ORIENTAL JOURNAL OF CHEMISTRY
An International Open Access, Peer Reviewed Research Journal

www.orientjchem.org

ISSN: 0970-020 X
CODEN: OJCHEG
2026, Vol. 42, No.(5):
Pg. 2172-2183

GWAS in Cancer Susceptibility and Therapeutic Target Discovery

ALANKAR SHRIVASTAV[1]*, PREETI KUMARI[2], KAJAL DEVI[3] and APOORV RASTOGI[4]

[1]Pharmacy Academy, Faculty of Pharmacy, IFTM University, Moradabad, U.P., India.
[2]Department of Pharmaceutical Chemistry, Radha Govind University, Chandausi, India
[3]Department of Pharmacognosy, Shri Satya College of Higher Education, Moradabad, India.
[4]Department of Pharmacognosy, Moradabad, India.
*Corresponding author E-mail: alankar1994.ss@gmail.com

http://dx.doi.org/10.13005/ojc/420533

(Received: January 23, 2026; Accepted: April 10, 2026)

ABSTRACT

Genome-wide association studies (GWAS) have emerged as a revolution in the discovery of genetic determinants underlying cancer susceptibility by revealing germline variants that influence risk, progression and treatment response. In recent years, large-scale GWAS consortia have identified thousands of cancer loci in a wide variety of tumor types and importantly have shed light on inherited risk for cancer and functional elements within the noncoding genome as major contributors to oncogenic processes. These discoveries have widened the horizon of cancer genetics beyond high-penetrance mutations by establishing and recognizing for the first time ever, gaining importance of cumulative effect of common low-effect alleles and polygenic risk. In addition, the integration of GWAS data with transcriptomic, epigenomic and proteomic information has expedited functional interpretation of risk loci revealing pathways involved in DNA repair, immune modulation, cell proliferation and hormone signaling. Aside from vulnerability, GWAS-based findings have also been serving to shape medical strategies, by revealing druggable targets, biomarkers for treatment selection and risk stratification models. However, challenges including missing heritability, lack of population representation, gene-environment interactions and poor clinical translation remain. In this review, we discuss the current GWAS findings in several common cancers, recent methodological and analytical developments of GWAS studies that help improve our understanding of cancer biology, knowledge gained from GWAS findings for potential clinical application for therapy development, and address challenges encountered by these studies toward precision oncology.

Keywords: Genome-wide association studies (GWAS), Cancer susceptibility, Germline variants, Polygenic risk, Genetic biomarkers, Therapeutic target discovery, Precision oncology, Translational genomics, Cancer risk prediction, Multi-omics integration.

This is an Open Access article licensed under a Creative Commons license: Attribution 4.0 International (CC- BY).
Published by Oriental Scientific Publishing Company © 2018

INTRODUCTION

GWAS have emerged as a successful avenue for mapping the polygenic components of complex diseases, including most types of cancers. In contrast to Mendelian cancer syndromes caused by high- penetrance germline mutations (eg, BRCA1/2 in breast cancer or APC in familial adenomatous polyposis), most sporadic cancers result from the combined effects of many low-penetrance risk alleles with environmental, hormonal, immunological and metabolic factors [1]. GWAS allow for a large-scale interrogation of millions of common single nucleotide polymorphisms (SNPs) throughout the genome for association with disease phenotypes in the absence of prior knowledge about gene function or biological pathways. Although it has limited power to detect rare variants, this hypothesis-agnostic approach has transformed cancer genomics through the discovery of novel predisposition loci, many localizing to non-coding regulatory regions that affect gene expression rather than protein sequence [2].

In recent years, the increasing number of large biobanks and international consortia, combined with high-density genotyping technologies, has greatly increased the strength and resolution of GWAS. Now that enormous resources like the UK Biobank, dbGaP, TCGA and cancer coordination projects are available, studies of hundreds of thousands of individuals have begun to be possible - at last with enough statistical power for low-effect variants [3]. Such initiatives have uncovered thousands of loci with involvement in breast, prostate, colorectal, lung and hematological cancers and other cancer sites, providing insights into mechanisms underlying carcinogenesis and knowledge to target for therapy. Of importance, GWAS findings have also been the basis for polygenic risk scoring, a clinical device that shows increasing promise in risk stratification, prioritization screening and preventive care [4].

Despite these advances, challenges persist. Many GWAS signals localize to enhancers, promoters, non-coding RNA genes or transcription factor (TF) binding sites and demand deep functional annotations to pinpoint causal genes. Furthermore, the population representation continues to be imbalanced towards those of European descent, with concerns on bias and decreased clinical relevance in diverse global populations [5]. As cancer prevention and precision oncology shift towards combining germline and somatic information, GWAS will be a vital conduit between inherited risk, molecular pathways and therapeutic targets.

Cancer Genetics and Hereditary Susceptibility

Cancer susceptibility is influenced by a continuum of genetic effects, from rare high-penetrance mutations to common low-effect alleles. High-penetrance mutations in BRCA1/2, TP53 (p53), PTEN, MLH1 and SMAD4 are associated with significant lifetime cancer risk but represent a minority of overall cancer cases. Polygenic models, on the other hand, suggest that cancer risk in the population overall is affected by the aggregate effect of many variants with each contributing a minor amount (odds ratios < 1.05–1.3) [6]. GWAS have been especially valuable as tools to define this polygenic milieu and identify sequence variants that correlate with germline predispositions for tumor initiation, immune surveillance, hormone signaling, and DNA damage repair [7].

Notably, the heritability of cancer is different by tumor type. In twin studies, the heritability of susceptibility to prostate cancer has been estimated as 57%, colorectal cancer 31%, and breast cancer 27% secretion, whereas that for lung cancer is greatly modified by environmental exposures such as smoking. GWAS strongly supports these patterns in that it has identified the strongest germline contributions to hormone-dependent cancers and immune-mediated diseases[8]. With the increasing number of loci identified, predictive models keep improving with personalized preventive oncology as a prospect.

Rationale for GWAS in Cancer Research

The rationale for applying GWAS to cancer includes several key motivations:

1. Identifying novel susceptibility genes and pathways,
2. Improving understanding of tumor biology,
3. Informing screening and prevention, and
4. Discovering potential therapeutic targets or repurposable drugs.

Many GWAS loci pinpoint genes not previously associated with cancer. For instance, these include variants near FGFR2 for breast cancer, HOXB13 and TERT for prostate as well as

multiple other cancers, and ASXL2 for hematologic malignancies. Some map to expression-regulating regions affecting transcription factors networks or chromatin modification components. These connections often intercept at common oncogenic mechanisms of action, including regulation of the cell cycle, induction of apoptosis, DNA repair and inflammation. From a translational perspective, these results support both target delineation for drug development, prognostic modeling and stratified clinical management in combination with patterns of somatic mutation and functional genomic data [9].

Principles of Genome-Wide Association Studies

GWAS use high-resolution genotyping chips to measure frequency differences across millions of SNPs between cases and controls. The main statistical goal is to identify variants that are distributed differently in cases and controls, which can be represented as being associated with disease susceptibility. GWAS generally needs a large number of samples as the effect sizes of common variants are usually very weak and because corrections for many comparisons require high stringency (e.g., $p < 5 \times 10^{-8}$) to avoid false positive errors [10].

Table 1: List of types of cancer, candidate genes and functional category 12

Cancer Type	Key Loci Identified	Candidate Gene(s)	Functional Category	Notes
Breast Cancer	10q26, 6q25, 19p13	FGFR2, ESR1, BABAM1	Hormone signaling / DNA repair	Multiple non-coding enhancer SNPs
Prostate Cancer	8q24, 17q12, 10q11.2	MYC, HNF1B, MSMB	Cell proliferation / secretory function	8q24 hotspot shared across cancers
Colorectal Cancer	8q23.3, 18q21, 11q23	EIF3H, SMAD7, POU2AF1	Wnt/TGF-β signaling	Strong polygenic component
Lung Cancer	5p15.33, 6p21, 15q25	TERT, HLA, CHRNA5	Telomerase / immunity / nicotine receptor	Gene–environment interaction (smoking)
Ovarian Cancer	9p22.2, 8q24.21, 19p13.11	BNC2, MYC, BABAM1	DNA damage repair / proliferation	Overlap with breast cancer
Melanoma	9p21, 16q24, 20q11	CDKN2A, MC1R, ASIP	Pigmentation / cell cycle	Pigment pathways dominate
AML/ALL	7p12.2, 10p12.31, 10q21.2	IKZF1, ARID5B, PIP4K2A	Hematopoietic differentiation	Germline influence in childhood ALL
Pancreatic Cancer	13q12.2, 5p15.33	PDX1, TERT	Developmental / telomerase	Suggests stemness mechanisms

The majority of modern GWAS pilots use genotype imputation to reference haplotype panels (e.g., 1000 Genomes, HRC) to impute untyped SNPs. thus, gaining greater genomic resolution. Moreover, meta-analytic frameworks enable combining results across sets of similar cohorts to increase the power. Since then, the field has advanced from single-variant approaches to methodologies involving gene-based and pathway analyses as well as Bayesian fine mapping for the discovery of causal variants and to prioritize functional studies downstream [11].

GWAS Study Design and Statistical Models

GWAS are typically conceived as large scale case–control association studies in which allele frequencies at millions of SNPs are compared between individuals with cancer and healthy controls. Several potential study designs are available [13], such as (1) classical case–control studies, (2) nested case–control within cohorts, (3) population-based and (4) meta-analytic combining across consortia. Design choices impact power, the degree to which confounding can be mitigated and the generalizability of findings. When applied to cancer

studies, case ascertainment may rely on histological subtype, tumor grade, age of onset or family history, permitting stratified genetic analyses [14].

The primary test in GWAS compares the significance of individual SNPs associated with disease risk, typically via logistic regression for binary cancer traits or the Cox proportional hazards model for time-to-event phenotypes such as cancer survival. These models can be adjusted by the covariates, including age, sex, principal components of ancestry, smoking dose, hormonal status and other potential confounders) to decrease population stratification. In order to control false positives due to multiple testing, genome-wide significance thresholds generally use a Bonferroni-corrected threshold or permutation-based thresholds of $p < 5 \times 10^{-8}$ as a standard criterion for the significant association[15].

Alternative association models have been developed to address complex scenarios, including mixed-model frameworks (e.g., GEMMA, REGENIE) that incorporate genetic relatedness matrices for family-structured or biobank data, Bayesian variable selection models for fine mapping of causal loci, and multi-trait GWAS methods (MTAG) that leverage genetic correlations among related cancers. More recently, polygenic additive models have been complemented by interaction tests exploring gene–gene (epistasis) and gene–environment interactions, particularly relevant to cancers influenced by lifestyle exposures (e.g., tobacco and lung cancer) [16].

Bioinformatics Pipelines and Analytical Frameworks in GWAS

The standard framework of a cancer GWAS analysis includes multiple stages, ranging from sample quality control (QC), genotype calling and imputation, to association testing, fine mapping and biological interpretation. These (QC) procedures filter according to criteria of call rate thresholds, Hardy–Weinberg disequilibrium, cryptic relatedness, sex discordance and sample contamination. Subsequent to QC, imputation of genotypes against large reference panels (e.g., 1000 Genomes, HRC, TOPMed) raises the SNP density from hundreds of thousands to tens of millions, in turn allowing for identification of regions not captured directly by genotyping arrays [17].

Downstream interpretation links the associated variants to regulation landscapes with the integration of multiple annotations, including but not limited to several (i) ENCODE (39), Roadmap Epigenomics (40), GTEx (41), eQTL databases and 3D chromosome conformation maps, and expression quantitative trait loci datasets. Functional annotation tools (FUMA, ANNOVAR, MAGMA and HaploReg) can reveal whether the variants map at promoters, enhancers, transcription factor binding sites (TFBS), histone marks and/or splice junctions. For cancer, this is particularly significant because the majority of risk SNPs that have been identified are in non-coding regulatory domains rather than exonic protein-coding sequences [18].

With a multi-omics integrative strategy, we can interpret this statistical association into biological mechanisms. PrediXcan/TWAS approaches estimate the effect of genetic variation on gene expression to connect germline associations with transcriptomic profiles. Chromatin conformation technologies (Hi-C, ChIA-PET, CaptureC) generate contact maps that connect the positions of regulatory SNPs to distant gene promoters, an essential feature given the modular nature of cancer enhancers. In addition, direct candidate locus perturbation can also be performed using CRISPR-Cas9 screening systems to confirm their functional roles in tumorigenesis [19].

GWAS Discoveries in Major Cancer Types

Genome-wide association studies have identified thousands of susceptibility loci across multiple tumor types, enabling insights into cancer biology, risk prediction, and therapeutic development. Below is a detailed examination of discoveries across the most extensively studied cancers [20].

Breast Cancer

Breast cancer genome-wide association studies (GWAS) have given rise to one of the most extensive sets of risk loci discovered thus far, with over 200 breast cancer susceptibility variants documented as of today. Numerous linked SNPs are located in the proximity of genes implicated in hormone signaling (for example, ESR1, FGFR2), DNA repair (BRCA1, BRCA2, RAD51), chromatin remodeling (SMAD3) androgenesis and proliferation. A significant proportion of risk loci are located in mammalian cell-active enhancers and regions harboring estrogen receptor binding sites, which

further emphasizes the endocrine-dependent nature of breast cancer [21].

The FGFR2 locus at 10q26 is one of the most significant common genetic risk factors, and functional studies have reported an enhancer-mediated influence on epithelial cell proliferation. Polygenic risk scores (PRS) for breast cancer derived from GWAS are beyond being tested for risk stratification and screening, and when combined with BRCA1/2 penetrance models, can enhance preventive decision making. Multi-ancestry GWAS studies have identified population-specific signals, and highlight the importance of diverse cohorts for equitable translation [22].

Prostate Cancer

Prostate cancer is one of the more heritable cancers and >170 risk loci have been identified through genome-wide association studies (GWAS). An interesting finding is the consistent detection of signals on 8q24, which contains regulatory elements modulating activity of MYC, a master oncogene controlling cellular expansion. Other loci suggest roles for genes involved in androgen signaling (HRH1, KCNQ1), immune function (COLEC12), apoptosis (MCF2L) and vesicular transport/smooth muscle contractility/regulation of prostate secretion (HNF1B, HNFB1, MSMB). The polygenic risk models for prostate cancer are one of the most clinically applicable, particularly for early detection and biopsy decisions [23].

Emerging studies show germline variants also modulate therapeutic response: for example, variants influencing androgen receptor activity may affect sensitivity to androgen deprivation therapy (ADT). These insights highlight the convergence of germline and somatic pathways within precision oncology [24].

Colorectal Cancer

Susceptibility loci related to SMAD7, POU2AF1, EIF3H and genes involved in the Wnt or TGF-βpathways have previously been reported from colorectal cancer GWAS. Interactions with the environment are of particular importance as diet, microbiome contents, obesity and chronic inflammation could all effect disease course. Of importance, variation at SMAD7 disrupts TGF-β antioncogenic signaling and thus provides a mechanistic link between germline variation and pathway imbalance [25]. Specific associations at the population level, including East Asian-specific loci near ZNF106 and STXBP5-AS1, reveal underlying genetic heterogeneity across ethnic groups. Integrating GWAS with microbiome data is an evolving horizon for colorectal cancer studies [26].

Lung Cancer

Lung cancer GWAS point to direct and indirect genetic dependence of risk. Variants in 15q25 near CHRNA3/CHRNA5 are associated with the propensity to smoke and SCZ, suggesting how germline genetics modify risk for an environmental exposure. Other loci are the telomerase regulators TERT on 5p15. 33 and immunity-related HLA genomic regions. Gene–environment interaction analyses highlight different germline contributions between smoker and nonsmoker subgroups and across histological subtypes (adenocarcinoma and squamous cell carcinoma) [27].

Melanoma

Melanoma GWAS highlights pigmentation genetics and light response pathways; pivotal loci include MC1R, ASIP, TYR and KITLG. Several exist as regulators of melanocyte differentiation, melanin production or UV-stimulated DNA repair. Further loci indicate immune surveillance and T cell function, providing insight into the immunobiology of melanoma. Germline risk scores associated with nevus counts and skin phototype -- combining behavioral and genetic markers of susceptibility [28].

Hematologic Malignancies

GWAS studies in blood malignancies, including acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), lymphomas, and myeloproliferative neoplasms have identified susceptibility loci that implicate processes of hematopoietic lineage commitment, immune signaling, and chromatin regulation. Childhood ALL, for example, has one of the strongest germline risk architectures across cancers including IKZF1, ARID5B/CEBPE and CDKN2A polymorphisms that impact B cell developmental pathways and age-related susceptibility to this disease262–264 [29]. These data are in favor of a model where alterations of early lymphoid differentiation play an essential role in predisposing to leukemic transformation.

For lymphomas, GWAS studies have identified associations in the HLA region confirming antigen presentation and immune dysregulation as drivers of lymphomagenesis. The CLL loci suggest roles in B-cell receptor signaling, apoptosis and chromatin regulation (for example, through IRF8 and BCL2L11) [30]. Clinically, these germline-based findings are consistent with the clinical success of immunomodulatory and kinase inhibitors (e.g., BTKs, anti-CD20s). Myeloproliferative neoplasms are correlated with JAK-STAT regulators and inflammatory pathways in addition to somatic driver mutations including JAK2 V617F. These overlaps further demonstrate the interplay between germline susceptibility and somatic evolution in determining malignant phenotypes [31].

Genetic Architecture of Cancer Susceptibility (Rewritten Manuscript Style)

The genetic basis of cancer susceptibility consists of an array of inherited factors that span rare high penetrance mutations to common low effect alleles that together control risk. This structure mirrors monogenic inherited cancer syndromes and polygenic background in sporadic cancers [32]. In a conceptual sense, cancer susceptibility can be partitioned into three broad groups: (i) highly penetrant pathogenic variants; (ii) common risk alleles contributing to polygenic liability; and (iii) gene–gene and gene–environment interactions that modify risk in particular biological or exposure contexts [32].

High-Penetrance Mutations

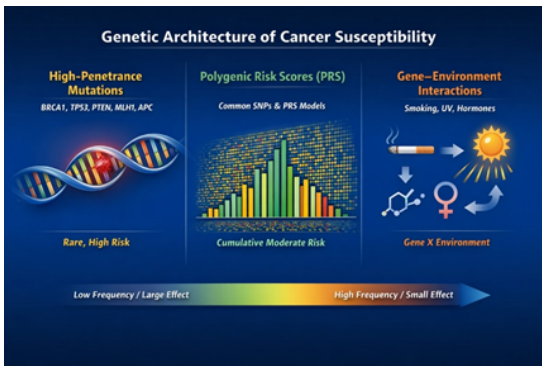
Highly penetrant germline mutations tend to be rare but carry a high level of lifetime risk and they may present as classic hereditary cancer syndromes. Prominent examples include pathogenic proteins encoded by BRCA1 and BRCA2 that provide a functionally compromised homologous recombination repair with consequent higher penetrance to breast and ovarian cancers, mutations in TP53 for which Li–Fraumeni is dominant, PTEN in Cowden syndrome and mismatch repair genes such as MLH1 giving rise to microsatellite instability-driven colorectal cancers. Although high-risk variants have clinical implications for surveillance, prevention interventions, and targeted therapy approaches, they explain a small fraction of the overall cancer burden in population-based studies indicating a substantial role of additional inherited genetic factors [33].

Common Low-Effect Variants and Polygenic Risk

Common SNPs have been found to contribute only modestly to cancer susceptibility in genome-wide association studies (GWAS). Individual risk alleles typically exhibit modest effect sizes (odds ratio ~1.05 – 1.3), but collectively they account for a large proportion of the heritable magnitude of the common malignancies such as breast and prostate cancers [34]. The combination of these alleles into polygenic risk scores (PRS) allows for the classification of subjects based on their inherited risk load. PRS are currently under consideration for different clinical uses including additional screening, risk-directed follow-up, preventive chemoprevention approaches and at-risk communication in genetic counseling13–15. The evidence is most mature for prostate cancer, where PRS-based screening strategies can reduce unnecessary biopsies, while maintaining the sensitivity of cancer detection [35].

Gene–Gene and Gene–Environment Interactions

Inherited genetic susceptibility does not act alone, it interacts with both endogenous biological pathways and exogenous exposures. Gene–gene interactions can explain increase or decrease of disease risk, nonetheless measuring them is hard because it requires statistical power. Gene–environment interaction is also crucial, in which inherited factors interact with environmental risk exposures [36]. For example, CHRNA5 polymorphisms affect nicotine addiction and therefore indirectly lung cancer risk; pigmentation genes (e.g., MC1R) affect sensitivity to ultraviolet radiation and susceptibility melanoma; hormone-related genetic variants determine the online but also environmental impact of endogenous as well exogenous estrogen exposure on breast cancer [37]. However, while



recent studies have shed new light on these interactions, a complete understanding of the basis for them is complicated by exposure heterogeneity, measurement error, and the requirement of large, carefully phenotyped cohorts [38].

Functional Interpretation of GWAS Signals

Most cancer GWAS-discovered variants occur in non-coding regions of the human genome, therefore making it difficult to connect statistical associations with molecular mechanisms. Given that these variants do not perturb protein-coding regions, their functional interpretation necessitates the merging of other informatics bioresources (regulatory genomics and chromatin biology among others) and molecular assays to characterize the effect of causal variants, target genes, or pathways involved in driving oncogenic phenotypes [40].

Linking Non-Coding Variants to Gene Regulation

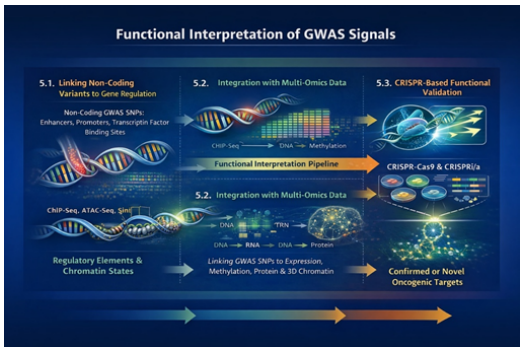
Many other non-coding variants increase the risk of cancer and are localized within enhancers, promoters, insulators, or motifs with potential to bind TFs responsible for regulation specific to a certain type of cell. These variants may modulate chromatin accessibility, hormone-stimulated transcriptional responses or the function of lineage-specific regulatory programs [41]. High-resolution techniques such as chromatin immunoprecipitation sequencing (ChIP-seq), assay for transposase-accessible chromatin sequencing (ATAC-seq) and emerging genomics methods for single cell chromatin profiling have shown that in many instances, breast and prostate cancer risk loci map to hormone-responsive enhancers associated with epithelial cell compartments. These observations underscore the impact of cis-regulatory architecture on inherited susceptibility [42].

Integration with Multi-Omics Data

The functional annotation of GWAS variants is being increasingly addressed by multi-omics integration to capture intermediate molecular phenotypes. Expression quantitative trait loci (eQTL) and methylation QTL (meQTL) analyses connect germline variation to transcriptional and epigenomic readouts, and protein QTL (pQTL) studies further this frame for translational abundance. Transcriptome-wide association pipelines (TWASs), eg, PrediXcan, estimate expression-mediated genetic effects, whereas chromatin conformation methods (eg, Hi-C and ChIA-PET) determine 3D gene regulatory interactions between enhancers and promoters. Taken together, these modes of integration increase confidence in target gene prioritization and provide mechanistic hypotheses linking germline variation to tumor phenotypes [43].

CRISPR-Based Functional Validation

Technical developments in genome editing have allowed for the actual experimental measurements of candidate regulatory variants. (1) CRISPR-Cas9 nuclease editing and CRISPR interference/activation (CRISPRi/a) technologies offer single-locus targeting of enhancers, promoters, non-coding RNAs and potential regulatory gene targets at high resolution, whereas pooled screens provide scalable investigation of variant function across a range of genomic loci. These tools have confirmed susceptibility gene candidates, including FGFR2 and TERT, and identified super-enhancer hubs that coordinate oncogenic signaling pathways [44]. As such, CRISPR-based methods represent a vital link between statistical association and laboratory-determined molecular causation.



GWAS and Therapeutic Target Discovery

Genome-wide association studies are increasingly having a major impact on translational oncology by identifying associations or their functional mechanisms that, directly and indirectly, lead to therapeutic target identification, drug repurposing and clinical precision medicine. Germline susceptibility loci not only help to explain the etiology of cancer but can also identify biological pathways suitable for therapeutic intervention, thus increasingly acting as a bridge between human genetics and drug discovery pipelines [46].

Target Prioritization

GWAS signals often lead to genes and

pathways that can be targeted by current or future therapeutic strategies. These susceptibility loci to date (BRCA1, BRCA2, FGFR2 and TERT) demonstrate how germline variation can highlight vulnerabilities in DNA repair, tyrosine kinase receptor signaling as well as telomerase biology [47]. These considerations can be exploited for the rational development or refinement of small-molecule inhibitors, kinase-targeted agents, immune-based treatments, hormonal modulatory approaches to treatment and strategies rooted in synthetic lethality (e.g., PARP inhibition). Given that these associations are uncovered from human disease biology, and not animal models or in vitro perturbation studies, they tend to represent more therapeutically relevant dependencies with better translational potential [48].

Drug Repurposing

An important potential of GWAS-informed pathway analyses is to identify existing pharmacotherapies that can be repurposed for oncologic indications. GWAS signals localized to immune regulatory loci in melanoma are consistent with the widely appreciated response rates of immune checkpoint blockade, whereas susceptibility loci linked to hormone signaling in breast and prostate cancer validate the therapeutic relevance of endocrine manipulation [49]. Repurposing provides an opportunity for swift reapportioning shortening the time required to get therapies deployed, since drug safety studies during early-phase development are avoided and because the PK/PD of agents is well defined this can lead to reduced development time and cost as well as decreased attrition [50].

Biomarkers for Risk Stratification and Clinical Management

Inherited genetic markers also have value as biomarkers for individualized risk assessment and clinical management. Polygenic risk scores (PRS) are being evaluated to tailor screening intervals, imaging modality selection, and surveillance intensity. For example, high-PRS breast cancer individuals may benefit from earlier or more frequent mammographic or MRI surveillance; men with elevated prostate PRS can be prioritized for PSA-based screening; and individuals with germline pigmentation risk alleles associated with melanoma may warrant enhanced dermatologic monitoring [51]. The successful integration of PRS into clinical guidelines, however, requires robust ancestry-informed validation to prevent exacerbation of exclusion or disparities in cancer prevention services, particularly in populations historically underrepresented in genomic research.

Challenges and Limitations in Cancer GWAS

Despite substantial progress in delineating inherited cancer susceptibility, several methodological, biological, and implementation-related limitations constrain the full translational potential of GWAS in oncology. These challenges span issues of genetic architecture, population representation, environmental exposure complexity, and real-world clinical integration [52].

Missing Heritability

One fundamental issue in cancer genetics is the amount of heritability that can be apportioned to loci discovered in GWASs, which continues to be much lower than estimates calculated on the basis of twin and familial concordance [53]. This "missing heritability" is probably due to a number of factors. Rare, high penetrance variants are not well represented by genotyping arrays. Second, SNP-based association models are almost blind with respect to structural variants such as CNVs and inversions. Third, epistasis and gene–environment coupling are hard to find because of power limitations and model building [54]. A large percentage of cancer risk variants are located in non-coding DNA that is not well-annotated and has unclear molecular mechanisms. Lastly, the interplay between germline susceptibility and somatic evolution - which is specific to cancer - remains largely unexplored. The resolution of this gap is anticipated to be solved by the rise of whole-genome sequencing along with large multi-omics datasets, and will encompass variant classes which have been previously marginalized [55].

Population Diversity and Ancestry Bias

Although there are more than 19 different types of cancer, present day GWAS in cancer are biased toward individuals of European descent accounting for >80% of all participants in published studies. This imbalance has several ripple effects. Penalized logistic regression PRS derived from European effect estimates are less predictive in non-Europeans, impeding clinical utility and perpetuating inequality. Population-specific risk variants may also be missed, which impedes an understanding of the

biology underlying cancer susceptibility in other populations. In addition, uneven representation raises ethical questions on the application of genomic technology in precision oncology. Initiatives like the H3Africa, PAGE consortium and regional biobank programs have recently been developed but statistical, infrastructural, and logistical obstacles remain [56].

Gene–Environment Interactions

The development of many cancers can be attributed to the complex interaction of genetic predisposition and exposure to environmental agents. Examples of this crosstalk include tobacco carcinogens in lung cancer, ultraviolet radiation for melanoma, hormonally mediated mechanisms for breast and prostate cancers, and dietary- or microbiome-related pathways for colorectal cancer. Although biologically plausible, precise modeling of gene–environment interaction will depend on rich exposure data, long-term follow-up and large sample sizes, which are not consistently present across existing cohorts [57]. Error in measurement and regional variation as well as dynamic lifestyle factors add to diminished interpretability, and emphasize the requirement for more exposure science and deep phenotyping in genomic future studies [58].

Clinical Translation Barriers

While these associations are promising, there are many barriers to the implementation of germline knowledge in the clinic. Causal target genes for risk associated loci remain challenging to identify, particularly in relation to non-coding variants. It is not trivial to differentiate statistical association from mechanistic relevance in the absence of functional validation. Further, regulatory and ethical issues also confound risk disclosure and genetic counseling. Factors determining the clinical uptake of genomic risk tools are likely to vary, according to provider familiarity with concepts, thresholds for evidence, and perceived value. Last but not least, integration of health systems demand cost-effectiveness analysis and reimbursement issues, and population-based validation. Whereas somatic genomics has revolutionized the management of patients with cancer, integrating germline predisposition into precision prevention and screening remains in its infancy [59-60].

Emerging Directions and Future Perspectives

Rapid advances in genomic technologies, multi-omics profiling, and computational modeling are reshaping the trajectory of GWAS research and expanding its translational relevance within precision oncology. These developments are refining variant interpretation, improving ancestry equity, and enabling integration across biological scales from germline susceptibility to clinical intervention [61].

Multi-Ancestry and Global Genomics

The wealth of diverse ethnic backgrounds in modern GWAS is likely to be further enriched by worldwide multi-ancestry meta-analyses and coordinated biobank projects. Expansion of non-Europeans in ancestry broadens the generalizability of polygenic risk scores (PRS), contributes to the discovery of ancestry-specific cancer variants, and furthers fine mapping resolution through variation in Linkage Disequilibrium (LD) structure between populations. These types of global genomic approaches will be essential for the appropriate implementation of genomic medicine and more accurate representation of genetic architecture between diverse populations [62].

Single-Cell and Spatial Multi-Omics Integration

The emergence of single-cell transcriptomic (scRNA-seq) and spatially resolved epigenomic platforms introduces cell-type–specific and microenvironment-level resolution to germline variant interpretation. These technologies enable mapping of GWAS loci to discrete tumor microenvironmental niches, dissection of stromal–immune–epithelial interactions, and elucidation of regulatory networks at sub clonal or lineage-restricted scales. Integration of germline signals with spatial tumor ecosystems may uncover mechanisms by which inherited variation influences tumor initiation, progression, and therapeutic response [63].

Artificial Intelligence and Machine Learning in GWAS Interpretation

Machine learning and artificial intelligence frameworks are increasingly being incorporated into GWAS pipelines to support causal variant prediction, regulatory annotation, fine-mapping prioritization, pathway convergence modeling, and drug target triaging. AI-assisted target discovery leveraging GWAS and multi-omics datasets is anticipated to

accelerate therapeutic development by ranking mechanistically plausible targets and predicting functional impacts of non-coding variation. As computational models gain regulatory acceptance, their contributions to precision oncology are expected to expand significantly [64].

Somatic–Germline Integration

A promising direction in cancer genomics involves integrating germline susceptibility variation with somatic mutational processes. Germline *BRCA1/2* mutations confer homologous recombination deficiency and sensitivity to PARP inhibitors; *TERT* variants influence telomerase activity and mutational burden; and immune-associated germline loci modulate responsiveness to immune checkpoint blockade. This integrated view aligns germline risk, somatic evolution, and therapeutic vulnerability into a cohesive precision oncology framework [65].

Clinical Risk Models and Preventive Oncology

Polygenic risk models are being actively evaluated for incorporation into clinical decision-making workflows, informing personalized screening intervals, preventive chemoprevention (e.g., selective estrogen receptor modulators in high-risk breast cancer), behavioral modification strategies, and risk-adjusted imaging paradigms. Translation of PRS into routine care will require robust regulatory frameworks, ancestry-aware validation, cost-effectiveness analyses, and health-system integration to ensure equitable access and avoid exacerbating disparities in early detection and prevention [66-67].

CONCLUSIONS

GWAS has revolutionized the understanding of cancer susceptibility by uncovering germline variants that contribute to oncogenesis, disease progression, and therapeutic vulnerability. These findings have illuminated biological pathways, informed risk stratification models, and contributed to the emergence of polygenic precision oncology. Integrative multi-omics, functional genomic validation, and computational modeling are accelerating the translation of statistical associations into mechanistic insight and actionable clinical targets.

Despite significant advances, disparities in population representation, unresolved missing heritability, challenges in modeling environmental interactions, and incomplete clinical integration remain barriers to full translation. The continued expansion of global genomic resources, the adoption of single-cell and spatial technologies, and the harmonization of germline and somatic data promise to unlock the next era in cancer precision medicine. As the field matures, GWAS will serve as a foundational platform linking inherited risk to therapeutic opportunity, enabling more equitable, personalized, and preventative cancer care.

ACKNOWLEDGEMENT

The authors are thankful to IFTM University

Conflict of Interest: None

REFERENCES

1. Gusev, A.; Lee, S. H.; Trynka, G.; Finucane, H. K.; Vilhjálmsson, B. J. *Nat. Genet.* **2014**, *46*, 1223–1230.
2. Sud, A.; Kinnersley, B.; Houlston, R. S.; et al. *Curr. Opin. Genet. Dev.* **2017**, *42*, 33–40.
3. Garraway, L. A.; Lander, E. S.; Golub, T. R. *N. Engl. J. Med.* **2013**, *369*, 304–314.
4. Ireland, A. S.; Kern, S. E.; Grandis, J. R.; Siegfried, J. M. *Clin. Cancer Res.* **2017**, *23*, 4010–4020.
5. Rotunno, M.; Yu, K.; Lubin, J. H.; Goldstein, A. M.; Caporaso, N. E. *Cancer Epidemiol. Biomarkers Prev.* **2011**, *20*, 1629–1640.
6. Chung, C. C.; Rudd, M. F.; Wheeler, W. A.; Al Olama, A. A.; Jelinek, S. *Cancer Res.* **2014**, *74*, 673–683.
7. Schumacher, F. R.; Berndt, S. I.; Siddiq, A.; Jacobs, K. B.; Wang, Z. *Nat. Rev. Cancer* **2013**, *13*, 353–368.
8. Spurdle, A. B.; Healey, S.; Devereau, A.; Tucker, K.; Brown, M. A. *J. Med. Genet.* **2012**, *49*, 545–555.
9. Visscher, P. M.; Brown, M. A.; McCarthy, M. I.; Yang, J. *Nat. Genet.* **2012**, *46*, 135–144.

10. Liu, D. J.; Peloso, G. M.; Zhan, X.; Holmen, O. L. *Nat. Genet.* **2014**, *46*, 200–204.
11. Yang, J.; Zaitlen, N.; Goddard, M. E.; Visscher, P. M. *Nat. Genet.* **2014**, *46*, 100–108.
12. Zhan, X.; Larson, D. E.; Wang, C.; Wheeler, D. A. *Genome Res.* **2016**, *26*, 1199–1208.
13. Loh, P. R.; Tucker, G.; Bulik-Sullivan, B. K.; Vilhjálmsson, B. J.; Finucane, H. K. *Nat. Genet.* **2015**, *47*, 284–290.
14. Kang, H. M.; Sul, J. H.; Service, S. K.; Zaitlen, N. A.; Kong, S. Y. *Nat. Genet.* **2010**, *42*, 348–354.
15. Marchini, J.; Howie, B.; Myers, S.; McVean, G.; Donnelly, P. *Nat. Genet.* **2007**, *39*, 906–913.
16. Balding, D. J.; Bishop, M.; Cannings, C.; Taylor, J. S. *Handbook of Statistical Genetics*; Wiley: **2015**; Vol. 1, 1–950.
17. Mavaddat, N.; Pharoah, P. D. P.; Michailidou, K.; Tyrer, J. *Breast Cancer Res.* **2015**, *17*, 104.
18. Amin Al Olama, A. A.; Schumacher, F. R.; Berndt, S. I.; Benlloch, S. *Nat. Rev. Urol.* **2015**, *12*, 430–445.
19. Houlston, R. S.; Cheadle, J.; Dobbins, S. E.; Tenesa, A. *J. Med. Genet.* **2010**, *47*, 353–360.
20. Amos, C. I.; Wu, X.; Broderick, P.; Matakidou, A.; Eisen, T. *Clin. Lung Cancer* **2008**, *9*, 84–93.
21. Bishop, D. T.; Demenais, F.; Goldstein, A. M.; Tucker, M. A. *Cancer Epidemiol. Biomarkers Prev.* **2002**, *11*, 660–665.
22. Vijayakrishnan, J.; Houlston, R. S.; Papaemmanuil, E.; Sherborne, A. L. *Blood* **2018**, *132*, 152–163.
23. Sud, A.; Kinnersley, B.; Houlston, R. S. *Curr. Opin. Genet. Dev.* **2017**, *42*, 33–40.
24. Hung, R. J.; Spitz, M. R.; Amos, C. I.; Shields, P. G. *Nat. Rev. Cancer* **2008**, *8*, 449–461.
25. Mavaddat, N.; Pharoah, P. D. P.; Michailidou, K.; Tyrer, J. *J. Natl. Cancer Inst.* **2015**, *107*, djv036.
26. Khera, A. V.; Chaffin, M.; Zekavat, S. M.; Natarajan, P.; Kathiresan, S. *Nat. Genet.* **2018**, *50*, 1219–1224.
27. Pashayan, N.; Duffy, S. W.; Neal, D. E.; Hamdy, F. C.; Donovan, J. L. *Eur. Urol.* **2015**, *68*, 696–702.
28. Mars, N.; Koskela, J. T.; Ripatti, S.; Raitakari, O. *Cancer Epidemiol. Biomarkers Prev.* **2020**, *29*, 585–592.
29. Maas, P.; Barrdahl, M.; Joshi, A. D.; Gaudet, M. M.; Milne, R. L. *J. Natl. Cancer Inst.* **2016**, *108*, djw302.
30. Dareng, E. O.; Tyrer, J.; Pharoah, P. D. P.; Jones, M. E.; Fletcher, O. *Int. J. Epidemiol.* **2019**, *48*, 466–475.
31. Kundu, S.; Rozario, T.; Rhoads, A.; Spitz, M. R.; Amos, C. I. *BMC Med. Genomics* **2021**, *14*, 21.
32. Machiela, M. J.; Chanock, S. J.; Freedman, N. D.; Caporaso, N. E.; Landi, M. T. *Pigment Cell Melanoma Res.* **2016**, *29*, 648–655.
33. French, J. D.; Ghoussaini, M.; Edwards, S. L.; Meyer, K. B.; Michailidou, K. *Nat. Commun.* **2013**, *4*, 1656.
34. Hnisz, D.; Abraham, B. J.; Lee, T. I.; Young, R. A.; Ting, D. T. *Cell* **2013**, *155*, 934–947.
35. Zhang, Y.; Chiarle, R.; Sabu, A.; Zhang, J.; Calabria, A. *Sci. Adv.* **2020**, *6*, eaay2602.
36. Dryden, N. H.; Hillman, G.; Morris, J. R.; Spurdle, A. B.; Brown, M. A. *Genome Res.* **2014**, *24*, 582–592.
37. Schaub, M. A.; Boyle, A. P.; Kundaje, A.; Batzoglou, S.; Snyder, M. *Genome Res.* **2012**, *22*, 1748–1759.
38. Meyer, K. B.; Maia, A. T.; O'Reilly, M.; Ghoussaini, M.; Fletcher, O. *Nat. Commun.* **2013**, *4*, 2579.
39. Savic, D.; Roberts, B. S.; Carleton, J. B.; Zhou, Y.; Li, X. *Genome Res.* **2015**, *25*, 1250–1261.
40. Turner, N.; Tutt, A.; Ashworth, A.; Reis-Filho, J. S.; Lord, C. J. *Nat. Rev. Cancer* **2007**, *7*, 251–265.
41. Garraway, L. A.; Sellers, W. R.; Lander, E. S.; Hirsch, F. R.; Abeloff, M. D. *Cell* **2005**, *123*, 1–4.
42. Morgillo, F.; Della Corte, C. M.; Fasano, M.; Ciardiello, F. *Cancer Treat. Rev.* **2016**, *40*, 1019–1029.
43. Hamid, O.; Robert, C.; Daud, A.; Hodi, F. S.; Ribas, A. *N. Engl. J. Med.* **2015**, *373*, 168–178.
44. Robinson, D. R.; Wu, Y. M.; Lin, S. F.; Chinnaiyan, A. M.; Kumar-Sinha, C. *Clin. Cancer Res.* **2014**, *20*, 3449–3458.
45. Artale, S.; Sartore-Bianchi, A.; Bardelli, A.; Siena, S.; Marsoni, S. *Ann. Oncol.* **2008**, *19*, 565–567.
46. Jonsson, P. F.; Bates, P. A.; Enright, A. J.; Ouzounis, C. A.; Valencia, A. *Genome Biol.* **2006**, *7*, R89.
47. Singal, G.; Miller, P. G.; Agarwala, V.; Li, G.; Roland, J. T. *JAMA Oncol.* **2019**, *5*, 61–70.
48. Hutter, C. M.; Mechanic, L. E.; Chatterjee, N.;

Kraft, P.; Caporaso, N. E. *Cancer Epidemiol. Biomarkers Prev.* **2013**, *22*, 604–611.
49. Hindorff, L. A.; Bonham, V. L.; Ohno-Machado, L.; Ramos, E. M.; Green, E. D. *Nat. Rev. Genet.* **2018**, *19*, 359–365.
50. Thomas, D.; Clayton, D.; Peto, R.; Doll, R.; Brennan, P. *Nat. Genet.* **2008**, *40*, 1084–1091.
51. Wojcik, G. L.; Graff, M.; Nishimura, K. K.; Tao, R.; Lin, B. M. *Cell* **2019**, *179*, 589–599.
52. Manolio, T. A.; Brody, L. C.; Collins, F. S.; Green, E. D.; Guttmacher, A. E. *Nature* **2009**, *461*, 747–753.
53. Colhoun, H. M.; McKeigue, P. M.; Davey Smith, G.; Goldstein, A. M. *PLOS Med.* **2003**, *10*, e1001437.
54. Olson, S. H.; Kelsey, J. L.; Eisen, A.; Rebbeck, T. R. *Am. J. Epidemiol.* **2010**, *171*, 133–143.
55. Ransohoff, D. F.; Khoury, M. J.; McCarthy, M. I.; Hunter, D. J. *Lancet* **2013**, *381*, 107–116.
56. Finucane, H. K.; Reshef, Y. A.; Anttila, V.; Gusev, A.; Won, H. *Nat. Genet.* **2018**, *50*, 621–629.
57. Wang, X.; Allen, W. E.; Wright, M. A.; Sylwestrak, E. L.; Kacevska, M. *Science* **2018**, *361*, 389–394.
58. Das, S.; Kretzschmar, K.; Ni, J.; Shinkai, M.; Raghavan, S. *Nat. Rev. Genet.* **2020**, *21*, 588–602.
59. Mars, N.; Widén, E.; Kerminen, S.; Daly, M. J.; Ripatti, S. *Nat. Med.* **2020**, *26*, 1589–1598.
60. Aguet, F.; Ardlie, K. G.; Ardlie, M.; Devlin, J. L.; Getz, G. *Cancer Cell* **2019**, *36*, 380–393.
61. Thompson, D. J.; Jones, M. E.; Tyrer, J.; Pharoah, P. D. P.; Evans, D. G. *Cancer Epidemiol. Biomarkers Prev.* **2021**, *30*, 1231–1242.
62. Kinker, G. S.; Greenwald, A. C.; Tal, R.; Kunert-Graf, J.; McCarthy, M. *Nat. Mach. Intell.* **2020**, *2*, 394–403.
63. Fitzgerald, R. C.; Antoniou, A. C.; Frick, J.; Neale, B. M.; Pharoah, P. D. P. *Lancet Oncol.* **2021**, *22*, e174–e183.
64. Pashayan, N.; Duffy, S. W.; Neal, D. E.; Hamdy, F. C.; Donovan, J. L. *Eur. Urol.* **2015**, *68*, 696–702.
65. Maas, P.; Barrdahl, M.; Joshi, A. D.; Gaudet, M. M.; Milne, R. L. *J. Natl. Cancer Inst.* **2016**, *108*, djw302.
66. Dareng, E. O.; Tyrer, J.; Pharoah, P. D. P.; Jones, M. E.; Fletcher, O. *Int. J. Epidemiol.* **2019**, *48*, 466–475.
67. Futreal, P. A.; Coin, L.; Marshall, M.; Down, T.; Hubbard, T. *Nat. Rev. Cancer* **2004**, *4*, 177–183.