, rs1800896 (A-1082G) GG genotype has been found to be associated with high levels of IL-10 production [54]. Lipopolysaccharide (LPS) treatment increased

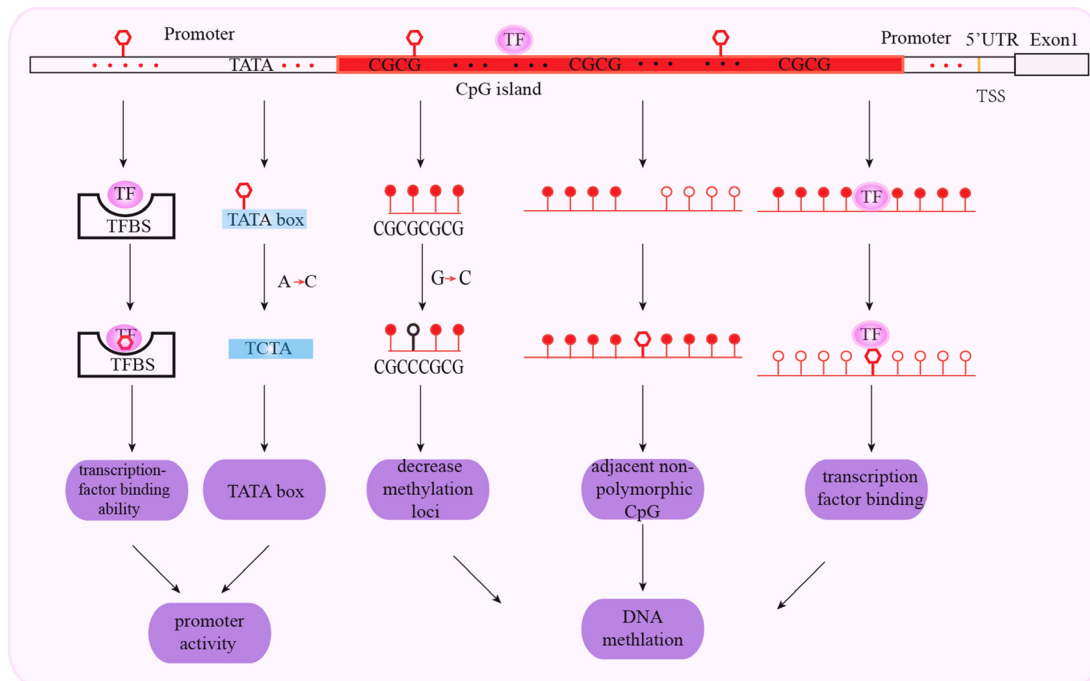
acetylation of H4 histones and methylation of histone H3 in lymphoblastoid cells with GG genotype than cells with AA genotype. Conversely, lymphoblastoid cells with AA genotype showed increased histone H3 acetylation than cells with GG genotype. In unstimulated lymphoblastoid cells, cells with GG genotype showed higher levels of acetylation and methylation of histone H3 than cells with AA genotype, whereas cells with AA genotype showed higher levels of acetylation of histone H4 [55].

## EXONAL SNPs AND CANCER SUSCEPTIBILITY

SNPs in exons are classified as non-synonymous and synonymous coding SNPs (cSNPs) based on their ability to replace the encoded amino acid. Exonal SNPs generally influence cancer susceptibility by genetic mechanisms (Figure 2).

### Non-synonymous cSNPs change protein structure and function

Non-synonymous cSNPs result in amino acid substitution that can affect protein function. Changes in the first two bases of a codon result in amino acid changes in most instances. Alterations in the amino acid sequence can alter the secondary structure of the protein by increasing or decreasing hydrogen bonding and phosphorylation, which affects protein interactions and functions. As a



**Figure 1: Schematic representation of mechanisms associated with promoter SNPs and cancer susceptibility.** SNPs in transcription factor binding sites affect transcription factor binding to the gene promoter. SNPs in the TATA box affect promoter activity with A to C substitutions decreasing the number of the TATA boxes. SNPs in the CpG islands decrease methylation, affecting adjacent non-polymorphic CpG and transcription factor binding. The red triangle represents SNP; red arrows show substitutions of SNPs; red hollow circle represents unmethylated loci; red solid circle represents methylated loci.

consequence, these changes alter cell signaling pathways as well as levels of oncogenic and tumor suppressor proteins. More than 13,000 known SNPs are in exons of various genes, of which 58% are non-synonymous cSNPs [56]. Non-synonymous SNPs influence cancer susceptibility due to changes in the structure and function of the encoded proteins. For example, a non-synonymous cSNP in the epidermal growth factor gene (*EGFR*) eliminates the tyrosine kinase domain (TKD), which is targeted by small-molecule tyrosine kinase inhibitors (TKIs) such as gefitinib and erlotinib. Erlotinib forms a single hydrogen bond with Met 769 in the EGFR, whereas gefitinib forms two hydrogen bonds with Gly 772. Gefitinib shows higher affinity with five polymorphic EGFR proteins than the wild-type EGFR. Polymorphisms in the EGFR-TKD lead to structural changes which increase protein activity and sensitivity to TKIs [57]. The *NAT2* (N-acetyltransferase 2) gene polymorphism rs752955201 results in substitution of valine by a bulkier isoleucine which decreases its affinity for NAT2 substrates and increases its affinity for NAT1 substrates [58].

A non-synonymous cSNP that changes amino acid sequence of the protein-protein interface (PPI) can alter protein interactions, affect stability, and alter post-translational modifications [59]. The Leu858Arg mutation increases the ability of EGFR to form dimers and is associated with cell proliferation [60]. Gain-of-function mutations in tumor suppressor gene p53 promote tumorigenesis [61]. Coincidentally, many

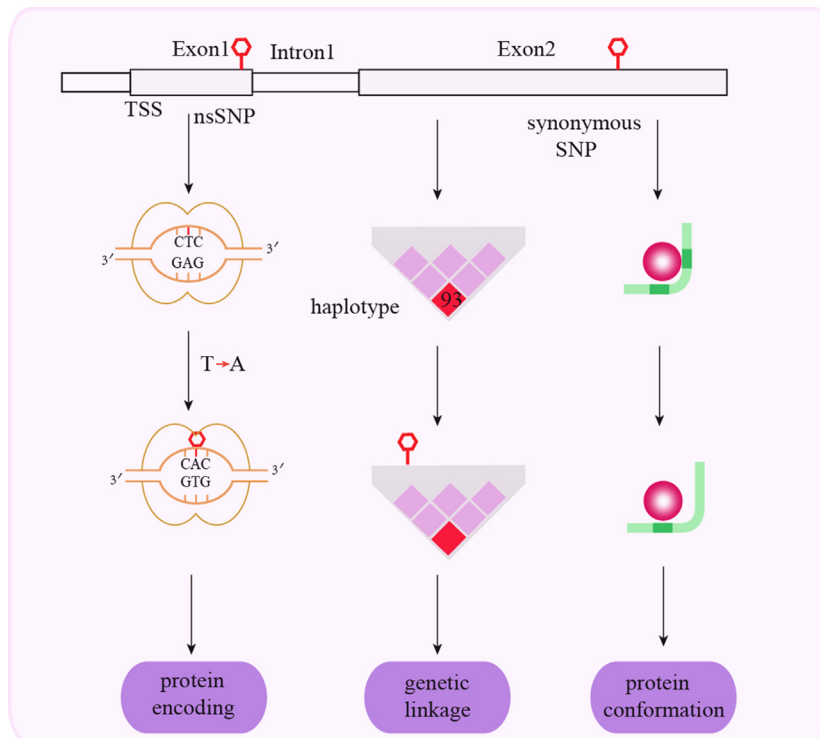
unclassified mutations are found at protein-protein interfaces [62].

### Synonymous cSNPs alter protein structure and function indirectly

Synonymous cSNPs do not alter the amino acid sequence of the encoded protein. In most cases, the nucleotide change is associated with substitution in the third base in a codon. Since the amino acid sequence of these proteins is the same as wild type, they were previously not considered important. However, recent studies show that synonymous cSNPs affect gene function and expression by changing the expression of neighbouring genes. Many studies have demonstrated that synonymous mutations alter the structure, function, and expression of proteins.

### Synonymous cSNPs change secondary structural conformation of mRNAs and proteins

Synonymous cSNPs form different haplotype SNPs that modulate the stability of mRNA secondary structure (local stem-ring structure), thereby reducing enzyme function. For example, two synonymous cSNPs and one non-synonymous cSNP in the Catechol-O-Methyltransferase (*COMT*) gene form different haplotype SNPs. The major *COMT* haplotypes show variations in the local stem-loop structures of messenger RNA. The mRNAs with stable secondary structures are



**Figure 2: Schematic representation of mechanisms associated with exonal SNPs and cancer susceptibility.** Non-synonymous exonal cSNPs change the amino acid sequence of the encoded protein. Synonymous exonal cSNPs change protein conformation and function via genetic linkage.

associated with low COMT expression and activity [63]. A synonymous SNP in multidrug transporter (*MDR1*) gene alters the expression of the critical drug transport protein, P-glycoprotein. This affects its expression and function, thereby impacting drug resistance. SNPs in *MDR1*, namely, 1236C→T, 3435C→T, and 2677G→T SNPs occur in 31-49% of Chinese, Malaysian, and Indian populations [64]. Strong linkage disequilibrium among these three SNPs generates a haplotype which decreases the function of the MDR1 protein with subtle changes in the conformation of its ATP-binding site. The T1236C and C3435T *MDR1* SNPs have the potential to influence non-synonymous SNPs. The C3435T polymorphism affects co-translational folding and insertion of P-glycoprotein into the membrane, thereby altering the structure of substrate-inhibitor interaction sites [65]. The three SNPs represent rare wild-type codons, which are associated with differences in translational rate and termination of translation [66]. The resulting differences in drug-substrate binding represent a possible mechanism by which synonymous SNPs regulate non-synonymous SNPs and result in diversity of clinical responses [67].

Synonymous polymorphisms affect messenger RNA splicing, stability, and structure as well as protein folding. These changes significantly affect function of proteins resulting in changes in cellular response to therapeutic targets, which explains differential responses of individual patients to medications [68]. For example, rs74090726 synonymous polymorphic variation in MCAD exon 5 inactivates the exonal splicing enhancer (ESE) causes exon skipping thus leading to loss of a functional protein and to MCAD deficiency [69]. The synonymous SNPs can lead to diverse diseases by changing microRNA-mediated gene regulation and further altering the gene expression based on bioinformatics analysis [70].

#### **Synonymous cSNPs alter translational rates via genetic linkage**

Synonymous cSNPs can also accelerate or decelerate the speed at which the ribosome moves along the mRNA, thus changing the dynamics of translation, and the subsequent protein structure and function. They may also result in different mRNA secondary structures and protein secondary structures such as  $\alpha$  helix  $\beta$  folding [71]. Moreover, codon usage in synonymous codon families is not random and base-structure preference is associated with the position of the second base in the codon [72]. In some cases, the base-structure preference is related to the third base [73].

## **INTRONAL SNPs AND CANCER SUSCEPTIBILITY**

Introns are involved in regulating tissue-specific gene expression, mRNA transcription, and translation. They contain enhancers or other cis-elements that promote transcription initiation or elongation. Intron splicing

increases mRNA stability in the nucleus. Introns are also involved in alternative splicing and genome imprinting. Functional SNPs in introns are sometimes linked to SNPs of nearby genes, affecting their mRNA splicing and lncRNA binding. This results in variations in the sequence and function of the mature proteins. Thus, intronal SNPs influence the genetic susceptibility to cancer by both genetic and epigenetic mechanisms (Figure 3).

#### **Intronal SNPs affect gene expression by cis-acting regulatory elements**

Intronal sequences contain many cis-acting regulatory elements such as transcription factors, enhancers, silencers, and insulators that positively regulate gene expression and are common “hot spots” in genetic risk variants [74]. The current focus of research is to understand the functional consequences of these loci.

A GWAS identified rs2981578 variant in fibroblast growth factor receptor 2 (FGFR2) as one of the highest ranking risk alleles in breast cancer. Heterozygous rs2981578 clones showed high levels of FOXA1 transcription factor binding to this intronal SNP [75]. The rs12343867 T > C variant in Janus kinase 2 intron 14 is associated with myeloproliferative neoplasms by acting as a transcriptional repressor [76].

Cowper-Salari *et al.* found that breast cancer risk variants targeted enhancers of the FoxA1 and ESR1 transcription factors. The rs4784227 SNP in the *TOX3* gene modulated the affinity of chromatin for FOXA1 in distal regulatory elements resulting in allele-specific gene expression [77]. Three independent variants (rs2981578, rs35054928 and rs45631563) of FGFR2 map to transcriptional silencer elements and augment silencer activity resulting in lower FGFR2 expression and increased estrogen responsiveness and breast cancer risk [78].

#### **Intronal SNPs regulate protein synthesis by mRNA splicing**

The mRNA splicing involves splice donor and acceptor sites, exon splicing enhancers, and splicing proteins. Sequence changes due to non-synonymous or synonymous SNPs modulate mRNA splicing activity resulting in the production of splice variants. For example, SNP resulting from G to A in *DMD* intron 32 deactivates splice donor sites [79]. A polymorphism at a splice donor site in *ZNF419* produces ZAPHIR, an alternatively spliced polymorphic histocompatibility antigen in renal cell carcinoma [80]. An intronal polymorphism in *IRF4* is associated with increased risk of male acute lymphoblastic leukemia, wherein the C to T substitution increases *IRF4* gene expression by transcriptional repression. The wild-type C allele shows two-fold stronger binding to the transcription factor AP-2 $\alpha$  than the variant T allele. AP-2 $\alpha$  plays a negative regulatory role in some tumor cells,

although it is known promote transcription. Wild-type *IRF4* significantly inhibits the *IRF4* promoter, but is reversed by the replacement of C by T [81].

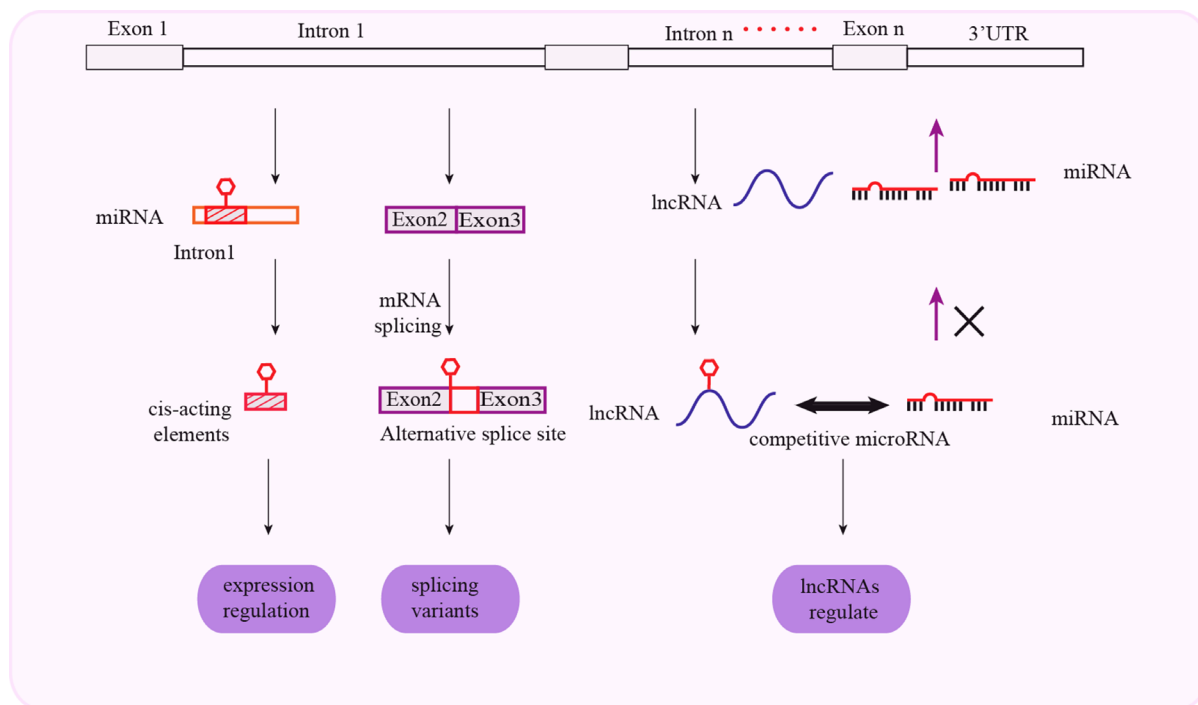
### Intronal SNPs influence genomic imprinting

Genomic imprinting refers to differential expression from maternal and paternal alleles due to variations in DNA methylation and histone acetylation. Polymorphisms in the imprinted regions affect gene expression. *H19* is an imprinted gene that codes for an oncofetal RNA, but is down-regulated postnatally. Heterozygotes for the *H19* rs2839698 TC intronal SNP are protected against bladder cancer, especially non-muscle-invasive bladder cancer. The *H19* RNA inhibits insulin-like growth factor 2 (IGF2) transcription and translation, whereas the *H19* SNP promotes mitotic IGF2 expression, thereby decreasing the risk of bladder cancer [82].

### Intronal SNPs regulate gene expression via lncRNAs

Intronal sequences contain many non-coding RNA motifs, which do not encode proteins but regulate protein expression through epigenetic, transcriptional and post-transcriptional regulation. As an important member of non-coding RNAs, lncRNAs (long non-coding RNA) are not only involved in chromatin remodelling, histone and chromatin modifications, but also participate in transcriptional activation or

repression, alternative splicing, endosome transport and oncogene/tumor suppressor activation or repression. lncRNAs are closely related to malignant tumorigenesis. Polymorphisms in lncRNA-encoding genes affect normal biological functions by competitive microRNAs, thereby influencing susceptibility to cancer. The rs2147578 polymorphism in lncLAMC2-1:1 transcript influences the binding of miR-128-3p and correlates with colorectal cancer (CRC) risk. The CG and GG genotypes of the rs2147578 polymorphism are associated with increased risk for CRC [83]. Numerous genetic polymorphisms in lncRNAs are associated with increased risk in gastric cancer, CRC and prostate cancer [84–86]. SNPs in the lncRNAs *H19*, *HOTAIR* (HOX transcript antisense RNA) and *PRNCR1* correlate with increased gastrointestinal cancer risk [87]. The SNP rs920778 with T allele increases the expression of lncRNA *HOTAIR* via a novel intron enhancer and is associated with ESCC susceptibility [88]. Many SNPs of *PRNCR1* are associated with prostate cancer susceptibility. *PRNCR1* expression is increased in prostate intraepithelial neoplasia and prostate cancer, whereas tumour cell viability and androgen receptor activity is decreased when *PRNCR1* is silenced [89]. The rs10680577 insertion/deletion polymorphism of *EGLN2* results in high expression of both *EGLN2* and RERT-lncRNA and positively correlates with liver cancer risk. It is a potential biomarker for early diagnosis of liver cancer and has been postulated to disrupt the structure of RERT-lncRNA resulting



**Figure 3: Schematic representation of mechanisms associated with intronal SNPs and cancer susceptibility.** Intronal SNPs influence gene expression through cis-acting regulatory elements. Intronal SNPs influence protein synthesis by mRNA splicing and regulation of lncRNA function.

in changes in EGLN2 expression [90]. The NBAT-1 lncRNA inhibits neuroblastoma progression by decreasing cell proliferation and promoting neuronal differentiation [91].

### Intronal SNPs affect chromatin looping

In male-specific childhood ALL, intronal SNP rs12203592, located in intron 4 of the IRF4 gene, enhances physical interaction of the enhancer with the IRF4 promoter through an allele-dependent chromatin loop resulting in higher IRF4 transcriptional rate [92, 93].

## UTR-RELATED SNPs AND CANCER SUSCEPTIBILITY

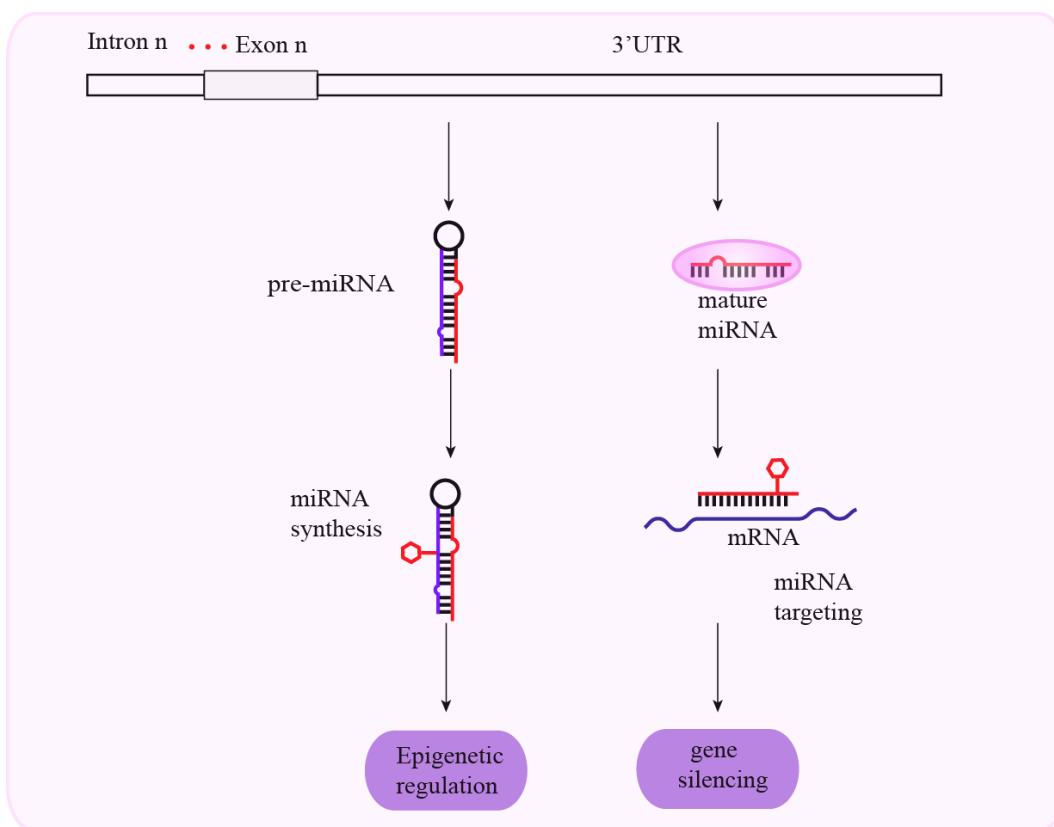
The 5' and 3' UTRs of mRNAs are critical because they control translation [94]. The 5'-UTR regulates translation initiation, whereas the 3'-UTR determines mRNA stability. Specific regulation of mRNA translation is an essential part of gene expression and can be modulated by sequence variations in the 5' and 3' UTRs.

Single nucleotide variations (SNVs) are highly disruptive as they change the secondary structure and miRNA target sites within UTRs [95]. These changes alter the expression of known cancer related genes and signaling pathways. Sequence changes in the UTR regions affect

mRNA folding that impacts transcript stability, mRNA processing and/or translational control. Thus, UTR-SNPs (non-coding SNPs located in the UTR) may have functional consequences on mRNA stability and/or expression [96].

### SNPs in 5'-UTRs affect transcription and protein translation

The 5'-UTR includes genetic elements that regulate gene expression. It starts at the transcription start site and ends at the nucleotide before the start codon. Polymorphisms in 5'-UTRs have been linked to many human diseases because they regulate mRNA processing, transport, stability, and translation. The overall translation rate is influenced by the length of the 5'-UTR, translational start site and its secondary structure, upstream AUGs, upstream open reading frame (ORF) and ribosome binding sites (internal ribosome entry sites, IRESs) [97]. Therefore, polymorphisms or mutations in the 5'UTR can affect translation efficiency. In mammalian genes, ribosomes bind directly to the IRES of 5'-UTR region [97]. In multiple myeloma, a +2756 C to T mutations in the 5'-UTR of the *c-Myc* gene increase the activity of IRES, thereby promoting *c-Myc* expression and tumorigenesis [98]. In the 5'UTR of *CDKN2A* gene, a -34 G to T substitution results in the addition of a new upstream ORF that inhibits translation, which results in loss of allele function and increases melanoma risk [99].



**Figure 4: Schematic representation of mechanisms associated with 3'-UTR SNPs and cancer susceptibility.** SNPs in the 3'-UTR affect miRNA synthesis and gene silencing by altering miRNA-mediated translational repression.

**Table 1: Molecular mechanisms of region-based SNP on cancer susceptibility**

| SNP regions          | Possible molecular mechanism   | Unclear issues  |
|----------------------|--|---|
| Promoter             | Genetic regulation: promoter activity (TATA box, transcription-factor binding ability)<br>Epigenetic regulation: DNA methylation, histone modification | the interaction between genetic and epigenetic elements<br>effect of SNPs on DNA methylation status       |
| Exons                | Non-synonymous cSNPs: coding protein structure and function<br>Synonymous cSNPs: secondary structure conformation translation dynamics                 | detail mechanism at biochemical and cellular level<br>mechanism of kinetics of translation                |
| Introns              | cis-regulatory elements<br>mRNA splicing<br>genomic imprinting<br>lncRNAs<br>chromatin looping   | detail functions of cis-regulatory elements and splicing  |
| UTRs                 | 5'-UTRs: protein translation and transcription activity.<br>3'-UTRs : Regulate mRNA degradation and translation  | how SNPs in the 5'-UTR affect the efficiency of translation<br>how the 3'-UTR affects miRNA binding sites |
| Non definite regions | long range cis regulation<br>tRNA and rRNA   | the ways polymorphisms affect long range cis regulation, tRNA and rRNA                                    |

The 5'-UTR also plays a role in transcription activity. The +24 T/C polymorphism in the 5'-UTR of the *CR2* gene is associated with nasopharyngeal carcinoma risk. Individuals with the minor allele C are at a higher risk for NPC than those carrying the T allele (OR = 1.81) due to increased transcriptional activity [100].

### SNPs in 3'-UTRs regulate mRNA degradation and translation

The 3'-UTRs regulate gene expression through mRNA degradation and translation. The 3'-UTR controls polyadenylation, subcellular localization, translation efficiency, and mRNA degradation. It also determines the fate of specific mRNAs and cell type-specific mRNA expression. Therefore, mutations in the 3'-UTR are involved in many diseases because they affect gene progression. As a regulatory region, the 3'-UTR is indispensable for normal gene expression. Therefore, polymorphisms in the 3'-UTR and miRNAs can alter miRNA binding sites and affect mRNA degradation and protein translation (Figure 4).

### SNPs in 3'-UTRs alter miRNA-mediated translational repression

The microRNAs (miRNAs) inhibit translation and destabilize their target mRNAs by binding to the 3'-UTR of the target transcript [101–103]. SNPs in 3'-UTRs alter target recognition of microRNAs by disrupting sequence complementarity. Some polymorphisms interfere with the function of the miRNA and affect the expression of the miRNA targets [104, 105]. The rs93410170 C > T SNP in the 3'-UTR of the estrogen receptor- $\alpha$  (*ER- $\alpha$* ) gene results in stringent miR-206-mediated regulation of *ER- $\alpha$*  expression and is associated with high risk of breast cancer [106]. Polymorphisms or mutations in the 3' end of target genes affect protein translation even in the absence of changes in mRNA expression. The C allele of rs10889677 A > C polymorphism in the 3'-UTR of *IL23R* is associated with breast cancer, lung cancer, and nasopharyngeal carcinoma in Chinese individuals; the A allele eliminates miR-let-7f binding sites, thereby increasing *IL23R* transcription [107].

SNPs in miRNAs can inhibit miRNA processing and targeting. Mature miRNAs include a 5'-seed region and a 3'-mismatch-tolerant region. The seed sequence is made



up of two to seven nucleotides at the 5' end of the miRNA and is involved in the specific identification of the miRNA targets. Therefore, SNPs in the seed sequence can interfere in the binding of the miRNA to its target mRNA. The rs11614913 T > C SNP in miR-196a2 and rs3746444 A > G SNP in miR-499 are associated with increased breast cancer risk [108]. Polymorphisms or mutations in miRNA alter binding sites, thereby inhibiting gene expression and protein synthesis in the absence of any changes in the encoding mRNA.

#### SNPs in miRNAs

The miRNA polymorphisms or mutations outside the pre-miRNA hairpin-shaped structure and the miRNA seed sequence can affect the synthesis of miRNAs. For example, pri-miRNA polymorphisms such as pri-miR-34 b/c rs4938723 [109], pri-miR-218 rs11134527 [110], and pri-miR-938 rs2505901 [111] are used as biomarkers to predict hepatocellular and gastric cancer risk. A polymorphism in the seed region of miR-125a inhibits the processing of the pri-miR-125 to pre-miR-125a, thereby reducing miR-125a mediated translational repression [112]. A C > T polymorphism in the primary transcript of miR-15a/miR-16 is associated with reduced miR-15 and miR-16 expression in familial chronic lymphocytic leukemia [113, 114].

### SNPs IN UNDEFINED GENETIC REGIONS AND CANCER SUSCEPTIBILITY

#### SNPs in undefined genetic regions modulate gene transcription through long range cis effects

SNPs in non-coding regions regulate their target genes through long-range chromatin interactions. Most of these interactions are located in sites with active histone modifications and transcription factor binding sites. Many candidate genes, such as CAPG, C2orf43, RFX6, NFASC, MYC and AGAP7P and their regulatory variants, including rs1446669, rs699664, rs1078004, rs13394027, rs10993994 and rs4631830 were identified in prostate cancer, thereby showing the role of long-range chromatin interactions [115]. Multiple functional variants in long-range enhancer elements are associated with SNP rs965513, regulated the expression of FOXE1 and PTSC2, and contribute to thyroid cancer risk [30]. Lo *et al.* demonstrated that multiple variants co-operated with the lead SNP and long-range enhancers to transcriptionally regulate FOXE1 and PTSC2 expression [30]. Functional variants at the 11q13 risk locus for breast cancer downregulate cyclin D1 expression through long-range enhancers [116]. A transcriptional enhancer in the G allele of rs554219 increases the risk of breast cancer by reducing the binding of ELK4 transcription factor and correlates with low cyclin D1 levels in breast cancer tissues [116]. The rs6983267 variant in a transcriptional enhancer

on chromosome 8q24 is associated with colorectal cancer pathogenesis because it binds differentially to transcription factor 7-like 2 (TCF7L2) and cMyc [117].

#### SNPs in tRNAs and rRNAs cause transcriptional and translational defects

In some cases, changes in tRNAs and rRNAs are associated with cancer susceptibility. Mutations in the mitochondrial tRNA genes alter the secondary and tertiary tRNA structure, leading to transcriptional and translational defects in the mitochondrial respiratory chain components. The A12308G is a polymorphic mutation in the V-loop of tRNA<sup>Leu</sup> (CUN), which is associated with colorectal cancer susceptibility [118]. Mutations in the mitochondrial tRNA genes are associated with mitochondrial dysfunction in breast cancer [119]. A cis non-coding rRNA (nc-rRNA), upstream of the 45S rRNA transcriptional start site is implicated in altered rRNA transcription in human lung epithelial cancer cells due to changes in its secondary structure [120]. Mitochondrial SNPs (mtSNPs) in the rRNA and tRNAs, MT-ND2 and haplogroup T are associated with CRC risk in European Americans (OR = 1.66, 95% CI: 1.19–2.33, *P* = 0.003) [121].

### CONCLUSIONS AND PERSPECTIVES

In this systematic review, we describe the role of SNPs in different regions of the gene (promoter, exons, introns and UTRs) in cancer susceptibility. SNPs in the promoter region influence promoter activity, transcription factor binding, DNA methylation and histone modifications. SNPs in the exons modulate gene transcription and translation. SNPs in the introns affect RNA splicing, genomic imprinting, and lncRNAs. The SNPs in 5'-UTR promote translation, whereas SNPs in the 3'-UTR modulate miRNA-dependent gene silencing. Moreover, SNPs in other regulatory regions like the enhancers regulate transcription, and translation as well as protein folding (Table 1).

Despite extensive research into the role of SNPs in genetic predisposition of cancer, the mechanisms remain complex. For example, while it's well known that promoter region SNPs affect methylation and histone modifications, the co-operative interaction between genetic and epigenetic mechanisms remains unclear. Moreover, while genetic evidence exists for association between exonal polymorphisms and cancer susceptibility, biochemical and cell biological outcomes of non-synonymous cSNPs are not known. Also, the mechanistic details regarding the effect of synonymous SNPs on the kinetics of translation need detailed investigation. The effect of DNA methylation, which mainly occurs in the promoter region and the first exon, needs to be studied in relation to SNPs. While the functions of intronal polymorphisms are well understood, their effects on splicing, cis-regulatory elements and lncRNAs require further study. Also, the

role of polymorphisms in untranslated regions in gene expression through epigenetic regulation of translation has been recognized. However, the effect of SNPs in the 5'- and 3'UTRs in translation, DNA methylation, miRNA binding, and tissue-specific gene expression needs to be further investigated. Polymorphisms in the pri-, pre-, and mature miRNAs influence selection of target miRNAs and affect expression of the numerous proteins and signaling pathways. The mechanisms related to gene polymorphisms and cancer susceptibility is very complex. Gene polymorphisms also change spatial structure, which affects mRNA stability, methylation and allele-specific expression. These mechanisms require further detailed investigation.

## Author contributions

YY conceived and designed this study and revised the manuscript. ND designed, analyzed, interpreted the literature and wrote the manuscript. HZ and HF collected data and participated in the revision of the manuscript. All authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

This work is supported by grants from the National Science and Technology Support Program (2015BAI13B07) and the National Natural Science Foundation (81772987).

## CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

## REFERENCES

- Devilee P, Rookus MA. A tiny step closer to personalized risk prediction for breast cancer. *N Engl J Med.* 2010; 362:1043–1045.
- Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, Kluth S, Bozic I, Lawrence M, Bottcher S, Carter SL, Cibulskis K, Mertens D, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature.* 2015; 526:525–530.
- Oldridge DA, Wood AC, Weichert-Leahey N, Crimmins I, Sussman R, Winter C, McDaniel LD, Diamond M, Hart LS, Zhu S, Durbin AD, Abraham BJ, Anders L, et al. Genetic predisposition to neuroblastoma mediated by a LMO1 super-enhancer polymorphism. *Nature.* 2015; 528:418–421.
- He C, Tu H, Sun L, Xu Q, Gong Y, Jing J, Dong N, Yuan Y. SNP interactions of *Helicobacter pylori*-related host genes PGC, PTPN11, IL1B, and TLR4 in susceptibility to gastric carcinogenesis. *Oncotarget.* 2015; 6:19017–26. <https://doi.org/10.18632/oncotarget.4231>.
- Xu Q, Liu JW, Yuan Y. Comprehensive assessment of the association between miRNA polymorphisms and gastric cancer risk. *Mutat Res Rev Mutat Res.* 2015; 763:148–160.
- Liu J, He C, Xing C, Yuan Y. Nucleotide excision repair related gene polymorphisms and genetic susceptibility, chemotherapeutic sensitivity and prognosis of gastric cancer. *Mutat Res.* 2014; 765:11–21.
- Zhang Y, Sun LP, Xing CZ, Xu Q, He CY, Li P, Gong YH, Liu YP, Yuan Y. Interaction between GSTP1 Val allele and *H. pylori* infection, smoking and alcohol consumption and risk of gastric cancer among the Chinese population. *PLoS One.* 2012; 7:e47178.
- Jing JJ, Li M, Yuan Y. Toll-like receptor 4 Asp299Gly and Thr399Ile polymorphisms in cancer: a meta-analysis. *Gene.* 2012; 499:237–242.
- Xu Q, Dong Q, He C, Liu W, Sun L, Liu J, Xing C, Li X, Wang B, Yuan Y. A new polymorphism biomarker rs629367 associated with increased risk and poor survival of gastric cancer in Chinese by up-regulated miRNA-let-7a expression. *PLoS One.* 2014; 9:e95249.
- Liu JW, He CY, Sun LP, Xu Q, Xing CZ, Yuan Y. The DNA repair gene ERCC6 rs1917799 polymorphism is associated with gastric cancer risk in Chinese. *Asian Pac J Cancer Prev.* 2013; 14:6103–6108.
- He C, Tu H, Sun L, Xu Q, Li P, Gong Y, Dong N, Yuan Y. *Helicobacter pylori*-related host gene polymorphisms associated with susceptibility of gastric carcinogenesis: a two-stage case-control study in Chinese. *Carcinogenesis.* 2013; 34:1450–1457.
- Ulaganathan VK, Sperl B, Rapp UR, Ullrich A. Germline variant FGFR4 p.G388R exposes a membrane-proximal STAT3 binding site. *Nature.* 2015; 528:570–574.
- Schirmer MA, Luske CM, Roppel S, Schaudinn A, Zimmer C, Pfluger R, Haubrock M, Rapp J, Gungor C, Bockhorn M, Hackert T, Hank T, Strobel O, et al. Relevance of Sp Binding Site Polymorphism in WWOX for Treatment Outcome in Pancreatic Cancer. *J Natl Cancer Inst.* 2016; 108.
- Wu H, Zhang K, Gong P, Qiao F, Wang L, Cui H, Sui X, Gao J, Fan H. A novel functional TagSNP Rs7560488 in the DNMT3A1 promoter is associated with susceptibility to gastric cancer by modulating promoter activity. *PLoS One.* 2014; 9:e92911.
- Fan H, Liu D, Qiu X, Qiao F, Wu Q, Su X, Zhang F, Song Y, Zhao Z, Xie W. A functional polymorphism in the DNA methyltransferase-3A promoter modifies the susceptibility in gastric cancer but not in esophageal carcinoma. *BMC Med.* 2010; 8:12.
- Zhu X, Zhang J, Fan W, Wang F, Yao H, Wang Z, Hou S, Tian Y, Fu W, Xie D, Zhu W, Long J, Wu L, et al. The rs391957 variant cis-regulating oncogene GRP78 expression contributes to the risk of hepatocellular carcinoma. *Carcinogenesis.* 2013; 34:1273–1280.
- Xu M, Gao Y, Yu T, Wang J, Cheng L, Cheng L, Cheng D, Zhu B. Functional promoter rs2295080 T>G variant in MTOR gene is associated with risk of colorectal cancer in a Chinese population. *Biomed Pharmacother.* 2015; 70:28–32.

18. Shao N, Li J, Xu B, Wang Y, Lu X, Feng N. Role of the functional variant (-652T>G) in the XRCC4 promoter in prostate cancer. *Mol Biol Rep.* 2014; 41:7463–7470.
19. Rintisch C, Heinig M, Bauerfeind A, Schafer S, Mieth C, Patone G, Hummel O, Chen W, Cook S, Cuppen E, Colome-Tatche M, Johannes F, Jansen RC, et al. Natural variation of histone modification and its impact on gene expression in the rat genome. *Genome Res.* 2014; 24:942–953.
20. He C, Xu Q, Tu H, Sun L, Gong Y, Liu J, Yuan Y. Polymorphic rs9471643 and rs6458238 upregulate PGC transcription and protein expression in overdominant or dominant models. *Mol Carcinog.* 2016; 55:586–99.
21. Yang L, Yang X, Ji W, Deng J, Qiu F, Yang R, Fang W, Zhang L, Huang D, Xie C, Zhang H, Zhong N, Ran P, et al. Effects of a functional variant c.353T>C in *snai1* on risk of two contextual diseases. Chronic obstructive pulmonary disease and lung cancer. *Am J Respir Crit Care Med.* 2014; 189:139–148.
22. Fang W, Qiu F, Zhang L, Deng J, Zhang H, Yang L, Zhou Y, Lu J. The functional polymorphism of NBS1 p.Glu185Gln is associated with an increased risk of lung cancer in Chinese populations: case-control and a meta-analysis. *Mutat Res.* 2014; 770:61–68.
23. Griseri P, Bourcier C, Hieblot C, Essafi-Benkhadir K, Chamorey E, Touriol C, Pages G. A synonymous polymorphism of the Tristetraprolin (TTP) gene, an AU-rich mRNA-binding protein, affects translation efficiency and response to Herceptin treatment in breast cancer patients. *Hum Mol Genet.* 2011; 20:4556–4568.
24. Pagenstecher C, Wehner M, Friedl W, Rahner N, Aretz S, Friedrichs N, Sengteller M, Henn W, Buettner R, Propping P, Mangold E. Aberrant splicing in MLH1 and MSH2 due to exonic and intronic variants. *Hum Genet.* 2006; 119:9–22.
25. Li L, Sun R, Liang Y, Pan X, Li Z, Bai P, Zeng X, Zhang D, Zhang L, Gao L. Association between polymorphisms in long non-coding RNA PRNCR1 in 8q24 and risk of colorectal cancer. *J Exp Clin Cancer Res.* 2013; 32:104.
26. Xiong HY, Alipanahi B, Lee LJ, Bretschneider H, Merico D, Yuen RK, Hua Y, Gueroussov S, Najafabadi HS, Hughes TR, Morris Q, Barash Y, Krainer AR, et al. RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science.* 2015; 347:1254806.
27. Xu Q, Liu JW, He CY, Sun LP, Gong YH, Jing JJ, Xing CZ, Yuan Y. The interaction effects of pri-let-7a-1 rs10739971 with PGC and ERCC6 gene polymorphisms in gastric cancer and atrophic gastritis. *PLoS One.* 2014; 9:e89203.
28. Xu Q, Chen MY, He CY, Sun LP, Yuan Y. Promoter polymorphisms in trefoil factor 2 and trefoil factor 3 genes and susceptibility to gastric cancer and atrophic gastritis among Chinese population. *Gene.* 2013; 529:104–112.
29. Dunna NR, Naushad SM, Vuree S, Anuradha C, Sailaja K, Surekha D, Rao DR, Vishnupriya S. Association of thymidylate synthase 5'-UTR 28bp tandem repeat and serine hydroxymethyltransferase C1420T polymorphisms with susceptibility to acute leukemia. *Asian Pac J Cancer Prev.* 2014; 15:1719–1723.
30. He H, Li W, Liyanarachchi S, Srinivas M, Wang Y, Akagi K, Wu D, Wang Q, Jin V, Symer DE, Shen R, Phay J, Nagy R, et al. Multiple functional variants in long-range enhancer elements contribute to the risk of SNP rs965513 in thyroid cancer. *Proc Natl Acad Sci U S A.* 2015; 112:6128–6133.
31. Peltoketo H, Piao Y, Mannermaa A, Ponder BA, Isomaa V, Poutanen M, Winqvist R, Vihko R. A point mutation in the putative TATA box, detected in nondiseased individuals and patients with hereditary breast cancer, decreases promoter activity of the 17 beta-hydroxysteroid dehydrogenase type 1 gene 2 (EDH17B2) *in vitro*. *Genomics.* 1994; 23:250–252.
32. Li LC, Chui RM, Sasaki M, Nakajima K, Perinchery G, Au HC, Nojima D, Carroll P, Dahiya R. A single nucleotide polymorphism in the E-cadherin gene promoter alters transcriptional activities. *Cancer Res.* 2000; 60:873–876.
33. Shin Y, Kim IJ, Kang HC, Park JH, Park HR, Park HW, Park MA, Lee JS, Yoon KA, Ku JL, Park JG. The E-cadherin -347G->GA promoter polymorphism and its effect on transcriptional regulation. *Carcinogenesis.* 2004; 25:895–899.
34. Wang M, Chu H, Wang S, Wang M, Wang W, Han S, Zhang Z. Genetic variant in APE1 gene promoter contributes to cervical cancer risk. *Am J Obstet Gynecol.* 2013; 209:360 e361–367.
35. Lo PH, Urabe Y, Kumar V, Tanikawa C, Koike K, Kato N, Miki D, Chayama K, Kubo M, Nakamura Y, Matsuda K. Identification of a functional variant in the MICA promoter which regulates MICA expression and increases HCV-related hepatocellular carcinoma risk. *PLoS One.* 2013; 8:e61279.
36. Zhang X, Miao X, Tan W, Ning B, Liu Z, Hong Y, Song W, Guo Y, Zhang X, Shen Y, Qiang B, Kadlubar FF, Lin D. Identification of functional genetic variants in cyclooxygenase-2 and their association with risk of esophageal cancer. *Gastroenterology.* 2005; 129:565–576.
37. Yang ZH, Zhou CL, Zhu H, Li JH, He CD. A Functional SNP in the MDM2 Promoter Mediates E2F1 Affinity to Modulate Cyclin D1 Expression in Tumor Cell Proliferation. *Asian Pac J Cancer Prev.* 2014; 15:3817–3823.
38. Tower GB, Coon CI, Belguise K, Chabos D, Brinckerhoff CE. Fra-1 targets the AP-1 site/2G single nucleotide polymorphism (ETS site) in the MMP-1 promoter. *Eur J Biochem.* 2003; 270:4216–4225.
39. Yu C, Zhou Y, Miao X, Xiong P, Tan W, Lin D. Functional haplotypes in the promoter of matrix metalloproteinase-2 predict risk of the occurrence and metastasis of esophageal cancer. *Cancer Res.* 2004; 64:7622–7628.
40. De Clerck YA, Darville MI, Eeckhout Y, Rousseau GG. Characterization of the promoter of the gene encoding human tissue inhibitor of metalloproteinases-2 (TIMP-2). *Gene.* 1994; 139:185–191.

41. Boidot R, Vegran F, Jacob D, Chevrier S, Cadouot M, Feron O, Solary E, Lizard-Nacol S. The transcription factor GATA-1 is overexpressed in breast carcinomas and contributes to survivin upregulation via a promoter polymorphism. *Oncogene*. 2010; 29:2577–2584.
42. Rakyan VK, Hildmann T, Novik KL, Lewin J, Tost J, Cox AV, Andrews TD, Howe KL, Otto T, Olek A, Fischer J, Gut IG, Berlin K, et al. DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. *PLoS Biol*. 2004; 2:e405.
43. Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E, Li K, Murty VV, Schupf N, Vilain E, Morris M, Haghghi F, Tycko B. Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. *Nat Genet*. 2008; 40:904–908.
44. Shoemaker R, Deng J, Wang W, Zhang K. Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Res*. 2010; 20:883–889.
45. Stirzaker C, Song JZ, Davidson B, Clark SJ. Transcriptional gene silencing promotes DNA hypermethylation through a sequential change in chromatin modifications in cancer cells. *Cancer Res*. 2004; 64:3871–3877.
46. Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W, Craig DW, Redman M, Gershon ES, Liu C. Genetic control of individual differences in gene-specific methylation in human brain. *Am J Hum Genet*. 2010; 86:411–419.
47. Zhang S, Lu J, Zhao X, Wu W, Wang H, Lu J, Wu Q, Chen X, Fan W, Chen H, Wang F, Hu Z, Jin L, et al. A variant in the CHEK2 promoter at a methylation site relieves transcriptional repression and confers reduced risk of lung cancer. *Carcinogenesis*. 2010; 31:1251–1258.
48. Su KJ, Lin CW, Chen MK, Yang SF, Yu YL. Effects of EZH2 promoter polymorphisms and methylation status on oral squamous cell carcinoma susceptibility and pathology. *Am J Cancer Res*. 2015; 5:3475–3484.
49. Tycko B. Allele-specific DNA methylation: beyond imprinting. *Hum Mol Genet*. 2010; 19:R210–220.
50. Butcher DT, Mancini-DiNardo DN, Archer TK, Rodenhiser DI. DNA binding sites for putative methylation boundaries in the unmethylated region of the BRCA1 promoter. *Int J Cancer*. 2004; 111:669–678.
51. Boumber YA, Kondo Y, Chen X, Shen L, Guo Y, Tellez C, Estecio MR, Ahmed S, Issa JP. An Sp1/Sp3 binding polymorphism confers methylation protection. *PLoS Genet*. 2008; 4:e1000162.
52. Ogino S, Hazra A, Tranah GJ, Kirkner GJ, Kawasaki T, Nosho K, Ohnishi M, Suemoto Y, Meyerhardt JA, Hunter DJ, Fuchs CS. MGMT germline polymorphism is associated with somatic MGMT promoter methylation and gene silencing in colorectal cancer. *Carcinogenesis*. 2007; 28:1985–1990.
53. McVicker G, van de Geijn B, Degner JF, Cain CE, Banovich NE, Raj A, Lewellen N, Myrthil M, Gilad Y, Pritchard JK. Identification of genetic variants that affect histone modifications in human cells. *Science*. 2013; 342:747–749.
54. Larsson L, Rymo L, Berglundh T. Sp1 binds to the G allele of the -1087 polymorphism in the IL-10 promoter and promotes IL-10 mRNA transcription and protein production. *Genes Immun*. 2010; 11:181–187.
55. Larsson L, Thorbert-Mros S, Rymo L, Berglundh T. Influence of epigenetic modifications of the interleukin-10 promoter on IL10 gene expression. *Eur J Oral Sci*. 2012; 120:14–20.
56. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, Kang HM, Jordan D, Leal SM, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science*. 2012; 337:64–69.
57. Raghav D, Sharma V, Agarwal SM. Structural investigation of deleterious non-synonymous SNPs of EGFR gene. *Interdiscip Sci*. 2013; 5:60–68.
58. Tsrka T, Boukouvala S, Agianian B, Fakis G. Polymorphism p.Val231Ile alters substrate selectivity of drug-metabolizing arylamine N-acetyltransferase 2 (NAT2) isoenzyme of rhesus macaque and human. *Gene*. 2014; 536:65–73.
59. Yates CM, Sternberg MJ. The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on protein-protein interactions. *J Mol Biol*. 2013; 425:3949–3963.
60. Shan Y, Eastwood MP, Zhang X, Kim ET, Arkhipov A, Dror RO, Jumper J, Kuriyan J, Shaw DE. Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell*. 2012; 149:860–870.
61. van Oijen MG, Slootweg PJ. Gain-of-function mutations in the tumor suppressor gene p53. *Clin Cancer Res*. 2000; 6:2138–2145.
62. David A, Razali R, Wass MN, Sternberg MJ. Protein-protein interaction sites are hot spots for disease-associated nonsynonymous SNPs. *Hum Mutat*. 2012; 33:359–363.
63. Nackley AG, Shabalina SA, Tchivileva IE, Satterfield K, Korchymskiy O, Makarov SS, Maixner W, Diatchenko L. Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science*. 2006; 314:1930–1933.
64. Koehn J, Fountoulakis M, Krapfenbauer K. Multiple drug resistance associated with function of ABC-transporters in diabetes mellitus: molecular mechanism and clinical relevance. *Infect Disord Drug Targets*. 2008; 8:109–118.
65. Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM. A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science*. 2007; 315:525–528.
66. Sauna ZE, Ambudkar SV. About a switch: how P-glycoprotein (ABCB1) harnesses the energy of ATP binding and hydrolysis to do mechanical work. *Mol Cancer Ther*. 2007; 6:13–23.
67. Lubelski J, Konings WN, Driessen AJ. Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria. *Microbiol Mol Biol Rev*. 2007; 71:463–476.

68. Hunt R, Sauna ZE, Ambudkar SV, Gottesman MM, Kimchi-Sarfaty C. Silent (synonymous) SNPs: should we care about them? *Methods Mol Biol.* 2009; 578: 23–39.
69. Nielsen KB, Sorensen S, Cartegni L, Corydon TJ, Doktor TK, Schroeder LD, Reinert LS, Elpeleg O, Krainer AR, Gregersen N, Kjems J, Andresen BS. Seemingly neutral polymorphic variants may confer immunity to splicing-inactivating mutations: a synonymous SNP in exon 5 of MCAD protects from deleterious mutations in a flanking exonic splicing enhancer. *Am J Hum Genet.* 2007; 80:416–432.
70. Wang Y, Qiu C, Cui Q. A Large-Scale Analysis of the Relationship of Synonymous SNPs Changing MicroRNA Regulation with Functionality and Disease. *Int J Mol Sci.* 2015; 16:23545–23555.
71. Shen LX, Basilion JP, Stanton VP Jr. Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proc Natl Acad Sci U S A.* 1999; 96:7871–7876.
72. Gupta SK, Majumdar S, Bhattacharya TK, Ghosh TC. Studies on the relationships between the synonymous codon usage and protein secondary structural units. *Biochem Biophys Res Commun.* 2000; 269:692–696.
73. Adzhubei AA, Adzhubei IA, Krashennnikov IA, Neidle S. Non-random usage of ‘degenerate’ codons is related to protein three-dimensional structure. *FEBS Lett.* 1996; 399:78–82.
74. Zhang X, Bailey SD, Lupien M. Laying a solid foundation for Manhattan--‘setting the functional basis for the post-GWAS era’. *Trends Genet.* 2014; 30:140–149.
75. Robbez-Masson LJ, Bodor C, Jones JL, Hurst HC, Fitzgibbon J, Hart IR, Grose RP. Functional analysis of a breast cancer-associated FGFR2 single nucleotide polymorphism using zinc finger mediated genome editing. *PLoS One.* 2013; 8:e78839.
76. Spasovski V, Tomic N, Nikcevic G, Stojiljkovic M, Zukic B, Radmilovic M, Karan-Djurasevic T, Srzentic S, Colovic M, Pavlovic S. The influence of novel transcriptional regulatory element in intron 14 on the expression of Janus kinase 2 gene in myeloproliferative neoplasms. *J Appl Genet.* 2013; 54:21–26.
77. Cowper-Salari R, Zhang X, Wright JB, Bailey SD, Cole MD, Eeckhoutte J, Moore JH, Lupien M. Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. *Nat Genet.* 2012; 44:1191–1198.
78. Campbell TM, Castro MAA, de Santiago I, Fletcher MNC, Halim S, Prathalingam R, Ponder BAJ, Meyer KB. FGFR2 risk SNPs confer breast cancer risk by augmenting oestrogen responsiveness. *Carcinogenesis.* 2016; 37:741–750.
79. Thi Tran HT, Takeshima Y, Surono A, Yagi M, Wada H, Matsuo M. A G-to-A transition at the fifth position of intron-32 of the dystrophin gene inactivates a splice-donor site both *in vivo* and *in vitro*. *Mol Genet Metab.* 2005; 85:213–219.
80. Broen K, Levenga H, Vos J, van Bergen K, Fredrix H, Greupink-Draaisma A, Kester M, Falkenburg JH, de Mulder P, de Witte T, Griffioen M, Dolstra H. A polymorphism in the splice donor site of ZNF419 results in the novel renal cell carcinoma-associated minor histocompatibility antigen ZAPHIR. *PLoS One.* 2011; 6:e21699.
81. Do TN, Ucisik-Akkaya E, Davis CF, Morrison BA, Dorak MT. An intronic polymorphism of IRF4 gene influences gene transcription *in vitro* and shows a risk association with childhood acute lymphoblastic leukemia in males. *Biochim Biophys Acta.* 2010; 1802:292–300.
82. Verhaegh GW, Verkleij L, Vermeulen SH, den Heijer M, Witjes JA, Kiemeny LA. Polymorphisms in the H19 gene and the risk of bladder cancer. *Eur Urol.* 2008; 54:1118–1126.
83. Gong J, Tian J, Lou J, Ke J, Li L, Li J, Yang Y, Gong Y, Zhu Y, Zhang Y, Zhong R, Chang J, Miao X. A functional polymorphism in lnc-LAMC2-1:1 confers risk of colorectal cancer by affecting miRNA binding. *Carcinogenesis.* 2016; 37:443–451.
84. Li L, Jia F, Bai P, Liang Y, Sun R, Yuan F, Zhang L, Gao L. Association between polymorphisms in long non-coding RNA PRNCR1 in 8q24 and risk of gastric cancer. *Tumour Biol.* 2015; 37:299–303.
85. Chu H, Xia L, Qiu X, Gu D, Zhu L, Jin J, Hui G, Hua Q, Du M, Tong N, Chen J, Zhang Z, Wang M. Genetic variants in noncoding PIWI-interacting RNA and colorectal cancer risk. *Cancer.* 2015; 121:2044–2052.
86. Cao DL, Gu CY, Zhu Y, Dai B, Zhang HL, Shi GH, Shen YJ, Zhu YP, Ma CG, Xiao WJ, Qin XJ, Lin GW, Ye DW. Polymorphisms at long non-coding RNAs and prostate cancer risk in an eastern Chinese population. *Prostate Cancer Prostatic Dis.* 2014; 17:315–319.
87. Lv Z, Xu Q, Yuan Y. A systematic review and meta-analysis of the association between long non-coding RNA polymorphisms and cancer risk. *Mutat Res.* 2017; 771:1–14.
88. Zhang X, Zhou L, Fu G, Sun F, Shi J, Wei J, Lu C, Zhou C, Yuan Q, Yang M. The identification of an ESCC susceptibility SNP rs920778 that regulates the expression of lncRNA HOTAIR via a novel intronic enhancer. *Carcinogenesis.* 2014; 35:2062–2067.
89. Chung S, Nakagawa H, Uemura M, Piao L, Ashikawa K, Hosono N, Takata R, Akamatsu S, Kawaguchi T, Morizono T, Tsunoda T, Daigo Y, Matsuda K, et al. Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. *Cancer Sci.* 2011; 102:245–252.
90. Zhu Z, Gao X, He Y, Zhao H, Yu Q, Jiang D, Zhang P, Ma X, Huang H, Dong D, Wan J, Gu Z, Jiang X, et al. An insertion/deletion polymorphism within RERT-lncRNA modulates hepatocellular carcinoma risk. *Cancer Res.* 2012; 72:6163–6172.
91. Pandey GK, Mitra S, Subhash S, Hertwig F, Kanduri M, Mishra K, Fransson S, Ganeshram A, Mondal T, Bandaru S, Ostensson M, Akyurek LM, Abrahamsson J, et al. The risk-associated long noncoding RNA NBAT-1 controls neuroblastoma progression by regulating cell proliferation and neuronal differentiation. *Cancer Cell.* 2014; 26:722–737.
92. Wright JB, Brown SJ, Cole MD. Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-

- associated single-nucleotide polymorphism in colorectal cancer cells. *Mol Cell Biol*. 2010; 30:1411–1420.
93. Visser M, Palstra RJ, Kayser M. Allele-specific transcriptional regulation of IRF4 in melanocytes is mediated by chromatin looping of the intronic rs12203592 enhancer to the IRF4 promoter. *Hum Mol Genet*. 2015; 24:2649–2661.
  94. Wilkie GS, Dickson KS, Gray NK. Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. *Trends Biochem Sci*. 2003; 28:182–188.
  95. Sabarinathan R, Wenzel A, Novotny P, Tang X, Kalari KR, Gorodkin J. Transcriptome-wide analysis of UTRs in non-small cell lung cancer reveals cancer-related genes with SNV-induced changes on RNA secondary structure and miRNA target sites. *PLoS One*. 2014; 9:e82699.
  96. Aouacheria A, Navratil V, Lopez-Perez R, Gutierrez NC, Churkin A, Barash D, Mouchiroud D, Gautier C. In silico whole-genome screening for cancer-related single-nucleotide polymorphisms located in human mRNA untranslated regions. *BMC Genomics*. 2007; 8:2.
  97. Mendell JT, Dietz HC. When the message goes awry:disease-producing mutations that influence mRNA content and performance. *Cell*. 2001; 107:411–414.
  98. Chappell SA, LeQuesne JP, Paulin FE, deSchoolmeester ML, Stoneley M, Soutar RL, Ralston SH, Helfrich MH, Willis AE. A mutation in the c-myc-IRES leads to enhanced internal ribosome entry in multiple myeloma: a novel mechanism of oncogene de-regulation. *Oncogene*. 2000; 19:4437–4440.
  99. Cazzola M, Skoda RC. Translational pathophysiology:a novel molecular mechanism of human disease. *Blood*. 2000; 95:3280–3288.
  100. Fan Q, He JF, Wang QR, Cai HB, Sun XG, Zhou XX, Qin HD, Shugart YY, Jia WH. Functional polymorphism in the 5'-UTR of CR2 is associated with susceptibility to nasopharyngeal carcinoma. *Oncol Rep*. 2013; 30:11–16.
  101. Selbach M, Schwanhaussner B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. Widespread changes in protein synthesis induced by microRNAs. *Nature*. 2008; 455:58–63.
  102. Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP. The impact of microRNAs on protein output. *Nature*. 2008; 455:64–71.
  103. Mishra PJ, Humeniuk R, Mishra PJ, Longo-Sorbello GS, Banerjee D, Bertino JR. A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. *Proc Natl Acad Sci U S A*. 2007; 104:13513–13518.
  104. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. *Nat Genet*. 2007; 39:1278–1284.
  105. Barnes MR, Deharo S, Grocock RJ, Brown JR, Sanseau P. The micro RNA target paradigm: a fundamental and polymorphic control layer of cellular expression. *Expert Opin Biol Ther*. 2007; 7: 1387–1399.
  106. Brendle A, Lei H, Brandt A, Johansson R, Enquist K, Henriksson R, Hemminki K, Lenner P, Forsti A. Polymorphisms in predicted microRNA-binding sites in integrin genes and breast cancer: ITGB4 as prognostic marker. *Carcinogenesis*. 2008; 29:1394–1399.
  107. Zheng J, Jiang L, Zhang L, Yang L, Deng J, You Y, Li N, Wu H, Li W, Lu J, Zhou Y. Functional genetic variations in the IL-23 receptor gene are associated with risk of breast, lung and nasopharyngeal cancer in Chinese populations. *Carcinogenesis*. 2012; 33:2409–2416.
  108. Hu Z, Liang J, Wang Z, Tian T, Zhou X, Chen J, Miao R, Wang Y, Wang X, Shen H. Common genetic variants in pre-microRNAs were associated with increased risk of breast cancer in Chinese women. *Hum Mutat*. 2009; 30:79–84.
  109. Xu Y, Liu L, Liu J, Zhang Y, Zhu J, Chen J, Liu S, Liu Z, Shi H, Shen H, Hu Z. A potentially functional polymorphism in the promoter region of miR-34b/c is associated with an increased risk for primary hepatocellular carcinoma. *Int J Cancer*. 2011; 128:412–417.
  110. Zhang LS, Liang WB, Gao LB, Li HY, Li LJ, Chen PY, Liu Y, Chen TY, Han JG, Wei YG, Zhang L. Association between pri-miR-218 polymorphism and risk of hepatocellular carcinoma in a Han Chinese population. *DNA Cell Biol*. 2012; 31:761–765.
  111. Arisawa T, Tahara T, Shiroeda H, Matsue Y, Minato T, Nomura T, Yamada H, Hayashi R, Saito T, Matsunaga K, Fukuyama T, Hayashi N, Otsuka T, et al. Genetic polymorphisms of IL17A and pri-microRNA-938, targeting IL17A 3'-UTR, influence susceptibility to gastric cancer. *Hum Immunol*. 2012; 73:747–752.
  112. Duan R, Pak C, Jin P. Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. *Hum Mol Genet*. 2007; 16:1124–1131.
  113. Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M, Iuliano R, Palumbo T, Pichiorri F, et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med*. 2005; 353:1793–1801.
  114. Raveche ES, Salerno E, Scaglione BJ, Manohar V, Abbasi F, Lin YC, Fredrickson T, Landgraf P, Ramachandra S, Huppi K, Toro JR, Zenger VE, Metcalf RA, et al. Abnormal microRNA-16 locus with synteny to human 13q14 linked to CLL in NZB mice. *Blood*. 2007; 109:5079–5086.
  115. Du M, Tillmans L, Gao J, Gao P, Yuan T, Dittmar RL, Song W, Yang Y, Sahr N, Wang T, Wei GH, Thibodeau SN, Wang L. Chromatin interactions and candidate genes at ten prostate cancer risk loci. *Sci Rep*. 2016; 6:23202.
  116. French JD, Ghousaini M, Edwards SL, Meyer KB, Michailidou K, Ahmed S, Khan S, Maranian MJ, O'Reilly M, Hillman KM, Betts JA, Carroll T, Bailey PJ, et al, and GENICA Network, and kConFab Investigators. Functional variants at the 11q13 risk locus for breast cancer regulate cyclin D1 expression through long-range enhancers. *Am J Hum Genet*. 2013; 92:489–503.

117. Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, Beckwith CA, Chan JA, Hills A, Davis M, Yao K, Kehoe SM, Lenz HJ, et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat Genet.* 2009; 41:882–884.
118. Mohammed F, Rezaee Khorasany AR, Mosaieby E, Houshmand M. Mitochondrial A12308G alteration in tRNA(Leu(CUN)) in colorectal cancer samples. *Diagn Pathol.* 2015; 10:115.
119. Grzybowska-Szatowska L, Slaska B. Polymorphisms in genes encoding mt-tRNA in female breast cancer in Poland. *Mitochondrial DNA.* 2012; 23:106–111.
120. Shiao YH, Lupascu ST, Gu YD, Kasprzak W, Hwang CJ, Fields JR, Leighty RM, Quinones O, Shapiro BA, Alvord WG, Anderson LM. An intergenic non-coding rRNA correlated with expression of the rRNA and frequency of an rRNA single nucleotide polymorphism in lung cancer cells. *PLoS One.* 2009; 4:e7505.
121. Li Y, Beckman KB, Caberto C, Kazma R, Lum-Jones A, Haiman CA, Le Marchand L, Stram DO, Saxena R, Cheng I. Association of Genes, Pathways, and Haplogroups of the Mitochondrial Genome with the Risk of Colorectal Cancer: The Multiethnic Cohort. *PLoS One.* 2015; 10: e0136796.