Review

SNP-Target Genes Interaction Perturbing the Cancer Risk in the Post-GWAS

Wenmin Yang [1,2,3,†], Te Zhang [1,2,3,†], Xuming Song [1,2,3], Gaochao Dong [1,2,*], Lin Xu [1,2,3,4] and Feng Jiang [1,2,3,*]

1   Department of Thoracic Surgery, Nanjing Medical University Affiliated Cancer Hospital and Jiangsu Cancer Hospital and Jiangsu Institute of Cancer Research, Nanjing 210009, China
2   Jiangsu Key Laboratory of Molecular and Translational Cancer Research, Cancer Institute of Jiangsu Province, Nanjing 210009, China
3   The Fourth Clinical College, Nanjing Medical University, Nanjing 210009, China
4   Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing 211116, China
*   Correspondence: ilsyvm@njmu.edu.cn (G.D.); fengjiang_nj@njmu.edu.cn (F.J.)
†   These authors contributed equally to this work.

Simple Summary: Genome-wide association studies have identified a vast number of cancer risk-associated loci harboring numerous single nucleotide polymorphisms that regulate gene expression and affect individual genetic susceptibility to cancer through different routes. Recently, there has been some progress made regarding the molecular and biological mechanisms underlying the ways that genetic variation affects gene regulation. This review summarizes the molecular and biological mechanisms of genetic variation that affect gene regulation by drawing from the findings provided by past studies.

Abstract: Cancer ranks as the second leading cause of death worldwide, and, being a genetic disease, it is highly heritable. Over the past few decades, genome-wide association studies (GWAS) have identified many risk-associated loci harboring hundreds of single nucleotide polymorphisms (SNPs). Some of these cancer-associated SNPs have been revealed as causal, and the functional characterization of the mechanisms underlying the cancer risk association has been illuminated in some instances. In this review, based on the different positions of SNPs and their modes of action, we discuss the mechanisms underlying how SNPs regulate the expression of target genes to consequently affect tumorigenesis and the development of cancer.

1. Introduction

Cancer is a primary cause of mortality and a major impediment to improving life expectancy in every country on the planet [1,2]. According to the latest cancer burden data released by the World Health Organization's International Agency for Research on Cancer (IARC), 19.3 million new cancers were diagnosed worldwide in 2020, resulting in nearly 10 million deaths, and it is predicted that over the next two decades, as populations age, unhealthy behaviors increase, and other factors influence, the number of cancer cases worldwide could increase by 60% [1]. In low or middle-income countries, the increase could be as high as 81%. In 2020, breast cancer surpassed lung cancer to become the most commonly diagnosed cancer, with an estimated 2.3 million new cases (11.7%); however, lung cancer still continues to be a major cause of mortality, accounting for an estimated 1.8 million deaths (18%) [1]. Modifiable risk factors account for more than 40% of cancer incidence and mortality in China, including tobacco exposure, obesity, unhealthy lifestyle and alcohol consumption. [3]. In addition, epidemiological research strongly suggests

that the etiology of common malignancies is hereditary. Elevated cancer incidence in affected families is attributable to a complex etiology, whereby several low-penetrance cancer predisposition genes interact with environmental variables to affect cancer risk [4,5].

In recent years, genome-wide association studies (GWAS) have emerged as an effective means to determine genetic contributions to cancer risk. GWAS provide an agnostic approach to the identification of genetic variations influencing cancer risk. GWAS research, as opposed to typical candidate-driven studies, consider the full genomic sequence obtained from a large sample of cancer patients, and then compare the data to those of a healthy population, thereby allowing for a more objective approach to identifying genetic aberrances in cancer. GWAS have been performed for most types of cancers, and hundreds of risk alleles have consequently been identified, most of which are common and individually confer a modest increase in cancer risk [6].

Single nucleotide polymorphism (SNP) is a DNA sequence polymorphism caused by a single nucleotide variation at the genomic level, which refers to a series of single nucleotide changes in DNA [7]. SNP is a variation in which more than 1% of a population does not carry the same nucleotide at a specific position in the DNA sequence. Furthermore, SNP is non-pathogenic, or not directly responsible [7,8]. However, mutation is an irreversible sequence variation in DNA, which in essence includes all variations that occur in the human genome either spontaneously or non-spontaneously [9]. The frequency of mutations is relatively low, generally less than 1%, and most mutations are pathogenic. The distribution of SNPs and mutations is region-specific and race-specific.

Multiple GWAS have been conducted during the past decade for each of the main types of malignancies, such as prostate [10–12], breast [13–15], colorectal [16–19], lung [20–24], gastric [25–28], renal [29–31], and ovarian [32,33] cancers. GWAS for thyroid cancer [34], glioma [35,35], and malignant melanoma [36–38] have also been reported. Currently, more than 430 cancer associations in 262 distinct genomic regions have been identified by GWAS, and increasingly more loci are being identified over time. However, due to the existence of linkage disequilibrium and the complex interactions that exist between genes and the environment, tag SNPs identified by GWAS are not necessarily true pathogenic variants. Therefore, it is important to interpret GWAS results with greater scrutiny. Freedman and Edwards et al. raised the point that post-GWAS research strategies aimed at screening functional sites and elucidating their underlying molecular mechanisms [39,40]. Up until now, there have been numerous studies conducted on the functional resolution of cancer risk SNPs. Herein, to better understand the role of functional SNPs in tumorigenesis and cancer development, we propose several working classifications for function SNPs based on region, and elaborate the mechanisms that tumor risk-associated SNP complex regulated target genes.

2. Functional Mechanisms of Coding Region SNPs

SNPs located in coding regions can be divided into two types: synonymous and non-synonymous mutations. Although synonymous mutations do not affect the amino acid sequence of the protein, they may change the expression of the protein by affecting post-transcriptional modifications, translation rates, and other processes. In contrast, non-synonymous SNPs (nsSNPs) cause the substitution of amino acids, thereby resulting in changes to the protein structure, its physical and chemical properties (stability, solubility, etc.), and its function. At present, there are many biological software packages (such as SIFT(Sorting Intolerant From Tolerant), F-SNP(the Functional Single Nucleotide Polymorphism) and PolyPhen) that can be used to predict the effect of nsSNPs on protein structure and function [41–43]. Compared to SNPs located in gene non-coding regions, the functional mechanism underlying tumor-associated nsSNPs is relatively simple [44]. Combined with whole exon analysis, several coding region SNPs have been identified to be associated with colorectal cancer development. For example, the missense mutation rs3184504 (p. trp263ARg) located in a domain of SH2B3 may change the function of the protein in the context of regulating cell division. Other coding variants may also affect

variable shear (RS16888728, UTP23) [45]. The mechanism by which SNPs located within the coding regions of genes affect the risk of disease is inseparable from the function of the resulting coded proteins.

Some risk loci exert an effect on the amino acid sequence of the produced protein. Examples include BRCA2 p.Lys3326Ter (rs11571833), and CHEK2 p.Ile157Thr (rs17879961) in lung [46] and breast [47] cancers. The mechanistic interpretation of such variants is presumed to be relatively simple. In addition to the aforesaid, coding SNPs can affect RNA processing; an example is rs78378222 in the 3′ untranslated region of TP53, whereby the risk-corrected variation alters the sequence AATAAA to AATACA, thereby changing the polyadenylation signal of TP53, and ultimately resulting in the impaired 3′-end processing of TP53 mRNA [48,49]. Some variants can also affect splicing. Tian and his colleagues identified a single-nucleotide variation in the *ELP2* gene that affect *ELP2* exon pre-mRNA splicing through splicing a quantitative trait locus (sQTL) [50].

Researchers have often focused on specific signaling pathways, genes, and genetic modifications of interest, while also performing whole-exon association analysis (GWAS) to find any relevant coding SNPs with large effects on these molecules and processes. For example, Li and colleagues used exon sequencing and conducted an association analysis of 12 important genes involved in TGF-β signaling to find that low-frequency causative variation in the TGF-β pathway contributes to colorectal carcinoma (CRC) susceptibility. They discovered that the missense variation rs3764482 (c. 83C>T; p. S28F) located in the gene *SMAD7* was consistently and strongly associated with CRC risk. The rs3764482 allele T was more effective compared to the dominant allele C in limiting TGF-β signaling and reducing the phosphorylation of receptor-regulated SMADs (R-SMADs) via impeding the activation of downstream genes, thereby promoting cancer cell proliferation and contributing to CRC pathogenesis [51].

Coding SNPs may also affect gene and protein modifications. The N6-methyladenosine (m6A) modification is critical for ensuring messenger RNA stability and is involved in many biological activities, including pre-mRNA splicing, 3′-end processing, nuclear export, translation regulation, mRNA degradation, and the DNA damage response [52,53]. The m6A methylation modification occurs in the messenger RNA(mRNA) and can be formed by methylation "writers" and removed by demethylation "erasers" [54]. Rs8100241, located in the gene *ANKLE1*, was identified to be associated with susceptibility to both CRC and breast cancer. The presence of the rs8100241 risk allele A (Figure 1a) combined with the m6A "writer" complex (comprised of the proteins METTL3, METTL14, and WTAP) and the m6A "reader" protein (YTHDF1) was found to increase the levels of the m6A modification on the gene *ANKLE1* and consequently increase its protein expression. Mechanistically, ANKLE1 functions as a potential tumor suppressor by decreasing CRC cell proliferation while maintaining genomic integrity, thereby contributing to a lower risk of CRC [55].

Notably, coding SNPs may interact with other SNPs to produce a stronger functional role [45]. The rs138649767 A allele (Figure 1b) located in the exon region of *TCF7L2* can activate the *MYC* enhancer containing rs6983267 allele G to promote the expression of MYC [56]. SNPs occurring in the exons and introns of *SMAD7* may affect its regulation and jointly affect downstream signaling pathways involving SMAD7 and TGFβ [51]. As a result, while examining SNPs in coding regions, the interactions between them should be taken into account to better understand their functional processes.
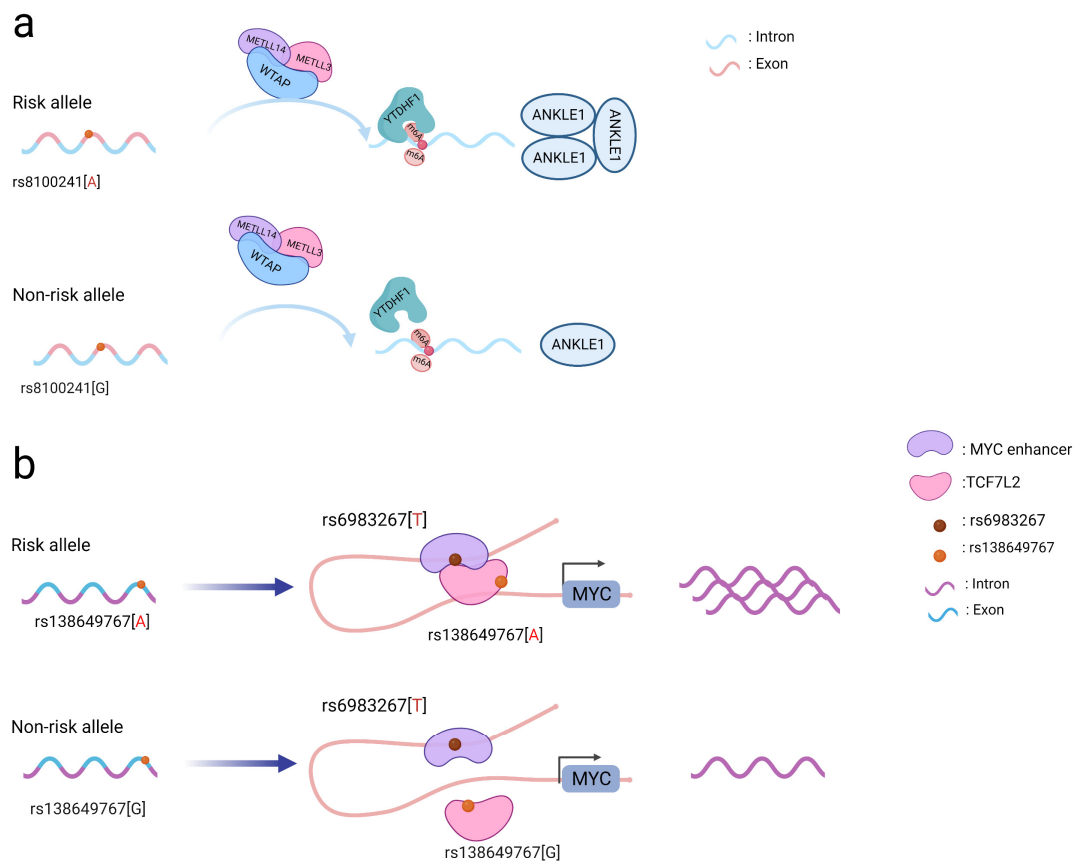
Figure 1. Schematic diagram of the action mechanism employed by coding SNPs. (**a**) The A allele of the rs8100241 variant, which is found in the *ANKLE1* second exon region, has been linked to a lower risk of CRC by increasing *ANKLE1* mRNA m6A levels and thus facilitating ANKLE1 protein expression, thereby potentially functioning as a negative regulator to hinder cell growth by maintaining genomic stability. (**b**) Interaction between the *TCF7L2* missense variant rs138649767 and a regulatory variant rs6983267 in the *MYC* enhancer and promoter on the expression of MYC.

3. Functional Mechanisms of Non-Coding Region SNPs

Accumulating evidence shows that a SNP in non-coding regions is the most common type of genetic variation in the human genome, accounting for 90% of inter-individual variation [6,57]. Depending on the location, the region can harbor a response element that is either proximal (promoter, enhancer, or super-enhancer) or distal (intergenic or intra-genic). The risk loci identified by GWAS were located in the genomic regions of cell type-specific active chromatin, and most of them were quantitative trait loci, methylation quantitative trait loci and transcription factor (TF) binding related loci. Chromatin conformational studies have helped to link regulatory regions localized by SNPs to their respective target genes [6,58,59]. These loci may be involved in gene transcription, post-transcriptional processing, translation, post-translational modifications, and other processes to regulate gene expression. Many target genes have been identified using expression quantitative trait loci (eQTL) to detect the relationship between SNPs and gene expression. Non-coding SNPs can regulate the transcription of target genes by sequence-proximal (*cis*)- or distal (*trans*)-interactions. Studies have found that histone modifications in the regions of such risk SNPs are particularly abundant, especially those related to promoter and enhancer activities (H3K4me3, H3K4me1, H3K27ac). Most SNPs are predicted to destroy the binding motifs of specific transcription factors. For example, rs6983267 may change the binding of transcription factors such as MYC, CTCF, and TCF7L2 [16]. In addition to affecting gene transcription levels by altering transcription factor-binding sites (TFBS), non-coding SNPs also change epigenetic modifications and/or the chromatin structure to influence

target gene expression. Through the above method, non-coding SNP participates in cell proliferation, apoptosis, migration, and invasion.

3.1. Genetic Variants That Alter Promoters

A promoter is a sequence of DNA that is recognized, bound, and serves to initiate transcription by RNA polymerase. Promoters contain variations of a conserved sequence required for the specific binding of RNA polymerase and transcription initiation. Most promoters are located upstream of the transcription initiation point of structural genes, and the promoter itself is not transcribed [60]. Promoters are located upstream of the 5' end of a given structural gene, and they activate RNA polymerase to bind accurately to the template DNA with specificity for inducing the initiation of transcription [60]. Promoters do not control gene activity themselves; rather, gene activity is regulated by binding to proteins called transcription factors (TF). SNPs within promoter regions generally play a regulatory role by influencing the binding of such transcription factors. A recently reported example is that of the SNP rs13278062 located in the promoter of death receptor 4 (DR4) which confers an altered risk of colorectal cancer. The study revealed that the rs13278062 G>T variant changed the binding affinity of the transcription factor Sp1/NF1, increased the expression of DR4, and thus suppressed carcinogenesis and metastasis of colorectal cancer [61]. The MPO promoter SNP rs2333227 increases the malignant characteristics of colorectal cancer by changing the promoter's affinity to AP-2α [62]. The variant SNP rs10993994 located in the upstream promoter of the gene MSMB is also found to be overrepresented in individuals with prostate cancer; this is attributed to stronger CREB binding and thus increased promoter activity [63]. Furthermore, the SNP rs11672691 is a risk locus associated with prostate cancer that is related to the lncRNA PCAT19. The non-risk variant rs11672691 and its linkage disequilibrium (LD) SNP rs887391 are more likely to bind the TFs NKX3.1 and YY1 to the PCAT19-short promoter, thereby leading to increased promoter but lower enhancer activity, which then activates PCAT19-short, and ultimately results in lower prostate cancer susceptibility [64]. SNPs in promoter regions of multiple genes, including TERT, KLHDC7A, PIDD1, and ESR1, have been discovered in breast cancer by GWAS, with reporter studies revealing that independent risk alleles change target promoter activity [65,66]. Most of the reported promoter changes exert their regulatory effects by altering TF binding. The SNP rs3824662 allele A (Figure 2a) increases chromatin accessibility by changing the TF GATA3 expression, promoting the binding of GATA3 with the CRLF promoter, and ultimately forming a chromatin loop [67].

3.2. Genetic Variants That Alter Enhancers

Enhancers are regions of DNA sequence that can increase the *cis*-acting transcription of their target gene sequences. Enhancers each differ in their distance from their target promoter(s); in mammalian species, an enhancer can be 100 bp to Mb away from their target gene [68]. Enhancers, unlike promoters, can be found anywhere in a gene; they can be positioned either upstream or downstream of their target genes, or even within another gene's gene body, and enhancer regulation can circumvent other genes irrespective of their orientation. Enhancers must bind to specific protein factors to enhance the transcription of their target. Enhancers generally have tissue or cell specificity, whereby they only show activity in certain cells or tissues, which is determined by the specific protein factors present in these cells or tissues [69]. Enhancers are typically recognized by the epigenetic marks H3K4me1 and H3K27ac, which are present in active enhancer elements. Conversely, H3K27me3 is regarded as a silent epigenetic mark associated with lower enhancer activity [70,71]. GWAS-identified risk loci for common illnesses are often found in non-coding areas, and many of these are thought to function as enhancers [72]. According to emerging data, these SNPs may influence gene regulation by changing the binding of important TFs to critical transcriptional enhancers [73].
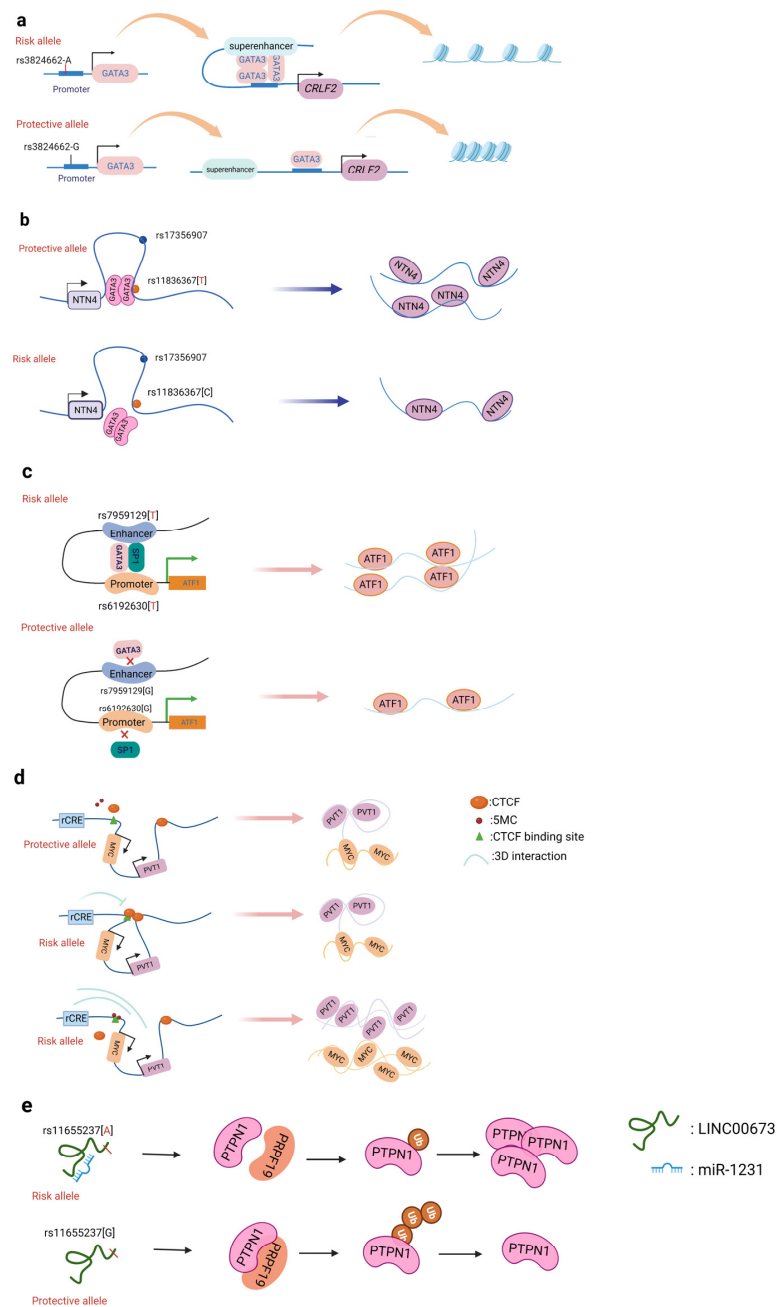
Figure 2. Schematic diagram of the action mechanism employed by non-coding SNPs. (**a**) The SNP rs3824662 allele A increases chromatin accessibility by inducing GATA3 expression, promoting the binding of GATA3 with the CRLF promoter, and ultimately forming a chromatin loop. (**b**) The NTN4 enhancer risk variant rs11836367 binds to the TF GATA3 to regulate NTN4 expression, ultimately promoting breast carcinoma initiation and progression. (**c**) Enhancer SNP rs7959129 risk allele G interacts with promoter SNP rs6192603 risk allele G contributing to ATF1 expression by binding TFs GATA3 and SP1. (**d**) The risk allele rs11986220 and higher methylation at −10 Kb synergistically function to confer a greater risk of tumor; however, when −20 Kb is hypomethylated, the function of the risk SNP is inhibited by the enhancer-blocking insulator loop mediated by CTCF. (**e**) The risk variant rs11655237 in LINC00673 creates a miR-1231–binding site that interferes with the expression of LINC00673 and contributes to pancreatic cancer susceptibility.

3.2.1. Breast Cancer

Of all cancers, breast cancer has so far yielded the greatest number of discovered risk loci [13]. Understanding the driving mechanism(s) underlying malignant transformation

provides the prospect of combating cancer recurrence and treatment resistance. Zhang et al. identified that the SNP rs4971059 resides in the sixth intron and within an active enhancer element of the TRIM46 gene. By using CRISPR/Cas9-mediated homologous recombination, they constructed the SNP rs4971059 with the allele G converted to allele A, thereby resulting in TRIM46 overexpression, boosting breast carcinoma cell growth, enhancing chemotherapy resistance in vitro, and hastening tumor development in vivo [74]. In addition, Yang and colleagues (Figure 2b) reported the noncoding regulatory variant rs11836367 at the NTN4 locus (12q22) and identified it to be associated with the risk of breast carcinoma as a causal variant. The rs11837367 protective T allele promotes GATA3 binding to the distal enhancer and increases NTN4 expression [75].

3.2.2. Prostate Cancer

Several studies have independently identified several genes in specific prostate cancer (PCa) susceptibility loci that are either controlled by causative SNPs containing a *cis*-regulatory element (CRE) or have been indicated as SNP-associated genes [76]. SNP rs339331 at 6q22 was found to be a prostate cancer risk-associated variant. The risk allele T of rs339331 has been found to augment the enhancer-binding of HOXB13, alter the level of the RFX6 protein in an allele-specific manner, and confer a predisposition to prostate cancer [77]. Recently, Huang et al. also identified that the PCa-associated rs11672691 located within an enhancer element can change the binding site of HOXA2, which in turn promotes oncogenesis by impacting the expression of nearby genes [78].

Notably, there are other cases of SNPs causing DNA-binding polymorphisms in distinct transcription factors. For example, a gastric cancer risk-associated polymorphism (rs2978980 T>G) that is situated in an intronic enhancer of *lncPSCA* has been found to disrupt the binding of the transcription factor RORA, thereby resulting in lower *lncPSCA* expression in an allele-specific manner [79]. As another example, the rs2647046 enhancer has been found to interact with the *HLA-DQB1-AS1* promoter to alter its expression via a CTCF-mediated long-range loop in an allele-specific manner, thereby conferring susceptibility to hepatocellular cancer (HCC) [80]. Another variation on chromosome 11q13.3 in a distant intergenic region has been characterized as a susceptibility locus for renal cell cancer. To control transcription, the 11q13.3 locus encodes a long-range enhancer that physically connects with the *CCDN1* promoter [81]. Interestingly, SNP sites can act as promoters and enhancers simultaneously, and their conversion is determined by the background genotype. As a result, one gene can produce several different RNAs that are involved in the development of diseases. The SNP rs11672691 mediates promoter and enhancer switching under different genotypes. A risk-associated sequence in the *PCAT19*-long enhancer interacts with the *PCAT19*-long promoter to enhance prostate cancer development through activating cell cycle genes [64].

3.2.3. Colorectal Cancer

GWAS have identified numerous colorectal cancer risk loci, but only a fraction of the target genes of these loci have been systematically interrogated. For example, Yu et al. identified a common SNP (rs7198799) in the intron of the gene *CDH1*. They demonstrated that the risk allele C of rs7198799 acts as an enhancer that can target the TF NFATC2 and remotely enhance ZFP90 expression [82]. A prominent mechanism by which SNP variants can affect cell-specific enhancer function is via altered TF binding, thus regulating the target gene's expression. Tian et al. identified two risk SNPs (rs61926301 and rs79591129) located in the ATF1 promoter and first intron, respectively. These are enriched in enhancer regions and open chromatin, which are also associated with H3K4me1, H3K27ac, and ATAC-seq peaks. The two variants increase the expression of ATF1 through preferentially binding to the two TFs SP1 and GATA3 [83]. Rs174575 can act as a specific remote enhancer of FADS2 and *lncRNA-AP002754.2* with the participation of the transcription factor E2F1. Interestingly, TF E2F1 can promote the expression of FADS2, form a chromatin loop, and affect the occurrence of colorectal cancer [84].

3.3. Genetic Variants That Affect Promoter–Enhancer Interactions

Promoter–enhancer interactions (PEIs) underlie differential transcriptional regulation. Several technologies (chromosome conformation capture (3C), Hi-c, and H3K27Ac-HiCHIP) allow for the study of long-range *cis*-regulation [85–87]. Promoter–enhancer interactions are essential events involved in the current theory of transcriptional control. So far, there is little evidence that PEIs are required for the transcriptional control of an enhancer's target gene. The insertion or deletion of promoters, the absence of certain PEI-associated proteins, and the inclusion of PEI-disrupting insulators all have an effect on the expression of target genes. Tian et al. found two risk variants (rs1926301 and rs7959129) located in the *ATF1* promoter and intron, respectively; the former binds the TF SP1 while the latter binds the TF GATA3 (Figure 2c). They found that these two risk sites increase the interaction between the promoter and enhancer by binding SP1 and GATA3, facilitating ATF1 expression, and conferring hereditary susceptibility to CRC [83]. Moreover, the SNP rs11672691 mediates promoter and enhancer switching in a manner dependent on different background genotypes. The risk is determined by the PCAT19-long enhancer interacting with the PCAT19-long promoter, thereby altering prostate cancer development through activating cell cycle genes [64].

3.4. Genetic Variants That Alter 3D Genome Architecture

Within the nucleus, genomic DNA folds into a three-dimensional structure organized at different levels by the formation of chromatin rings. These structures can bring distant enhancers near their target promoters to affect gene expression and regulation. The chromosomes fold into chromatin characterized by sequence-regulating spatial interactions that are key to maintaining normal cell status and function. In cancer genomes, structural variation typically results in changes to the genome's 3D structure and, as a result, alterations in genome-mediated transcriptional control [88]. Changes in the three-dimensional genome architecture or high-order chromatin structure are linked to the development and progression of several diseases [89,90]. Long-distance chromatin looping regulates cancer susceptibility genes either actively or passively. Enhancers frequently form long-range chromatin loops with their target gene promoter regions to affect gene expression. The 9q22 locus, for example, contains the thyroid cancer risk-related SNP rs965513, which demarcates a 33-kb linkage disequilibrium block (including the lead SNP rs965513) that is strongly linked with PTC risk. The chromatin characteristics and regulatory element signatures of this block indicate at least three regulatory elements that operate as enhancers. Using chromosomal conformation capture technology, researchers have observed the long-range looping connections of these elements with the promoter region shared by FOXE1 and PTCSC2 in a human papillary thyroid cancer cell line (KTC-1) and unaffected thyroid tissue [91]. Similarly, Zhang et al. discovered that the rs1859962 risk-associated LD block contains a PCa-specific enhancer that forms a 1-Mb chromatin loop with the *SOX9* gene. This study found that the rs1859962 PCa risk LD block contacts *SOX9* via a long-distance chromatin loop that connects it to the E1 enhancer [92].

CTCF is a transcription factor that promotes long-range chromosomal contact via looping. Hoffman et al. discovered that one allele in the Igf2/H19 imprinting control region (ICR) on chromosome 7 colocalized with one allele of Wsb1/Nf1 on chromosome 11. The lack of CTCF or the ablation of the maternal ICR was found to eliminate this connection and alter the expression of the *Wsb1/Nf1* gene [93]. This finding confirmed the importance of CTCF in the control of the shape of chromatin and the resulting gene expression. On the other hand, the unique contribution of CTCF is that of an insulator. Insulators are short nucleotide sequences that determine the boundaries of genomic areas that are close to one another [94]. When CTCF binds to an insulator region, it inhibits gene transcription by interfering with the communication between an enhancer and a gene promoter [95]. Ahmed M. et al. identified (Figure 2d) noncoding *cis*-regulatory elements (rCRE) by performing CRISPRi screens. They discovered that the 8q24.21 area is widely marked with H3K27ac and has a significant binding affinity to AR, FOXA1, and HOXB13, all of which are important

transcription regulators for PCa pathogenesis [96]. Using an integrated approach involving ChIP, Hi-C, CRISPR, and functional rescue, researchers also discovered that the rs11986220 containing the rCRE sequence interacts with the *MYC* promoter in V16A cells but not in 22Rv1 cells, as the promoter–CRE interaction is typically facilitated by a CTCF site in a 10 kb region upstream, which prevents chromatin looping [96]. Similarly, the rs6702619 region is inhabited by CTCF, which acts as an insulator with long-range physical interactions with CRC-relevant loci [97]. Understanding CTCF-mediated 3D genomic architecture will aid in understanding the mechanism of action underlying noncoding GWAS SNPs at either CTCF sites or regulatory enhancer sites [98].

3.5. Genetic Variants That Influence the Binding of miRNA

MicroRNAs (miRNAs) are noncoding RNA molecules that influence gene expression via regulating messenger RNA degradation and translation. MicroRNAs are normally excised by the RNase iii enzyme Dicer from 60–110 nucleotide long hairpin precursor (folded) RNA structures (pre-miRNAs), which are then integrated into the RNA-induced silencing complex (RISC). The pro-miRNA sequence is transcribed by Pol-II [99]. Accumulating evidence suggests that miRNAs play a key role in carcinogenesis by binding to the 3′-UTR of target mRNAs [100]. MiRNA mutations or their misexpression have been associated with human malignancies and alterations in cancer-associated gene expression [101]. Hoffman et al. detected a variant (rs11614913) in *has-miR-196a-2* using GWAS to screen genetic variants in 15 miRNAs. This SNP was identified to be associated with decreased breast cancer risk [102]. Previous research has confirmed that the methylation of [103] islands in miRNA regions may change miRNA function, thereby influencing carcinogenic pathways. The author and his colleagues found that a CpG island in the region upstream of the miRNA precursor is associated with breast cancer risk [102]. The *ATF1* rs11169571 variant was shown to be strongly related to ATF1 expression by influencing *hsa-miR-1283* and *hsa-miR-520d-5p* binding, which may increase susceptibility to colorectal cancer [56]. In addition, SNPs located in the 3′UTR region of *MDM4*, *CD44*, *LAMC1*, and other genes exert a similar mechanism [104–106].

Some SNPs within long non-coding RNA can also change their binding affinity to miRNAs. The variant loci rs1317082, discovered at exon 1 of lncRNA *RP11-362K14.5* (CC-Slnc362), establishes a binding site for *miR-4658*, which consequently reduces CCSlnc362 expression and confers lowered susceptibility to CRC [107]. The link between rs140618127 in the lncRNA *LOC146880* with non-small cell lung cancer involves a *miR-539-5p* binding site. The combination of *miR-539-5p* and *LOC146880* has been found to result in the reduced activation of the oncogene *ENO1*. Reduced *ENO1* phosphorylation also results in lower PI3K and Akt activation, which is linked to decreased cell proliferation and tumor formation [108]. Moreover, the SNP rs11655237 allele G in LINC00673 exon can create a miRNA binding site that increases the function of LINC00667 expression (Figure 2e). Furthermore, rs67311347 in RCC [109], rs12982687 in CRC [110], and rs16854802 in neck squamous cell carcinoma (HNSCC) [111] are SNPs in lncRNA sequences that affect target gene expression by binding with miRNA. If a SNP occurs within miRNA, it will consequently affect the binding affinity of the miRNA to target genes.

4. Bioinformatic Methods for SNP Functional Analysis

It has become abundantly obvious that SNPs perform different functions as a result of improvements in the mapping of SNPs and cataloging of data from numerous meta-analysis studies conducted in various parts of the world [112,113]. Numerous SNPs have been found to play a direct or indirect role in modifying the susceptibility to a number of diseases, including cancer [114]. Bioinformatics related algorithms were used to predict the functional of risk variants and the possible signaling pathways involved, thus contributing to identify potential genes and loci. In the post-genomic era, the SNP data are developing at a high speed, and they enable the functions analysis of the coding region SNPs and non-coding region SNPs by using a computational method. Currently, there is a large number

of biological software and databases that can be used to annotate and predict the function of SNPs and annotate them for further study. For example, biological software packages such as Annovar, SnpEff, VEP (Variant Effect Predictor), and Oncotator can be used to annotate the influence of nsSNPs on protein structure and SNP target genes [115–118]. The programs of SIFT(Sorting Intolerant From Tolerant), F-SNP(the Functional Single Nucleotide Polymorphism) and PolyPhen that predict whether the non-synonymous SNPs have adverse effects on protein structure or function, perturb the translation of protein and post-transcriptional modification [41–43]. These softwares help us to explain the molecular mechanism further.

Besides the annotation software, there are plenty of Human SNP/mutation databases to understand the SNP information. For example, the HGVbase (Human Genome Variation Database), the dbSNP (Database of Single Nucleotide Polymorphisms), the HGMD (Human Gene Mutation Database), the UCSC browser, and the Ensembl genome browser can provide information about DNA sequences near SNP sites, allele frequencies in populations, as well as provide abundant biology annotations. In addition to the above, researchers can integrate transcriptomic and epigenomic data to annotate relevant loci, screen potential causal loci and further annotate candidate loci. For example, databases such as ENCODE and Roadmap provide information on methylation, histone modifications, chromatin opening, etc; eQTL data can help identify target genes that may be affected by SNPs; databases such as Cistrome and JASPAR can be used to predict whether SNPs affect transcription factor binding. These softwares and databases contribute to our understanding and explain the SNPs' pathogenic mechanism.

5. The Variables of Tumor Risk

Cancer risk is decided by multiple factors such as environmental exposure and heritable genetics. In the past decades, GWAS provided plentiful germline genetic variations for various cancers. This research has revealed the genetic mechanism of cancer hereditary susceptibility. However, identifying a variant is not sufficient to make an association with the tumor risk. Besides genetic risk factors (family history and high-penetrance genes, genetic polymorphisms), adiposity, poor diet, tobacco exposure, alcohol consumption, occupational exposures and air pollution are all closely related to tumorigenesis (Table 1). For example, tobacco exposure is an important carcinogenic factor, many organs in the human body are the tobacco-related cancer sites, such as lung and bronchus, esophagus, stomach, liver, pancreas, cervix and ovaries [119,120]. There is no risk factor known to have as great an effect on the number of deaths as tobacco exposure. Heavy alcohol consumption is a known risk factor for many tumors including hepatocellular carcinoma, colorectal cancer, breast cancer and pancreatic cancer [121–124]. Genetic constitution and environmental exposure are both indispensable causal factors when assessing cancer risk.

Table 1. Risk factors of environmental exposure, lifestyle factors.

Cancer-Risk Factors	Cancer Site	References
Environmental Exposure	-	
Alcohol	Oral cavity, oropharynx, and hypopharynx; larynx; esophagus; liver; colon and rectum; and female breast	[121–127]
Smoking	Lung and bronchus, pancreas, stomach, liver, myeloid leukemia	[120,128–130]
Ionizing Radiation	Chronic lymphocytic leukemia, Hodgkin lymphoma, malignant melanoma, uterine cervical cancer, testicular cancer, and rectal cancer	[120,131–133]
Solar Radiation	Cutaneous squamous cell carcinoma; cutaneous and ocular melanoma and basal cell carcinoma	[120,134,135]
Obesity	Stomach; colorectal; breast; endometrial; kidney	[136–139]
Infectious Agents	-	
HPV	Uterine cervix, oropharyngeal, anogenital	[140–142]
HBV/HCV	Liver	[143,144]
EBV	Lymphoid tissues; Epithelial tissues; nasopharynx	[120,145–147]
Helicobacter pylori	Stomach; B-cell MALT lymphoma	[148–150]
Occupational Exposure	Urinary bladder, bone marrow (leukemia), lung, pleura/peritoneum (mesothelioma), nasal sinus, and liver (angiosarcoma)	[151–154]

6. Conclusions and Future Directions

GWAS have identified a large number of genetic variants affecting the risk of cancer. How to find real SNPs associated with a tumor, and from the biological interpretations of its function and its relationship with the tumor is one of the great challenges of the post-GWAS era. Post-genome-wide association study analysis, such as sequencing technologies, data interpretation, and function studies have enhanced the capacity to unscramble the potential biological role of danger loci in cancer. This review highlights the molecular mechanism of the transcriptional dysregulation of genes expression, and chromatin structure change by tumor-related genetic variants.

Despite the fact that GWAS identified large numbers of disease-associated SNPs, GWAS and post-GWAS analysis can only reveal a fraction of the genetic factors associated with complex diseases. Currently, most of the SNPs are linked to other SNPs or with particular genetic regions, which makes it incredibly challenging to precisely describe the impact of a particular SNP on a specific phenotype when other related SNPs may also impact that characteristic, through linkage disequilibrium and other potential mechanisms. GWAS and post-GWAS mostly focus on SNP location and the relation between cancer development and progression and genetic variants, and lose sight of the interaction between environmental factors and genes. Because of the genes–genes, genes–environment interplay, and the existence of other variants of low frequency, rare variant, genetic deficiencies are evidently in the GWAS researches. It is necessary to establish a complementary strategy (consider genetic epigenetic and environmental factor simultaneously) that is efficient and affordable to make an association with the cancer risk and explain the biological function of risk loci, SNP-mediated signaling pathways and regulatory mechanisms. With the technological advances, a larger population research becomes possible in the future. More genetic variants and SNP-mediated biological functions will be found. They can provide more susceptibility loci and a base to explore cancer development and progression.

Author Contributions: Writing—original draft preparation, W.Y.; writing—review and editing, T.Z., X.S.; project administration, G.D.; funding acquisition, F.J., L.X. All authors have read and agreed to the published version of the manuscript.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef] [PubMed]
2. CP, W.; E, W.; BW, S. World Cancer Report: Cancer Research for Cancer Prevention. ISBN 978-92-832-0447-3.
3. Zheng, R.; Zhang, S.; Zeng, H.; Wang, S.; Sun, K.; Chen, R.; Li, L.; Wei, W.; He, J. Cancer incidence and mortality in China. *J. Natl. Cancer Cent.* **2022**, *2*, 1–9. [CrossRef]
4. Turnbull, C.; Hodgson, S. Genetic predisposition to cancer. *Clin. Med.* **2005**, *5*, 491–498. [CrossRef] [PubMed]
5. Lichtenstein, P.; Holm, N.V.; Verkasalo, P.K.; Iliadou, A.; Kaprio, J.; Koskenvuo, M.; Pukkala, E.; Skytthe, A.; Hemminki, K. 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. [CrossRef]
6. Sud, A.; Kinnersley, B.; Houlston, R. Genome-wide association studies of cancer: Current insights and future perspectives. *Nat. Cancer* **2017**, *17*, 692–704. [CrossRef]
7. Brookes, A.J. The essence of SNPs. *Gene* **1999**, *234*, 177–186. [CrossRef]
8. Shastry, B.S. SNPs: Impact on Gene Function and Phenotype. In *Single Nucleotide Polymorphisms: Methods and Protocols*; Komar, A.A., Ed.; Humana Press: Totowa, NJ, USA, 2009; pp. 3–22.

9. Brown, T.A. *Mutation, Repair and Recombination*; Wiley-Liss: Hoboken, NJ, USA, 2002.

10. Eeles, R.A.; The COGS–Cancer Research UK GWAS–ELLIPSE (part of GAME-ON) Initiative; Olama, A.A.A.; Benlloch, S.; Saunders, E.J.; Leongamornlert, D.; Tymrakiewicz, M.; Ghoussaini, M.; Luccarini, C.; Dennis, J.; et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. *Nat. Genet.* **2013**, *45*, 385–391. [CrossRef]

11. Amundadottir, L.; Sulem, P.; Gudmundsson, J.; Helgason, A.; Baker, A.; Agnarsson, B.A.; Sigurdsson, A.; Benediktsdottir, K.R.; Cazier, J.-B.; Sainz, J.; et al. A common variant associated with prostate cancer in European and African populations. *Nat. Genet.* **2006**, *38*, 652–658. [CrossRef]

12. Eeles, A.R.; The UK Genetic Prostate Cancer Study Collaborators; Kote-Jarai, Z.; Giles, G.; Al Olama, A.A.; Guy, M.; Jugurnauth, S.K.; Mulholland, S.; Leongamornlert, D.; Edwards, S.M.; et al. Multiple newly identified loci associated with prostate cancer susceptibility. *Nat. Genet.* **2008**, *40*, 316–321. [CrossRef]

13. Michailidou, K.; Beesley, J.; Lindstrom, S.; Canisius, S.; Dennis, J.; Lush, M.J.; Maranian, M.J.; Bolla, M.K.; Wang, Q.; Shah, M.; et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat. Genet.* **2015**, *47*, 373–380. [CrossRef]

14. Garcia-Closas, M.; Couch, F.J.; Lindstrom, S.; Michailidou, K.; Schmidt, M.K.; Brook, M.N.; Orr, N.; Rhie, S.K.; Riboli, E.; Buring, J.E.; et al. Genome-wide association studies identify four ER negative–specific breast cancer risk loci. *Nat. Genet.* **2013**, *45*, 392–398. [CrossRef]

15. Couch, F.J.; Kuchenbaecker, K.B.; Michailidou, K.; Mendoza-Fandino, G.A.; Nord, S.; Lilyquist, J.; Olswold, C.; Hallberg, E.; Agata, S.; Ahsan, H.; et al. Identification of four novel susceptibility loci for oestrogen receptor negative breast cancer. *Nat. Commun.* **2016**, *7*, 11375. [CrossRef] [PubMed]

16. Law, P.J.; The PRACTICAL consortium; Timofeeva, M.; Fernandez-Rozadilla, C.; Broderick, P.; Studd, J.; Fernandez-Tajes, J.; Farrington, S.; Svinti, V.; Palles, C.; et al. Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat. Commun.* **2019**, *10*, 2154. [CrossRef] [PubMed]

17. Zeng, C.; Matsuda, K.; Jia, W.-H.; Chang, J.; Kweon, S.-S.; Xiang, Y.-B.; Shin, A.; Jee, S.H.; Kim, D.-H.; Zhang, B.; et al. Identification of Susceptibility Loci and Genes for Colorectal Cancer Risk. *Gastroenterology* **2016**, *150*, 1633–1645. [CrossRef] [PubMed]

18. Lu, Y.; Kweon, S.-S.; Tanikawa, C.; Jia, W.-H.; Xiang, Y.-B.; Cai, Q.; Zeng, C.; Schmit, S.L.; Shin, A.; Matsuo, K.; et al. Large-Scale Genome-Wide Association Study of East Asians Identifies Loci Associated with Risk for Colorectal Cancer. *Gastroenterology* **2019**, *156*, 1455–1466. [CrossRef]

19. Schmit, S.L.; Edlund, C.K.; Schumacher, F.R.; Gong, J.; Harrison, T.A.; Huyghe, J.R.; Qu, C.; Melas, M.; Van Den Berg, D.J.; Wang, H.; et al. Novel Common Genetic Susceptibility Loci for Colorectal Cancer. *J. Natl. Cancer Inst.* **2019**, *111*, 146–157. [CrossRef]

20. Hu, Z.; Wu, C.; Shi, Y.; Guo, H.; Zhao, X.; Yin, Z.; Yang, L.; Dai, J.; Hu, L.; Tan, W.; et al. A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. *Nat. Genet.* **2011**, *43*, 792–796. [CrossRef]

21. Shiraishi, K.; Kunitoh, H.; Daigo, Y.; Takahashi, A.; Goto, K.; Sakamoto, H.; Ohnami, S.; Shimada, Y.; Ashikawa, K.; Saito, A.; et al. A genome-wide association study identifies two new susceptibility loci for lung adenocarcinoma in the Japanese population. *Nat. Genet.* **2012**, *44*, 900–903. [CrossRef]

22. Dong, J.; Hu, Z.; Wu, C.; Guo, H.; Zhou, B.; Lv, J.; Lu, D.; Chen, K.; Shi, Y.; Chu, M.; et al. Association analyses identify multiple new lung cancer susceptibility loci and their interactions with smoking in the Chinese population. *Nat. Genet.* **2012**, *44*, 895–899. [CrossRef]

23. Lan, Q.; Hsiung, C.A.; Matsuo, K.; Hong, Y.-C.; Seow, A.; Wang, Z.; Hosgood, H.D.; Chen, K.; Wang, J.-C.; Chatterjee, N.; et al. Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. *Nat. Genet.* **2012**, *44*, 1330–1335. [CrossRef]

24. Amos, C.I.; Wu, X.; Broderick, P.; Gorlov, I.P.; Gu, J.; Eisen, T.; Dong, Q.; Zhang, Q.; Gu, X.; Vijayakrishnan, J.; et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25. *Nat. Genet.* **2008**, *40*, 616–622. [CrossRef] [PubMed]

25. Shi, Y.; Hu, Z.; Wu, C.; Dai, J.; Li, H.; Dong, J.; Wang, M.; Miao, X.; Zhou, Y.; Lu, F.; et al. A genome-wide association study identifies new susceptibility loci for non-cardia gastric cancer at 3q13.31 and 5p13. *Nat. Genet.* **2011**, *43*, 1215–1218. [CrossRef] [PubMed]

26. Hu, N.; Wang, Z.; Song, X.; Wei, L.; Kim, B.S.; Freedman, N.D.; Baek, J.; Burdette, L.; Chang, J.; Chung, C.; et al. Genome-wide association study of gastric adenocarcinoma in Asia: A comparison of associations between cardia and non-cardia tumours. *Gut* **2016**, *65*, 1611–1618. [CrossRef] [PubMed]

27. Wang, Z.; Dai, J.; Hu, N.; Miao, X.; Abnet, C.; Yang, M.; Freedman, N.D.; Chen, J.; Burdette, L.; Zhu, X.; et al. Identification of new susceptibility loci for gastric non-cardia adenocarcinoma: Pooled results from two Chinese genome-wide association studies. *Gut* **2017**, *66*, 581–587. [CrossRef]

28. Yan, C.; Zhu, M.; Ding, Y.; Yang, M.; Wang, M.; Li, G.; Ren, C.; Huang, T.; Yang, W.; He, B.; et al. Meta-analysis of genome-wide association studies and functional assays decipher susceptibility genes for gastric cancer in Chinese populations. *Gut* **2020**, *69*, 641–651. [CrossRef]

29. Gudmundsson, J.; Sulem, P.; Gudbjartsson, D.; Masson, G.; Petursdottir, V.; Hardarson, S.; Gudjonsson, S.A.; Johannsdottir, H.; Helgadottir, H.; Stacey, S.N.; et al. A common variant at 8q24.21 is associated with renal cell cancer. *Nat. Commun.* **2013**, *4*, 2776. [CrossRef]

30. Scelo, G.; Purdue, M.P.; Brown, K.M.; Johansson, M.; Wang, Z.; Eckel-Passow, J.E.; Ye, Y.; Hofmann, J.; Choi, J.; Foll, M.; et al. Genome-wide association study identifies multiple risk loci for renal cell carcinoma. *Nat. Commun.* **2017**, *8*, 15724. [CrossRef]

31. Purdue, M.; Johansson, M.; Zelenika, D.; Toro, J.R.; Scelo, G.; Moore, L.E.; Prokhortchouk, E.; Wu, X.; Kiemeney, L.; Gaborieau, V.; et al. Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13. *Nat. Genet.* **2011**, *43*, 60–65. [CrossRef]

32. Pharoah, P.D.P.; Australian Cancer Study; Tsai, Y.-Y.; Ramus, S.; Phelan, C.M.; Goode, E.L.; Lawrenson, K.; Buckley, M.; Fridley, B.L.; Tyrer, J.P.; et al. GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. *Nat. Genet.* **2013**, *45*, 362–370. [CrossRef]

33. Kuchenbaecker, K.B.; Ramus, S.J.; Tyrer, J.; Lee, A.; Shen, H.C.; Beesley, J.; Lawrenson, K.; McGuffog, L.; Healey, S.; Lee, J.M.; et al. Identification of six new susceptibility loci for invasive epithelial ovarian cancer. *Nat. Genet.* **2015**, *47*, 164–171. [CrossRef] [PubMed]

34. Gudmundsson, J.; Thorleifsson, G.; Sigurdsson, J.K.; Stefansdottir, L.; Jonasson, J.G.; Gudjonsson, S.A.; Gudbjartsson, D.F.; Masson, G.; Johannsdottir, H.; Halldorsson, G.H.; et al. A genome-wide association study yields five novel thyroid cancer risk loci. *Nat. Commun.* **2017**, *8*, 14517. [CrossRef]

35. Kinnersley, B.; Labussière, M.; Holroyd, A.; Di Stefano, A.-L.; Broderick, P.; Vijayakrishnan, J.; Mokhtari, K.; Delattre, J.-Y.; Gousias, K.; Schramm, J.; et al. Genome-wide association study identifies multiple susceptibility loci for glioma. *Nat. Commun.* **2015**, *6*, 8559. [CrossRef] [PubMed]

36. Bishop, D.T.; Demenais, F.; Iles, M.; Harland, M.; Taylor, J.C.; Corda, E.; Randerson-Moor, J.; Aitken, J.; Avril, M.-F.; Azizi, E.; et al. Genome-wide association study identifies three loci associated with melanoma risk. *Nat. Genet.* **2009**, *41*, 920–925. [CrossRef]

37. Genome-Wide Association Study Identifies Three New Melanoma Susceptibility Loci—PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/21983787/ (accessed on 26 August 2022).

38. Law, M.H.; Bishop, D.T.; Lee, J.E.; Brossard, M.; Martin, N.G.; Moses, E.K.; Song, F.; Barrett, J.H.; Kumar, R.; Easton, D.F.; et al. Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. *Nat. Genet.* **2015**, *47*, 987–995. [CrossRef] [PubMed]

39. Freedman, M.L.; Monteiro, A.N.A.; Gayther, S.A.; Coetzee, G.A.; Risch, A.; Plass, C.; Casey, G.; De Biasi, M.; Carlson, C.; Duggan, D.; et al. Principles for the post-GWAS functional characterization of cancer risk loci. *Nat. Genet.* **2011**, *43*, 513–518. [CrossRef] [PubMed]

40. Edwards, S.L.; Beesley, J.; French, J.D.; Dunning, A.M. Beyond GWASs: Illuminating the Dark Road from Association to Function. *Am. J. Hum. Genet.* **2013**, *93*, 779–797. [CrossRef]

41. Ng, P.C.; Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **2003**, *31*, 3812–3814. [CrossRef]

42. Lee, P.H.; Shatkay, H. F-SNP: Computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res.* **2008**, *36*, D820–D824. [CrossRef]

43. Ritchie, G.R.; Flicek, P. Computational approaches to interpreting genomic sequence variation. *Genome Med.* **2014**, *6*, 1760. [CrossRef]

44. Theodoratou, E.; Farrington, S.M.; Timofeeva, M.; Din, F.V.; Svinti, V.; Tenesa, A.; Liu, T.; Lindblom, A.; Gallinger, S.; Campbell, H.; et al. Genome-wide scan of the effect of common nsSNPs on colorectal cancer survival outcome. *Br. J. Cancer* **2018**, *119*, 988–993. [CrossRef]

45. Timofeeva, M.N.; Kinnersley, B.; Farrington, S.M.; Whiffin, N.; Palles, C.; Svinti, V.; Lloyd, A.; Gorman, M.; Ooi, L.-Y.; Hosking, F.; et al. Recurrent Coding Sequence Variation Explains Only a Small Fraction of the Genetic Architecture of Colorectal Cancer. *Sci. Rep.* **2015**, *5*, 16286. [CrossRef] [PubMed]

46. Wang, Y.; McKay, J.D.; Rafnar, T.; Wang, Z.; Timofeeva, M.N.; Broderick, P.; Zong, X.; Laplana, M.; Wei, Y.; Han, Y.; et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. *Nat. Genet.* **2014**, *46*, 736–741. [CrossRef] [PubMed]

47. Michailidou, K.; The Breast and Ovarian Cancer Susceptibility Collaboration; Hall, P.; Gonzalez-Neira, A.; Ghoussaini, M.; Dennis, J.; Milne, R.L.; Schmidt, M.; Chang-Claude, J.; Bojesen, S.E.; et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat. Genet.* **2013**, *45*, 353–361. [CrossRef] [PubMed]

48. Stacey, S.N.; Sulem, P.; Jonasdottir, A.; Masson, G.; Gudmundsson, J.; Gudbjartsson, D.F.; Magnusson, O.T.; Gudjonsson, S.A.; Sigurgeirsson, B.; Thorisdottir, K.; et al. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat. Genet.* **2011**, *43*, 1098–1103. [CrossRef]

49. Enciso-Mora, V.; Hosking, F.J.; Di Stefano, A.L.; Zelenika, D.; Shete, S.; Broderick, P.; Idbaih, A.; Delattre, J.-Y.; Hoang-Xuan, K.; Marie, Y.; et al. Low penetrance susceptibility to glioma is caused by the TP53 variant rs. *Br. J. Cancer* **2013**, *108*, 2178–2185. [CrossRef]

50. Tian, J.; Chen, C.; Rao, M.; Zhang, M.; Lu, Z.; Cai, Y.; Ying, P.; Li, B.; Wang, H.; Wang, L.; et al. Aberrant RNA Splicing Is a Primary Link between Genetic Variation and Pancreatic Cancer Risk. *Cancer Res.* **2022**, *82*, 2084–2096. [CrossRef]

51. Li, J.; Zou, L.; Zhou, Y.; Li, L.; Yang, Y.; Gong, Y.; Lou, J.; Ke, J.; Zhang, Y.; Tian, J.; et al. A low-frequency variant in SMAD7 modulates TGF-β signaling and confers risk for colorectal cancer in Chinese population. *Mol. Carcinog.* **2017**, *56*, 1798–1807. [CrossRef]

52. Roundtree, I.A.; Evans, M.E.; Pan, T.; He, C. Dynamic RNA Modifications in Gene Expression Regulation. *Cell* **2017**, *169*, 1187–1200. [CrossRef] [PubMed]

53. Frye, M.; Harada, B.T.; Behm, M.; He, C. RNA modifications modulate gene expression during development. *Science* **2018**, *361*, 1346–1349. [CrossRef]

54. Yue, Y.; Liu, J.; He, C. RNA N^6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes Dev.* **2015**, *29*, 1343–1355. [CrossRef]

55. Tian, J.; Ying, P.; Ke, J.; Zhu, Y.; Yang, Y.; Gong, Y.; Zou, D.; Peng, X.; Yang, N.; Wang, X.; et al. *ANKLE1* N^6-Methyladenosine-related variant is associated with colorectal cancer risk by maintaining the genomic stability. *Int. J. Cancer* **2020**, *146*, 3281–3293. [CrossRef] [PubMed]

56. Chang, J.; Tian, J.; Yang, Y.; Zhong, R.; Li, J.; Zhai, K.; Ke, J.; Lou, J.; Chen, W.; Zhu, B.; et al. A Rare Missense Variant in TCF7L2 Associates with Colorectal Cancer Risk by Interacting with a GWAS-Identified Regulatory Variant in the MYC Enhancer. *Cancer Res.* **2018**, *78*, 5164–5172. [CrossRef] [PubMed]

57. Maurano, M.T.; Humbert, R.; Rynes, E.; Thurman, R.E.; Haugen, E.; Wang, H.; Reynolds, A.P.; Sandstrom, R.; Qu, H.; Brody, J.; et al. Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science* **2012**, *337*, 1190–1195. [CrossRef] [PubMed]

58. Wei, G.-H.; Liu, D.-P.; Liang, C.-C. Charting gene regulatory networks: Strategies, challenges and perspectives. *Biochem. J.* **2004**, *381*, 1–12. [CrossRef]

59. Wei, G.H.; Liu, D.P.; Liang, C.C. Chromatin domain boundaries: Insulators and beyond. *Cell Res.* **2005**, *15*, 292–300. [CrossRef]

60. Haberle, V.; Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 621–637. [CrossRef]

61. Wu, S.; Meng, Q.; Zhang, C.; Sun, H.; Lu, R.; Gao, N.; Yang, H.; Li, X.; Aschner, M.; Chen, R. DR4 mediates the progression, invasion, metastasis and survival of colorectal cancer through the Sp1/NF1 switch axis on genomic locus. *Int. J. Cancer* **2018**, *143*, 289–297. [CrossRef]

62. Meng, Q.; Wu, S.; Wang, Y.; Xu, J.; Sun, H.; Lu, R.; Gao, N.; Yang, H.; Li, X.; Tang, B.; et al. *MPO* Promoter Polymorphism rs2333227 Enhances Malignant Phenotypes of Colorectal Cancer by Altering the Binding Affinity of AP-2α. *Cancer Res.* **2018**, *78*, 2760–2769. [CrossRef]

63. Lou, H.; Yeager, M.; Li, H.; Bosquet, J.G.; Hayes, R.B.; Orr, N.; Yu, K.; Hutchinson, A.; Jacobs, K.B.; Kraft, P.; et al. Fine mapping and functional analysis of a common variant in *MSMB* on chromosome 10q11.2 associated with prostate cancer susceptibility. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7933–7938. [CrossRef]

64. Hua, J.T.; Ahmed, M.; Guo, H.; Zhang, Y.; Chen, S.; Soares, F.; Lu, J.; Zhou, S.; Wang, M.; Li, H.; et al. Risk SNP-Mediated Promoter-Enhancer Switching Drives Prostate Cancer through lncRNA PCAT. *Cell* **2018**, *174*, 564–575. [CrossRef]

65. Bojesen, S.E.; Pooley, K.A.; Johnatty, S.E.; Beesley, J.; Michailidou, K.; Tyrer, J.P.; Edwards, S.L.; Pickett, H.A.; Shen, H.C.; Smart, C.E.; et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat. Genet.* **2013**, *45*, 371–384. [CrossRef] [PubMed]

66. Michailidou, K.; Lindstrom, S.; Dennis, J.; Beesley, J.; Hui, S.; Kar, S.; Lemacon, A.; Soucy, P.; Glubb, D.; Rostamianfar, A.; et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* **2017**, *551*, 92–94. [CrossRef] [PubMed]

67. Noncoding Genetic Variation in GATA3 Increases Acute Lymphoblastic Leukemia Risk through Local and Global Changes in Chromatin Conformation | Nature Genetics. Available online: https://www.nature.com/articles/s41588-021-00993-x (accessed on 11 October 2022).

68. Williamson, I.; Hill, R.E.; Bickmore, W.A. Enhancers: From Developmental Genetics to the Genetics of Common Human Disease. *Dev. Cell* **2011**, *21*, 17–19. [CrossRef] [PubMed]

69. Li, G.; Ruan, X.; Auerbach, R.K.; Sandhu, K.S.; Zheng, M.; Wang, P.; Poh, H.M.; Goh, Y.; Lim, J.; Zhang, J.; et al. Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. *Cell* **2012**, *148*, 84–98. [CrossRef] [PubMed]

70. Sur, I.; Taipale, I.S.J. The role of enhancers in cancer. *Nat. Cancer* **2016**, *16*, 483–493. [CrossRef] [PubMed]

71. Yan, J.; Chen, S.-A.A.; Local, A.; Liu, T.; Qiu, Y.; Dorighi, K.M.; Preissl, S.; Rivera, C.M.; Wang, C.; Ye, Z.; et al. Histone H3 lysine 4 monomethylation modulates long-range chromatin interactions at enhancers. *Cell Res.* **2018**, *28*, 204–220. [CrossRef] [PubMed]

72. Corradin, O.; Scacheri, P.C. Enhancer variants: Evaluating functions in common disease. *Genome Med.* **2014**, *6*, 85. [CrossRef]

73. Ward, L.D.; Kellis, M. Interpreting noncoding genetic variation in complex traits and human disease. *Nat. Biotechnol.* **2012**, *30*, 1095–1106. [CrossRef]

74. Zhang, Z.; Liu, X.; Li, L.; Yang, Y.; Yang, J.; Wang, Y.; Wu, J.; Wu, X.; Shan, L.; Pei, F.; et al. SNP rs4971059 predisposes to breast carcinogenesis and chemoresistance via TRIM46-mediated HDAC1 degradation. *EMBO J.* **2021**, *40*, e107974. [CrossRef]

75. Yang, H.; Ting, X.; Geng, Y.-H.; Xie, Y.; Nierenberg, J.L.; Huo, Y.-F.; Zhou, Y.-T.; Huang, Y.; Yu, Y.-Q.; Yu, X.-Y.; et al. The risk variant rs11836367 contributes to breast cancer onset and metastasis by attenuating Wnt signaling via regulating *NTN4* expression. *Sci. Adv.* **2022**, *8*, eabn3509. [CrossRef]

76. Tian, P.; Zhong, M.; Wei, G.-H. Mechanistic insights into genetic susceptibility to prostate cancer. *Cancer Lett.* **2021**, *522*, 155–163. [CrossRef] [PubMed]

77. Huang, Q.; Whitington, T.; Gao, P.; Lindberg, J.; Yang, Y.; Sun, J.; Väisänen, M.-R.; Szulkin, R.; Annala, M.; Yan, J.; et al. A prostate cancer susceptibility allele at 6q22 increases RFX6 expression by modulating HOXB13 chromatin binding. *Nat. Genet.* **2014**, *46*, 126–135. [CrossRef] [PubMed]

78. Gao, P.; Xia, J.-H.; Sipeky, C.; Dong, X.-M.; Zhang, Q.; Yang, Y.; Zhang, P.; Cruz, S.P.; Zhang, K.; Zhu, J.; et al. Biology and Clinical Implications of the 19q13 Aggressive Prostate Cancer Susceptibility Locus. *Cell* **2018**, *174*, 576–589. [CrossRef] [PubMed]

79. Zheng, Y.; Lei, T.; Jin, G.; Guo, H.; Zhang, N.; Chai, J.; Xie, M.; Xu, Y.; Wang, T.; Liu, J.; et al. LncPSCA in the 8q24.3 risk locus drives gastric cancer through destabilizing DDX. *EMBO Rep.* **2021**, *22*, e52707. [CrossRef] [PubMed]

80. Hepatocellular Carcinoma Risk Variant Modulates lncRNA HLA-DQB1-AS1 Expression via a Long-Range Enhancer–Promoter Interaction | Carcinogenesis | Oxford Academic. Available online: https://academic.oup.com/carcin/article/42/11/1347/ (accessed on 25 August 2022).

81. Schödel, J.; Bardella, C.; Sciesielski, L.; Brown, J.M.; Pugh, C.; Buckle, V.; Tomlinson, I.P.; Ratcliffe, P.; Mole, D.R. Common genetic variants at the 11q13.3 renal cancer susceptibility locus influence binding of HIF to an enhancer of cyclin D1 expression. *Nat. Genet.* **2012**, *44*, 420–425. [CrossRef]

82. Yu, C.-Y.; Han, J.-X.; Zhang, J.; Jiang, P.; Shen, C.; Guo, F.; Tang, J.; Yan, T.; Tian, X.; Zhu, X.; et al. A 16q22.1 variant confers susceptibility to colorectal cancer as a distal regulator of ZFP. *Oncogene* **2020**, *39*, 1347–1360. [CrossRef]

83. Tian, J.; Chang, J.; Gong, J.; Lou, J.; Fu, M.; Li, J.; Ke, J.; Zhu, Y.; Gong, Y.; Yang, Y.; et al. Systematic Functional Interrogation of Genes in GWAS Loci Identified ATF1 as a Key Driver in Colorectal Cancer Modulated by a Promoter-Enhancer Interaction. *Am. J. Hum. Genet.* **2019**, *105*, 29–47. [CrossRef]

84. Tian, J.; Lou, J.; Cai, Y.; Rao, M.; Lu, Z.; Zhu, Y.; Zou, D.; Peng, X.; Wang, H.; Zhang, M.; et al. Risk SNP-Mediated Enhancer–Promoter Interaction Drives Colorectal Cancer through Both *FADS2* and *AP002754*. *Cancer Res.* **2020**, *80*, 1804–1818. [CrossRef]

85. Capturing Chromosome Conformation | Science. Available online: https://www.science.org/doi/10.1126/science.1067799?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed (accessed on 25 August 2022).

86. Tolhuis, B.; Palstra, R.-J.; Splinter, E.; Grosveld, F.; de Laat, W. Looping and Interaction between Hypersensitive Sites in the Active β-globin Locus. *Mol. Cell* **2002**, *10*, 1453–1465. [CrossRef]

87. Giambartolomei, C.; Seo, J.-H.; Schwarz, T.; Freund, M.K.; Johnson, R.D.; Spisak, S.; Baca, S.C.; Gusev, A.; Mancuso, N.; Pasaniuc, B.; et al. H3K27ac HiChIP in prostate cell lines identifies risk genes for prostate cancer susceptibility. *Am. J. Hum. Genet.* **2021**, *108*, 2284–2300. [CrossRef]

88. Zhu, Y.; Gujar, A.D.; Wong, C.-H.; Tjong, H.; Ngan, C.Y.; Gong, L.; Chen, Y.-A.; Kim, H.; Liu, J.; Li, M.; et al. Oncogenic extrachromosomal DNA functions as mobile enhancers to globally amplify chromosomal transcription. *Cancer Cell* **2021**, *39*, 694–707. [CrossRef] [PubMed]

89. Zheng, H.; Xie, W. The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 535–550. [CrossRef] [PubMed]

90. Gorkin, D.U.; Leung, D.; Ren, B. The 3D Genome in Transcriptional Regulation and Pluripotency. *Cell Stem Cell* **2014**, *14*, 762–775. [CrossRef]

91. He, H.; Li, W.; Liyanarachchi, S.; Srinivas, M.; Wang, Y.; Akagi, K.; Wang, Y.; Wu, D.; Wang, Q.; Jin, V.; et al. Multiple functional variants in long-range enhancer elements contribute to the risk of SNP rs965513 in thyroid cancer. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6128–6133. [CrossRef] [PubMed]

92. Zhang, X.; Cowper-Sal·lari, R.; Bailey, S.D.; Moore, J.H.; Lupien, M. Integrative functional genomics identifies an enhancer looping to the *SOX9* gene disrupted by the 17q24.3 prostate cancer risk locus. *Genome Res.* **2012**, *22*, 1437–1446. [CrossRef] [PubMed]

93. Ling, J.Q.; Li, T.; Hu, J.F.; Vu, T.H.; Chen, H.L.; Qiu, X.W.; Cherry, A.M.; Hoffman, A.R. CTCF Mediates Interchromosomal Colocalization Between *Igf2/H19* and *Wsb1/Nf1*. *Science* **2006**, *312*, 269–272. [CrossRef]

94. Insulators: Exploiting Transcriptional and Epigenetic Mechanisms | Nature Reviews Genetics. Available online: https://www.nature.com/articles/nrg (accessed on 26 August 2022).

95. Yusufzai, T.M.; Tagami, H.; Nakatani, Y.; Felsenfeld, G. CTCF Tethers an Insulator to Subnuclear Sites, Suggesting Shared Insulator Mechanisms across Species. *Mol. Cell* **2004**, *13*, 291–298. [CrossRef]

96. Ahmed, M.; Soares, F.; Xia, J.-H.; Yang, Y.; Li, J.; Guo, H.; Su, P.; Tian, Y.; Lee, H.J.; Wang, M.; et al. CRISPRi screens reveal a DNA methylation-mediated 3D genome dependent causal mechanism in prostate cancer. *Nat. Commun.* **2021**, *12*, 1781. [CrossRef]

97. Statkiewicz, M.; Maryan, N.; Kulecka, M.; Kuklinska, U.; Ostrowski, J.; Mikula, M.; Czyżowska, A.; Barbasz, A. Functional analyses of a low-penetrance risk variant rs6702619/1p21.2 associating with colorectal cancer in Polish population. *Acta Biochim. Pol.* **2019**, *66*, 305–313. [CrossRef]

98. Claussnitzer, M.; Cho, J.H.; Collins, R.; Cox, N.J.; Dermitzakis, E.T.; Hurles, M.E.; Kathiresan, S.; Kenny, E.E.; Lindgren, C.M.; MacArthur, D.G.; et al. A brief history of human disease genetics. *Nature* **2020**, *577*, 179–189. [CrossRef]

99. Rahman, M.A.; Krainer, A.R.; Abdel-Wahab, O. SnapShot: Splicing Alterations in Cancer. *Cell* **2020**, *180*, 208–208.e1. [CrossRef] [PubMed]

100. He, L.; Hannon, G.J. MicroRNAs: Small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **2004**, *5*, 522–531. [CrossRef] [PubMed]

101. Esquela-Kerscher, A.; Slack, F. Oncomirs—MicroRNAs with a role in cancer. *Nat. Cancer* **2006**, *6*, 259–269. [CrossRef] [PubMed]

102. Hoffman, A.E.; Zheng, T.; Yi, C.; Leaderer, D.; Weidhaas, J.; Slack, F.; Zhang, Y.; Paranjape, T.; Zhu, Y. microRNA miR-196a-2 and Breast Cancer: A Genetic and Epigenetic Association Study and Functional Analysis. *Cancer Res.* **2009**, *69*, 5970–5977. [CrossRef] [PubMed]

103. Chen, J.; Jiang, Y.; Zhou, J.; Liu, S.; Qin, N.; Du, J.; Jin, G.; Hu, Z.; Ma, H.; Shen, H.; et al. Evaluation of CpG-SNPs in miRNA promoters and risk of breast cancer. *Gene* **2018**, *651*, 1–8. [CrossRef] [PubMed]

104. Gao, F.; Xiong, X.; Pan, W.; Yang, X.; Zhou, C.; Yuan, Q.; Zhou, L.; Yang, M. A Regulatory MDM4 Genetic Variant Locating in the Binding Sequence of Multiple MicroRNAs Contributes to Susceptibility of Small Cell Lung Cancer. *PLoS ONE* **2015**, *10*, e0135647. [CrossRef]

105. Wu, X.-M.; Yang, H.-G.; Zheng, B.-A.; Cao, H.-F.; Hu, Z.-M.; Wu, W.-D. Functional Genetic Variations at the microRNA Binding-Site in the CD44 Gene Are Associated with Risk of Colorectal Cancer in Chinese Populations. *PLoS ONE* **2015**, *10*, e0127557. [CrossRef]

106. Ke, J.; Tian, J.; Li, J.; Gong, Y.; Yang, Y.; Zhu, Y.; Zhang, Y.; Zhong, R.; Chang, J.; Gong, J. Identification of a Functional Polymorphism Affecting Microrna Binding in the Susceptibility Locus 1q25.3 for Colorectal Cancer. Wiley Online Library, 2021. Available online: https://onlinelibrary.wiley.com/doi/org/10.1002/mc.22649 (accessed on 26 August 2022).

107. Shen, C.; Yan, T.; Wang, Z.; Su, H.-C.; Zhu, X.; Tian, X.; Fang, J.-Y.; Chen, H.; Hong, J. Variant of SNP rs1317082 at CCSlnc362 (RP11-362K14.5) creates a binding site for miR-4658 and diminishes the susceptibility to CRC. *Cell Death Dis.* **2018**, *9*, 1177. [CrossRef]

108. Feng, T.; Feng, N.; Zhu, T.; Li, Q.; Zhang, Q.; Wang, Y.; Gao, M.; Zhou, B.; Yu, H.; Zheng, M.; et al. A SNP-mediated lncRNA (LOC146880) and microRNA (miR-539-5p) interaction and its potential impact on the NSCLC risk. *J. Exp. Clin. Cancer Res.* **2020**, *39*, 157. [CrossRef]

109. Wang, J.; Zou, Y.; Du, B.; Li, W.; Yu, G.; Li, L.; Zhou, L.; Gu, X.; Song, S.; Liu, Y.; et al. SNP-mediated lncRNA-ENTPD3-AS1 upregulation suppresses renal cell carcinoma via miR-155/HIF-1α signaling. *Cell Death Dis.* **2021**, *12*, 672. [CrossRef]

110. Fu, Y.; Zhang, Y.; Cui, J.; Yang, G.; Peng, S.; Mi, W.; Yin, X.; Yu, Y.; Jiang, J.; Liu, Q.; et al. SNP rs12982687 affects binding capacity of lncRNA UCA1 with miR-873-5p: Involvement in smoking-triggered colorectal cancer progression. *Cell Commun. Signal.* **2020**, *18*, 37. [CrossRef] [PubMed]

111. Hou, Y.; Zhou, M.; Li, Y.; Tian, T.; Sun, X.; Chen, M.; Xu, W.; Lu, M. Risk SNP-mediated LINC01614 upregulation drives head and neck squamous cell carcinoma progression via PI3K/AKT signaling pathway. *Mol. Carcinog.* **2022**, *61*, 797–811. [CrossRef]

112. The International HapMap Consortium A haplotype map of the human genome. *Nature* **2005**, *437*, 1299–1320. [CrossRef] [PubMed]

113. The 1000 Genomes Project Consortium A map of human genome variation from population-scale sequencing. *Nature* **2010**, *467*, 1061–1073. [CrossRef]

114. Lander, E.S. Initial impact of the sequencing of the human genome. *Nature* **2011**, *470*, 187–197. [CrossRef]

115. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [CrossRef]

116. Martin, F.J.; Amode, M.R.; Aneja, A.; Austine-Orimoloye, O.; Azov, A.G.; Barnes, I.; Flicek, P. Ensembl. *Nucleic Acids. Res.* **2022**, gkac958. [CrossRef]

117. McLaren, W.; Pritchard, B.; Rios, D.; Chen, Y.; Flicek, P.; Cunningham, F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **2010**, *26*, 2069–2070. [CrossRef]

118. Ramos, A.H.; Lichtenstein, L.; Gupta, M.; Lawrence, M.S.; Pugh, T.J.; Saksena, G.; Meyerson, M.; Getz, G. Oncotator: Cancer Variant Annotation Tool. *Hum. Mutat.* **2015**, *36*, E2423–E2429. [CrossRef]

119. Pan-Cancer Analysis Demonstrates that Integrating Polygenic Risk Scores with Modifiable Risk Factors Improves Risk Prediction | Nature Communications. Available online: https://www.nature.com/articles/s41467-020-19600- (accessed on 3 November 2022).

120. Schottenfeld, D.; Beebe-Dimmer, J.L.; Buffler, P.A.; Omenn, G.S. Current Perspective on the Global and United States Cancer Burden Attributable to Lifestyle and Environmental Risk Factors. *Annu. Rev. Public Health* **2013**, *34*, 97–117. [CrossRef]

121. Im, P.K.; Millwood, I.Y.; Kartsonaki, C.; Guo, Y.; Chen, Y.; Turnbull, I.; Yu, C.; Du, H.; Pei, P.; Lv, J.; et al. Alcohol drinking and risks of liver cancer and non-neoplastic chronic liver diseases in China: A 10-year prospective study of 0.5 million adults. *BMC Med.* **2021**, *19*, 216. [CrossRef] [PubMed]

122. Dam, M.K.; Hvidtfeldt, U.A.; Tjønneland, A.; Overvad, K.; Grønbæk, M.K.; Tolstrup, J. Five year change in alcohol intake and risk of breast cancer and coronary heart disease among postmenopausal women: Prospective cohort study. *BMJ* **2016**, *353*, i2314. [CrossRef] [PubMed]

123. Loftfield, E.; Stepien, M.; Viallon, V.; Trijsburg, L.; Rothwell, J.A.; Robinot, N.; Biessy, C.; Bergdahl, I.A.; Bodén, S.; Schulze, M.B.; et al. Novel Biomarkers of Habitual Alcohol Intake and Associations with Risk of Pancreatic and Liver Cancers and Liver Disease Mortality. *JNCI J. Natl. Cancer Inst.* **2021**, *113*, 1542–1550. [CrossRef] [PubMed]

124. Mayén, A.-L.; Viallon, V.; Botteri, E.; Proust-Lima, C.; Bagnardi, V.; Batista, V.; Cross, A.J.; Laouali, N.; MacDonald, C.J.; Severi, G.; et al. A longitudinal evaluation of alcohol intake throughout adulthood and colorectal cancer risk. *Eur. J. Epidemiol.* **2022**, *37*, 915–929. [CrossRef] [PubMed]

125. Key, T.J.; Reeves, G.K. Alcohol, Diet, and Risk of Breast Cancer—PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/27170004/ (accessed on 7 November 2022).

126. Boffetta, P.; Hashibe, M. Alcohol and Cancer. *Lancet Oncol.* **2006**, *7*, 149–156. [CrossRef]

127. Gomella, L.G. Prostate Cancer: Alcohol, Cancer and 5α-Reductase Inhibitors-Is There a Link? *Nat. Rev. Urol* **2014**, *11*, 253–254. [CrossRef]

128. Kuper, H.; Boffetta, P.; Adami, H. Tobacco Use and Cancer Causation: Association by Tumour Type. *J. Intern. Med.* **2002**, *252*. [CrossRef]

129. Hecht, S.S.; Hatsukami, D.K. Smokeless Tobacco and Cigarette Smoking: Chemical Mechanisms and Cancer Prevention. *Nat. Rev. Cancer* **2022**, *22*, 143–155. [CrossRef]

130. Malhotra, J.; Malvezzi, M.; Negri, E.; Vecchia, C.L.; Boffetta, P. Risk Factors for Lung Cancer Worldwide. *Eur. Respir. J.* **2016**, *48*, 889–902. [CrossRef]

131. Little, M.P.; Wakeford, R.; Zablotska, L.B.; Borrego, D.; Griffin, K.T.; Allodji, R.S.; de Vathaire, F.; Lee, C.; Brenner, A.V.; Miller, J.S.; et al. Lymphoma and Multiple Myeloma in Cohorts of Persons Exposed to Ionising Radiation at a Young Age. *Leukemia* **2021**, *35*, 2906–2916. [CrossRef]

132. Grady, W.M.; Russell, K. Ionizing Radiation and Rectal Cancer: Victims of Our Own Success. *Gastroenterology* **2005**, *128*, 1114–1117. [CrossRef] [PubMed]

133. McGlynn, K.A.; Trabert, B. Adolescent and Adult Risk Factors for Testicular Cancer. *Nat. rev. Urol.* **2012**, *9*. [CrossRef] [PubMed]

134. Hiller, T.W.R.; O'Sullivan, D.E.; Brenner, D.R.; Peters, C.E.; King, W.D. Solar Ultraviolet Radiation and Breast Cancer Risk: A Systematic Review and Meta-Analysis. *Environ. Health Perspect.* **2020**, *128*, 16002. [CrossRef] [PubMed]

135. Teng, Y.; Yu, Y.; Li, S.; Huang, Y.; Xu, D.; Tao, X.; Fan, Y. Ultraviolet Radiation and Basal Cell Carcinoma: An Environmental Perspective. *Front. Public Health* **2021**, *9*, 666528. [CrossRef]

136. Iyengar, N.M.; Gucalp, A.; Dannenberg, A.J.; Hudis, C.A. Obesity and Cancer Mechanisms: Tumor Microenvironment and Inflammation. *J. Clin. Oncol.* **2016**, *34*, 4270–4276. [CrossRef]

137. Avgerinos, K.I.; Spyrou, N.; Mantzoros, C.S.; Dalamaga, M. Obesity and Cancer Risk: Emerging Biological Mechanisms and Perspectives—PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/30445141/ (accessed on 7 November 2022).

138. Bardou, M.; Barkun, A.N.; Martel, M. Obesity and Colorectal Cancer. *Gut* **2013**, *62*, 933–947. [CrossRef]

139. Lega, I.C.; Lipscombe, L.L. Review: Diabetes, Obesity, and Cancer-Pathophysiology and Clinical Implications. *Endocr. Rev.* **2020**, *41*, bnz014. [CrossRef]

140. Shibata, T.; Lieblong, B.J.; Sasagawa, T.; Nakagawa, M. The Promise of Combining Cancer Vaccine and Checkpoint Blockade for Treating HPV-Related Cancer. *Cancer Treat. Rev.* **2019**, *78*, 8–16. [CrossRef]

141. Leemans, C.R.; Braakhuis, B.J.M.; Brakenhoff, R.H. The Molecular Biology of Head and Neck Cancer. *Nat. Rev. Cancer* **2011**, *11*, 9–22. [CrossRef]

142. Chaturvedi, A.K.; Engels, E.A.; Pfeiffer, R.M.; Hernandez, B.Y.; Xiao, W.; Kim, E.; Jiang, B.; Goodman, M.T.; Sibug-Saber, M.; Cozen, W.; et al. Human Papillomavirus and Rising Oropharyngeal Cancer Incidence in the United States—PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/21969503/ (accessed on 7 November 2022).

143. Levrero, M.; Zucman-Rossi, J. Mechanisms of HBV-Induced Hepatocellular Carcinoma. *J. Hepatol.* **2016**, *64*, S84–S101. [CrossRef]

144. McGlynn, K.A.; Petrick, J.L.; El-Serag, H.B. Epidemiology of Hepatocellular Carcinoma. *Hepatology* **2021**, *73* (Suppl. 1), 4–13. [CrossRef] [PubMed]

145. Ko, Y.-H. EBV and Human Cancer. *Exp. Mol. Med.* **2015**, *47*, e130. [CrossRef] [PubMed]

146. Chen, B.; Zhang, Y.; Dai, S.; Zhou, P.; Luo, W.; Wang, Z.; Chen, X.; Cheng, P.; Zheng, G.; Ren, J.; et al. Molecular Characteristics of Primary Pulmonary Lymphoepithelioma-like Carcinoma Based on Integrated Genomic Analyses. *Signal Transduct Target Ther.* **2021**, *6*, 6. [CrossRef] [PubMed]

147. Murray, P.G.; Young, L.S. An Etiological Role for the Epstein-Barr Virus in the Pathogenesis of Classical Hodgkin Lymphoma. *Blood* **2019**, *134*, 591–596. [CrossRef] [PubMed]

148. Raderer, M.; Kiesewetter, B.; Ferreri, A.J.M. Clinicopathologic Characteristics and Treatment of Marginal Zone Lymphoma of Mucosa-Associated Lymphoid Tissue (MALT Lymphoma). *CA Cancer J. Clin.* **2016**, *66*, 153–171. [CrossRef]

149. Kao, Y.-S.; Chen, C.-W.; Wei, J.C.-C. Helicobacter Pylori Infection and Risk of Gastric Cancer. *Lancet Public Health* **2022**, *7*, e302. [CrossRef]

150. Yang, L.; Kartsonaki, C.; Yao, P.; de Martel, C.; Plummer, M.; Chapman, D.; Guo, Y.; Clark, S.; Walters, R.G.; Chen, Y.; et al. The Relative and Attributable Risks of Cardia and Non-Cardia Gastric Cancer Associated with Helicobacter Pylori Infection in China: A Case-Cohort Study. *Lancet Public Health* **2021**, *6*, e888–e896. [CrossRef]

151. Würtz, E.T.; Hansen, J.; Røe, O.D.; Omland, Ø. Asbestos Exposure and Haematological Malignancies: A Danish Cohort Study. *Eur. J. Epidemiol.* **2020**, *35*, 949–960. [CrossRef]

152. Pelucchi, C.; La Vecchia, C.; Bosetti, C.; Boyle, P.; Boffetta, P. Exposure to Acrylamide and Human Cancer–a Review and Meta-Analysis of Epidemiologic Studies. *Ann. Oncol.* **2011**, *22*, 1487–1499. [CrossRef]

153. Li, P.; Deng, S.-S.; Wang, J.-B.; Iwata, A.; Qiao, Y.-L.; Dai, X.-B.; Boffetta, P. Occupational and Environmental Cancer Incidence and Mortality in China. *Occup Med* **2012**, *62*, 281–287. [CrossRef]

154. Bosetti, C.; Boffetta, P.; La Vecchia, C. Occupational Exposures to Polycyclic Aromatic Hydrocarbons, and Respiratory and Urinary Tract Cancers: A Quantitative Review to 2005. *Ann. Oncol.* **2007**, *18*, 431–446. [CrossRef] [PubMed]