


Review

Towards the Genetic Architecture of Complex Gene Expression Traits: Challenges and Prospects for eQTL Mapping in Humans

Chaeyoung Lee 

Department of Bioinformatics and Life Science, Soongsil University, 369 Sangdo-ro, Dongjak-gu, Seoul 06978, Korea; clee@ssu.ac.kr

Abstract: The discovery of expression quantitative trait loci (eQTLs) and their target genes (eGenes) has not only compensated for the limitations of genome-wide association studies for complex phenotypes but has also provided a basis for predicting gene expression. Efforts have been made to develop analytical methods in statistical genetics, a key discipline in eQTL analysis. In particular, mixed model- and deep learning-based analytical methods have been extremely beneficial in mapping eQTLs and predicting gene expression. Nevertheless, we still face many challenges associated with eQTL discovery. Here, we discuss two key aspects of these challenges: 1, the complexity of eTraits with various factors such as polygenicity and epistasis and 2, the voluminous work required for various types of eQTL profiles. The properties and prospects of statistical methods, including the mixed model method, Bayesian inference, the deep learning method, and the integration method, are presented as future directions for eQTL discovery. This review will help expedite the design and use of efficient methods for eQTL discovery and eTrait prediction.

Keywords: complex phenotype; expression quantitative trait locus; regulation of gene expression; statistical genetics; target gene



Citation: Lee, C. Towards the Genetic Architecture of Complex Gene Expression Traits: Challenges and Prospects for eQTL Mapping in Humans. *Genes* **2022**, *13*, 235. <https://doi.org/10.3390/genes13020235>

Academic Editor: Albert Jordan

Received: 20 December 2021

Accepted: 25 January 2022

Published: 26 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Despite considerable progress in the 20 years since the completion of the Human Genome Project, a complete understanding of the functions of the human genome sequence, comprising >3 billion base pairs, remains still elusive. Genome-wide association studies (GWASs) have identified thousands of quantitative trait loci (QTLs) as association signals for human complex traits. Because the QTLs have mostly been mapped in the noncoding regions of the human genome, knowledge on their underlying genetic mechanisms are limited. Indeed, candidate causal genes corresponding to QTLs are often falsely assigned because the gene closest to a given QTL is suspected as a target. Studies on expression QTL (eQTL) mapping have aimed to identify their target genes (eGenes) and understand the genetic mechanisms underlying their expression (eTraits) and phenotypic traits. However, these studies are just the tip of the iceberg when considering the eQTL studies required in the future. Herein, human eQTL mapping will be discussed, focusing on challenges and prospects. Detailed theories and equations of methodologies will not be discussed in depth in order to appeal to a wider range of geneticists and fields, including those with practical applications.

2. Ground-Breaking Approach for eQTL Discovery and eTrait Prediction

Various approaches have been used in the past decade to develop analytical models and methods for eQTL discovery. In particular, efficient methods have been intensively studied along with advances in technology, from hybridization-based microarray to next-generation sequencing (NGS)-based RNA-seq. Two landmark approaches will be presented herein as examples.

The mixed model method for genetic analysis was originally devised using pedigree information to explain the polygenic effects treated as random effects in the analytical model [1]. Here, the random effects infer that their parameters are assumed to have random variables (for details, see [2]). The use of the model was then extended to the genome-wide identification of loci associated with phenotypic traits and further with eTraits, explaining the random polygenic effects with genome-wide genotypic information instead of pedigree information [2]. In other words, the mixed model method enables us to discover eQTLs, not only via linkage analysis, but also via association analysis. More importantly, it helps rationally and efficiently correct for population structure and elucidate polygenic effects. As a result, it has greatly contributed to accurate eQTL discovery, and avoiding spurious eQTLs, which is considered one of the main concerns in eQTL studies. This method provides a portion of eTrait variance explained by eQTLs and partitions the portion into subportions according to the classified eQTLs; for example, we can separately estimate the portions for cis-acting eQTLs and trans-acting eQTLs. The mixed models can be used to simultaneously analyse multiple eTraits using genome-wide efficient mixed model association (GEMMA) [3]. Using efficient algorithms such as Cholesky decomposition (a method to produce the product of a lower triangular matrix and its transpose, equivalently to a positive-definite matrix) and Gauss–Seidel iteration (a method to iteratively solve linear equations by successive displacement), eTraits can be predicted with a mixed model framework without incurring a heavy computing burden of inverting a huge matrix [4]. In addition, the mixed model has been used in integrative analyses of eTraits and phenotypic traits to prioritize causal genes of phenotypic variation [5].

More recently, deep learning has been garnering increased popularity in predicting gene expression. Deep learning does not have a statistical framework, unlike other methods, thus, it offers great potential at gaining results that are difficult to obtain from other eQTL studies. Moreover, the rapid growth of relevant data will vastly increase the contribution of deep learning to eQTL studies. Deep learning generally refers to supervised or unsupervised learning using advanced artificial neural networks, often known as deep neural networks. Supervised learning represents predicting categorical or continuous variables using a training data set, whereas unsupervised learning represents studying intrinsic patterns and clustering them based on pattern similarity. In this respect, supervised deep learning is a good choice for predicting gene expression. In particular, the large amount of data generated by high-throughput techniques can provide an absolute condition for the use of supervised deep learning. For example, convolutional neural networks are feed-forward deep neural networks in which every unit in a layer is connected to all the units in the previous layer without forming any cycle of the unit connections. Convolutional neural networks have been used to predict gene expression using proximal promoter sequences and distal enhancer sequences obtained by HiChIP [6]. Recurrent neural networks are deep neural networks in which connections between units have a cyclic structure; these networks have been used to predict differences in gene expression between two cell types using histone modification profiles [7].

These approaches have been applied to eQTL studies and have contributed greatly to the discovery of eQTLs for specific genes and offered further insights into general regulation mechanisms of gene expression (Table 1). Often, these approaches are customized to fit specific study objectives. For example, a mixed model incorporating ancestry effects was applied to identify eQTLs in multi-ethnic, or admixed, populations to avoid confounding with the eQTL effect [8]. Deep learning was also applied using a hierarchical Bayesian model with the posterior of parameters for different tasks such as transcription factor binding, chromatin accessibility, and histone marks [9]. This is useful to identify causal variants among candidate nucleotide sequence variants in a strong linkage.

Table 1. Examples of eQTL studies and reviews using a mixed model or a deep learning.

eQTL	Methods	Results	Reference
Mixed model			
cis-neQTL	GEMMA	Inflammation-dependent cis-neQTLs in intestine	[10]
neQTL	GaLA-QTLM	neQTL mapping with ancestry data in multi-ethnic population	[8]
neQTL	MM with pedigree-based covariance matrix	Cell type-specific neQTLs in brain and blood for Alzheimer's disease	[11]
neQTL	BOLT-LMM, GEMMA	Cis-/trans-neQTLs in peripheral blood and their contribution to heritability	[12]
cis-neQTL	StructLMM	Cell-context interaction with environmental variables	[13]
neQTL, rQTL, pQTL	AIREML, BLUP	Efficient translational control of pQTLs for ribosomal protein genes	[14]
neQTL, rQTL, pQTL	AIREML	Inclusion of optimal number of eQTLs in constructing polygenic covariance matrix	[15]
General	REML and others	Review on frequentist mixed model methodology	[2]
General	Gibbs sampling, MCMC	Review on Bayesian mixed model methodology	[16]
Deep learning			
meQTL	CpGenie (CNN)	Prediction of the allele-specific impact of nucleotide variants on proximal CpG methylation	[17]
cQTL	DeepHiC (CNN)	Functional prediction of nucleotide substitution on chromatin interaction using Hi-C data	[18]
neQTL, dsQTL, atacQTL	MtBNN (CNN, RNN)	Incorporating a Bayesian approach to assessing functional impact of non-coding variants	[9]
General	DNN	Introduction to deep learning for genomics covering more than eQTL mapping	[19]

neQTL, narrow-sense expression quantitative trait locus; rQTL, ribosome occupancy QTL; pQTL, protein abundance QTL; meQTL, methylation QTL; cQTL, chromatin interaction QTL; dsQTL, DNase sensitivity QTL; atacQTL, assay for transposase accessible chromatin QTL; GEMMA, genome-wide efficient mixed-model association; MM, mixed model; AIREML, average information restricted maximum likelihood; BLUP, best linear unbiased prediction; MCMC, Markov chain Monte Carlo; CNN, convolutional neural network; RNN, recurrent neural network; DNN, deep neural network.

3. Complexity of eTraits

eQTL discovery is challenging primarily due to its intrinsic complexity. The expression of a single gene in a cell is regulated by various gene products. Each of the various regulating genes in a cell is also regulated by various gene products. This hierarchical regulatory mechanism suggests that many genes expressed in a cell are generally involved in the expression of a single gene (Figure 1), thus having a relatively small effect as an eQTL.

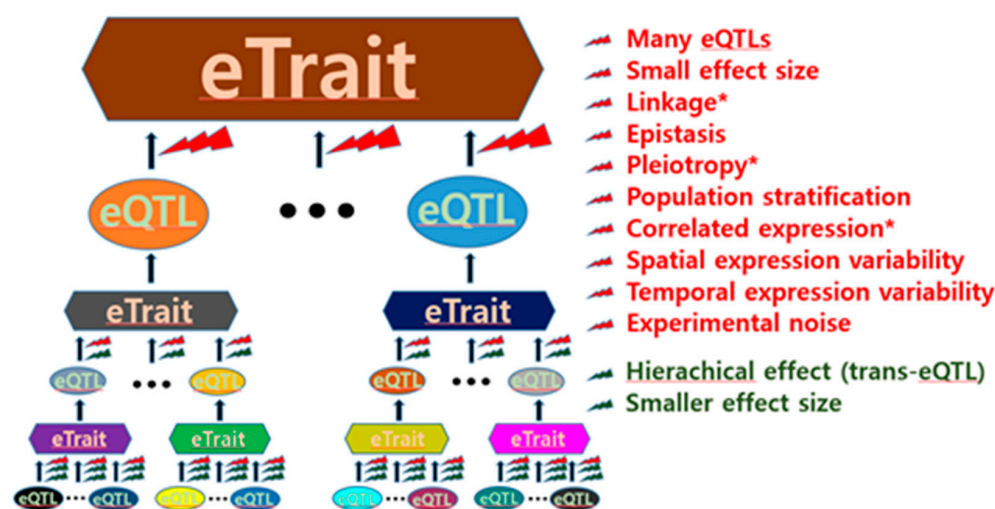


Figure 1. Obstacles to expression quantitative trait locus (eQTL) identification. General obstacles to both cis-eQTL and trans-eQTL identification are presented in red, whereas obstacles to only trans-eQTL identification are presented in dark green. The asterisk indicates obstacles that increase the difficulty of inference of causality.

In particular, the effect sizes of trans-eQTLs tend to decrease depending on the degree of trans-acting indirection (Figure 1); therefore, geneticists have experienced difficulties in finding trans-eQTLs. Although the cell environments determined by various conditions are likely genetic, identifying eQTLs can be formidably difficult. Thus, the expressional genetic architecture of many single genes is considerably complex. This was supported by a study where no single eQTL determined mRNA transcription, ribosomal occupancy, or protein abundance of any gene [15]. Hence, the expression of most genes has been referred to as “complex eTraits” hereafter in this article. The complexity of the regulatory function of eQTLs in gene expression can also be attributed to various other factors, as shown in Figure 1. Thus, these factors can act as obstacles to eQTL identification.

Linkages can help researchers find causal sequence variants linked to representative variants associated with gene expression. However, the effect size can be influenced by multiple functional variants in a linkage, often resulting in false positive/negative eQTLs. Many geneticists are therefore reluctant to search for functional sequence variants in the major histocompatibility complex region, where such characteristics resonate extremely well.

Epistasis among eQTLs is infrequently examined in eQTL mapping because it requires expensive computing power. Epistasis is a natural mechanism that contributes to gene expression regulation by binding DNA and DNA-derived substances (RNA and proteins). For example, eukaryotic transcription is usually initiated by the interplay among transcription factors, activators, mediators, RNA polymerase, enhancers, and proximal or distal promoters.

Population stratification is also ignored in many eQTL analyses, even though this reduces the accuracy of eQTL mapping and can result in spurious eQTLs [2,20,21].

Spatial expression variability ranges from variations among nearby cells to those among organs. In particular, cellular resolution of eQTLs in the brain might be critical for understanding transcriptional heterogeneity of pyramidal cells in addition to regional functions/misfunctions [22,23]. Likewise, temporal expression variability can result from short- to long-term differences, including variability arising from changes in the external environment.

The aforementioned factors prevent eQTL mapping. In particular, eQTL discovery can be difficult if there are small effect sizes because multiple testing is required to reduce false positives. Inference of causality is also interfered by obstacles such as linkage, pleiotropy, and correlated expression (Figure 1). Further, it is not easy to identify functional nucleotide variants in regions with strong linkages. For example, the human leukocyte antigen (HLA)

complex that spans 3.6 megabase pairs covering 224 genes on the short arm of chromosome 6 has many strong linkage disequilibrium (LD) blocks, and the LD blocks are often linked to nucleotide variants outside within the HLA complex [12]. Thus, careful attention is needed to interpret functionality of nucleotide variants, such a complex region in strong linkages.

4. Voluminous Work of eQTL Mapping

Owing to the considerable complexity of the eTraits, another critical issue regarding eQTL mapping is the construction of voluminous studies producing a vast amount of data. A genome-wide eQTL analysis produces a genetic profile of regulatory signals for a single gene. In modern studies, eQTL profiles are simultaneously produced for all transcriptome-wide genes. Moreover, integrated analysis of eQTL mapping with GWAS may reveal the regulatory mechanisms underlying phenotypic traits; this is known as a transcriptome-wide association study (TWAS). Profiles can be extended to various types of eQTLs—additive, dominant, recessive, haplotypic, and epistatic. Variations within these profiles can be produced by spatial and temporal gene expression and by various expression molecules across expression processes. Thus, eQTLs are characterized by the expression molecules used for their identification and categorized by expression processes such as transcription and translation (Table 2). The molecular layers can be further extended to DNA or chromatin modification, chromatin interaction, and metabolism. A significant number of profiles beyond those mentioned above are needed in order to understand complexity. For example, eQTL profiles of 13 brain parts are now available from the Common Fund’s Genotype-Tissue Expression Program [24], and more situation-specific single-cell products are becoming available. Profiles of disease states are also necessary to reveal the genetic aetiology; these profiles can be further generated on a population basis, as considered in GWAS for complex diseases.

Table 2. Types of expression quantitative trait loci (eQTLs) associated with various molecular layers.

Name	Abbrev.	Molecular Phenotype Associated with eQTL (Method)	Ref.
Chromatin modification			
chromatin accessibility QTL	caQTL	Active and potential regulatory DNA elements, e.g., dsQTL (DNase-seq, ATAC-seq)	[25–27]
methylation QTL	meQTL	DNA methylation for altering chromatin structure, mainly in regulatory regions such as promoters and intron–exon boundaries (ChIP-seq)	[28]
histone QTL	hQTL	Magnitude of histone post-translational modifications for chromosomal packaging, e.g., H3K4me3 for promoters, H3K4me1 for enhancers, H3K27ac for promoters and enhancers (ChIP-seq)	[27,29,30]
TF binding QTL	bQTL	Transcription factor binding, e.g., NF- κ B, PU.1/Spi1, Stat1, JunD, and Pou2f1/Oct1 (ChIP-seq)	[30,31]
Chromatin interaction			
promoter interacting eQTL	pieQTL	eQTLs overlapping active cis-regulatory elements that interact with their target gene promoters (HiChIP)	[32]
chromatin interaction QTL	cQTL	Allelic differences of chromatin interactions between two homologous chromosomes mediated by CTCF and RNAPII (ChIA-PET)	[33]
promoter enhancer interaction QTL	peQTL	Allele-specific RNAPII-mediated chromatin interactions with phased transcript (ChIA-PET)	[33]

Table 2. Cont.

Name	Abbrev.	Molecular Phenotype Associated with eQTL (Method)	Ref.
Transcription			
narrow-sense eQTL	neQTL	Gene expression level as the sum of all transcripts of each gene. We used the “neQTL” to differentiate this from eQTL, a generic term for all kinds (RNA-seq)	[34,35]
miRNA eQTL	miR-eQTL	Expression level of miRNA for post-transcriptional and translational regulation (small RNA-seq)	[34,36]
lncRNA eQTL	lncR-eQTL	Expression level of lncRNA for transcriptional, post-transcriptional, and epigenetic regulation (RNA-seq)	[34,37,38]
circRNA-eQTL	Circ-eQTL	Expression level of circRNA for sequestration of miRNAs/proteins, splicing interference, and transcriptional and translational regulation (RNA-seq)	[39,40]
response eQTL	reQTL	Transcriptomic response to external stimuli (RNA-seq)	[41,42]
repeat eQTL	repeat-eQTL	Retrotransposon-derived repeat element as a source for evolution of new transcripts (RNA-seq)	[34]
splicing QTL	sQTL	Relative abundance of the transcript isoforms of a gene or the intron excision ratios of an intron cluster for regulation of alternative splicing (RNA-seq)	[34,43]
transcript ratio QTL	trQTL	Ratio of each transcript to the total gene expression for transcript usage, splicing, and transcript structure (RNA-seq)	[34]
Allele specific expression QTL	aseQTL	Transcription differences between two different haplotypes in a heterozygous individual (RNA-seq)	[34]
poly(A) ratio QTL	apaQTL	Alternative polyadenylation for mRNA stability and translation efficiency (RNA-seq)	[44]
RNA editing QTL	edQTL	RNA editing level for post-transcriptional processes such as RNA splicing, localization, stability, and translational efficiency (RNA-seq)	[45]
m ⁶ A QTL	m ⁶ A-QTL	N ⁶ -methyladenosine level in mRNA transcript for mRNA processing. (m ⁶ A-seq)	[46]
RNA synthesis rate QTL	rsQTL	Transcription rates (4sU-seq)	[47]
RNA decay QTL	rdQTL	mRNA decay rates for modulating steady-state transcript levels (RNA-seq)	[48]
transcription initiation QTL	tiQTL	Activity of transcribed transcriptional regulatory elements (tTREs) in promoter and enhancer region (PRO-seq)	[49]
directional initiation QTL	diQTL	Directionality of divergent bidirectional transcription at tTREs using log ratio of plus strand reads over minus-strand reads (PRO-seq)	[49]
Translation			
ribosome occupancy QTL	rQTL	Ribosome occupancy for translational regulation and translation efficiency (Ribo-seq)	[50]
protein abundance QTL	pQTL	Protein expression level for post-transcriptional regulation (mass spectrometry)	[50]

Table 2. Cont.

Name	Abbrev.	Molecular Phenotype Associated with eQTL (Method)	Ref.
Metabolism			
Metabolite QTL	mQTL	Small endogenous molecules or metabolites that reflects the dynamic response to physiological, pathophysiological, and/or developmental stimuli (NMR or mass spectroscopy)	[51,52]
microbiome QTL	mbQTL	Microbial composition in multiple host tissues such as gut and skin (16S rRNA and ITS sequencing)	[53,54]

5. Miscellaneous Issues and Prospects

The issues regarding the qualities and quantities of eQTL maps addressed in this article represent important challenges in statistical genetics. Comprehending the complexities of eTraits is a challenge that transcends the currently available mapping facilities and tools. Recently, there have been substantial improvements in computing environments, which have enabled the analysis of considerable amounts of genomic data, and such developments are constantly accelerating. However, the development of high-resolution eQTL maps requires further marked developments in computing memory and speed. Even now, limitations exist in mixed models that can explain background polygenic effects during eQTL mapping [2]. This is even more problematic when Bayesian inference, which requires intensive computing to marginalize multidimensional joint posterior distribution through the Markov chain Monte Carlo method, is used [16]. Limitations in computing power also render geneticists reluctant to conduct GWASs into the interactions among eQTLs [55]. The computing burden exponentially increases as the interaction order increases. Moreover, while the number of nucleotide variants to be tested increases markedly, interactive variants with low minor allele frequency also likely increase. Identifying higher-order epistasis among eQTLs exhaustively is barely conceivable. Therefore, interactive eQTL mapping relies almost exclusively on experimental studies that may increase the likelihood of interactive functions. For example, a cis-acting eQTL study in which a HiChIP assay for the histone modification of H3K27ac was employed to identify cis-regulatory elements that interacted with the promoters of their target genes, for five types of immune cells, was conducted [32]. More efficient analytical algorithms, methods, and/or designs are required to identify a data-specific workaround, increase the accuracy of eQTL mapping, and decrease the required computing power.

In addition to the abovementioned ones, many other analytical methods have greatly contributed to eQTL mapping for complex eTraits. Nevertheless, the fact that it will likely be impossible to create an ideal eQTL map remains; this is attributed to the great complexity of eQTL maps, as shown in Figure 1. Furthermore, an eQTL map is a moving target that interacts with various environments and is driven by continuous evolution. Innovative progress should be achievable owing to advances in statistical genetics and other disciplines. Future advances in statistical methods for eQTL mapping are highly anticipated, with modifications and extensions of ground-breaking methodologies such as the mixed model, Bayesian inference, deep learning method, and integration method. These provide a critical basis in terms of reducing spurious eQTLs, integrating multiple data, or identifying epistatic eQTLs.

Spurious eQTLs are largely attributed to genetic, environmental, and experimental factors that are ignored or uncounted from the study. Genetic factors include major eQTLs, cis-eQTLs, and trans-eQTLs. A method that can directly explain these genetic factors is the mixed model. As discussed above, the mixed model uses genomic covariance among individuals to explain polygenic effects. The mixed model method reduces residuals unexplained by the analytical model and increases the accuracy of eQTLs [20,56]. It is considered the best statistical tool needed to go one step further to understand the genetic architecture of complex eTraits in the most direct way, without losing any degree of

freedom. However, including excessive loci in the construction of genomic covariance or employing the infinitesimal model can lead to spurious eQTLs [56]. The importance of an optimal genomic covariance structure in the mixed model should be stressed. Additional modifications are required in the mixed model for multiple eQTLs with major effects. An example is the multi-locus mixed model analysis in which the analytical model may include additional cofactors by stepwise regression of forward inclusion and backward exclusion [57]. Appropriate adjustment and filtration are required to reduce the errors generated by ignoring environmental and experimental factors. In addition, we need to be wary of eQTLs that can be falsely generated via correlated expression between genes [58].

Expensive computing costs to invert a large matrix is a stumbling block in the application of the mixed model method to eQTL mapping, and this block has been encountered in many studies. In particular, because a continuous increase in the sample size in eQTL studies is expected with the development of sequencing technologies, methods for solving or avoiding this problem are required. For example, a reduced animal model equivalent to an animal model has been widely used in genetic analyses for animal breeding; these analyses use a mixed model with pedigree information to reduce the size of the numerator relationship matrix and the number of equations to be solved [59]. Another concern is the violation of the assumption that known variance components are required for the best linear unbiased estimator (BLUE) of the eQTL effect [2]. When we estimate the fixed eQTL effect using polygenic and residual variance component estimates rather than true values, no penalties are imposed, resulting in increased error variability.

Bayesian inference has several good properties that make it suitable for eQTL mapping. In particular, the Bayesian approach incorporated with a mixed model may overcome the non-BLUE problem raised from the frequentist approach. For example, the Bayesian approach implemented with Gibbs sampling yields polygenic and residual variance components based on polygenic effects and residuals of individuals at every round of the Gibbs chain, finally providing samples of the eQTL effect. Thus, we can directly obtain various point estimates without assuming any distribution. This provides an empirical Bayes estimate of the eQTL effect, corresponding to the BLUE [16]. Nevertheless, mixed model-based Bayesian inference has barely been applied to eQTL analysis. This might be largely attributable to the intensive computing required for the Gibbs chain, resulting in computing costs that would be more expensive than those for the frequentist approach. Hamiltonian Monte Carlo is an efficient Markov chain Monte Carlo method used for the quick convergence of stationary probability distributions; it works by reducing autocorrelation between consecutive samples, which can greatly reduce the computing burden [60].

Deep learning is not based on statistical properties; however, it can be used to approach the complexity of eTraits. Deep learning has an advantage in eQTL studies using large amount of data, and it is also a niche approach that is difficult to be used within a statistical framework. However, in the case of deep learning, close attention should be paid to possible issues, such as parameter overfitting, data imbalance, and subtle variances in input data. This is critical to reducing noise and, thus, to unleashing the enormous potential of eQTLs [19]. To apply deep learning, it is also necessary to try to supplement the shortcomings of the difficulty in interpreting the results due to its black box nature.

Various integration algorithms for gene expression analysis have been developed by simultaneously dealing with data from multiple independent studies with comparable designs (horizontal integration) [61] and from multiple molecular measurements on the same subjects (vertical integration) [62]. An integrated analysis combining vertical and horizontal integration is also expected. Another integration analysis helps identify eQTLs simultaneously using multiple tissues, and the analysis may show improved accuracy as well as heterogeneity of eQTLs by tissues. An example is the hierarchical Bayesian model called MT-eQTL used for multiple tissue cis-eQTL analysis [63]. Moreover, further attempts will be made for more diverse types of multidimensional integrated analyses. An example is the TWAS, in which eQTLs and GWAS signals are integrated to identify genes associated with a complex phenotype [64]. TWAS has been extended to the identification

of pathways associated with a complex phenotype by aggregating functional annotations across the genes [65]. That is, gene set enrichment analysis can be integrated into gene-based TWAS to generate the pathway-based TWAS. A significant contribution of a pathway to complex phenotypes may be revealed from the set of genes with a small effect size on the pathway [66]. Many extended setups pose challenges because the total number of configurations is likely to grow exponentially, making implementation excessively slow and expensive.

Finally, for researchers wanting to apply analytical methods to eQTL analysis, finding a software with an efficient method suitable for the study purpose is critical to reducing computational costs and appropriately inferencing the results. Analytical methods need to be compared and characterized with various conditions to provide updated and proper guidance, and researchers are desperately in need of a thorough review of analytical methods and software prior to data analysis. For example, if a study on eQTL epistasis is planned, considering the eQTL accuracy and time efficiency, a different optimal method may be chosen according to the number of loci in epistasis and the covariates included in the analytical model. Furthermore, the computing environment should be considered.

This review has primarily focused on analytical methods from the perspective of statistical genetics. Along with the improvement of methods, increasing sample size and advancing sequencing technology are essential to accelerate the development of eQTL studies. Strategies that prioritize increasing sample size may include establishing standards for data and expanding shareable data repositories. Sequencing techniques such as Hi-C, Capture-C, and 3C-seq may improve eQTL studies for chromatin interactions [67]. Finally, advances in fundamental technologies that can be used more broadly could further change the paradigm of the field, as previously shown by microarrays and next generation sequencing.

6. Conclusions

Statistical genetics is a key discipline in mapping eQTLs and predicting eTraits. Now, it is time to accelerate the development of analytical methods for solving problems or mitigating limitations. To date, eQTL studies have been strongly biased toward the discovery of cis-eQTLs. This is largely attributable to the easy biological interpretation on direct regulation, the relatively large effect size, and the reduced number of variants to be tested within a certain region. However, we should also make efforts toward developing efficient methods to overcome the limitations of trans-eQTL discovery. Thirty-seven percent of GWAS signals ($p \leq 5 \times 10^{-8}$) for the phenotypic traits corresponded to trans-eQTLs that were recently reported using blood samples of 31,684 individuals by the eQTLGen consortium [68]. This proportion will increase with the trans-eQTLs of other tissues and with a larger sample size. Comparing this large-scale analysis to the second-largest analysis for blood eQTLs, where 8% of the GWAS signals were found to be trans-eQTLs using blood samples of 5311 individuals [69], a considerable advantage of larger sample sizes is expected. As great amounts of data are accumulated in the future, the concerns regarding trans-eQTL discovery will reduce. The trans-eQTL profile may reflect a characteristic of tissue-specific regulation that differentiates body parts and developing efficient analytical methods can increase the accuracy.

In conclusion, the genetic architecture of complex eTraits, produced at a higher resolution through a more rational analysis of enormous amounts of gene expression data, will contribute to the understanding of the genetic architecture of complex diseases; this will constitute an important basis in precision medicine.

Funding: This work was supported by a grant from the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (No. NRF-2018R1A2B6004867).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The author would like to thank three anonymous reviewers for their helpful comments on the first version of the manuscript and Jihye Ryu for collecting data on the various types of eQTLs outlined in Table 2.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Henderson, C.R. Estimation of Variance and Covariance Components. *Biometrics* **1953**, *9*, 226–252. [[CrossRef](#)]
2. Lee, C. Genome-Wide Expression Quantitative Trait Loci Analysis Using Mixed Models. *Front. Genet.* **2018**, *9*, 341. [[CrossRef](#)] [[PubMed](#)]
3. Zhou, X.; Stephens, M. Efficient Multivariate Linear Mixed Model Algorithms for Genome-Wide Association Studies. *Nat. Methods* **2014**, *11*, 407–409. [[CrossRef](#)] [[PubMed](#)]
4. Lee, C. Best Linear Unbiased Prediction of Individual Polygenic Susceptibility to Sporadic Vascular Dementia. *J. Alzheimers Dis.* **2016**, *53*, 1115–1119. [[CrossRef](#)]
5. Yang, Y.; Shi, X.; Jiao, Y.; Huang, J.; Chen, M.; Zhou, X.; Sun, L.; Lin, X.; Yang, C.; Liu, J. CoMM-S2: A Collaborative Mixed Model Using Summary Statistics in Transcriptome-Wide Association Studies. *Bioinformatics* **2020**, *36*, 2009–2016. [[CrossRef](#)]
6. Zeng, W.; Wang, Y.; Jiang, R. Integrating Distal and Proximal Information to Predict Gene Expression via a Densely Connected Convolutional Neural Network. *Bioinformatics* **2020**, *36*, 496–503. [[CrossRef](#)]
7. Sekhon, A.; Singh, R.; Qi, Y. DeepDiff: DEEP-Learning for Predicting DIFFerential Gene Expression From Histone Modifications. *Bioinformatics* **2018**, *34*, i891–i900. [[CrossRef](#)]
8. Zhong, Y.; Perera, M.A.; Gamazon, E.R. On Using Local Ancestry to Characterize the Genetic Architecture of Human Traits: Genetic Regulation of Gene Expression in Multiethnic or Admixed Populations. *Am. J. Hum. Genet.* **2019**, *104*, 1097–1115. [[CrossRef](#)]
9. Xu, C.; Liu, Q.; Zhou, J.; Xie, M.; Feng, J.; Jiang, T. Quantifying Functional Impact of Non-coding Variants with Multi-task Bayesian Neural Network. *Bioinformatics* **2020**, *36*, 1397–1404. [[CrossRef](#)]
10. Hu, S.; Uniken Venema, W.T.; Westra, H.J.; Vich Vila, A.; Barbieri, R.; Voskuil, M.D.; Blokzijl, T.; Jansen, B.H.; Li, Y.; Daly, M.J.; et al. Inflammation Status Modulates the Effect of Host Genetic Variation on Intestinal Gene Expression in Inflammatory Bowel Disease. *Nat. Commun.* **2021**, *12*, 1122. [[CrossRef](#)]
11. Patel, D.; Zhang, X.; Farrell, J.J.; Chung, J.; Stein, T.D.; Lunetta, K.L.; Farrer, L.A. Cell-type-specific Expression Quantitative Trait Loci Associated with Alzheimer Disease in Blood and Brain Tissue. *Transl. Psychiatry* **2021**, *11*, 250. [[CrossRef](#)] [[PubMed](#)]
12. Lloyd-Jones, L.R.; Holloway, A.; McRae, A.; Yang, J.; Small, K.; Zhao, J.; Zeng, B.; Bakshi, A.; Metspalu, A.; Dermitzakis, M.; et al. The Genetic Architecture of Gene Expression in Peripheral Blood. *Am. J. Hum. Genet.* **2017**, *100*, 228–237. [[CrossRef](#)] [[PubMed](#)]
13. Moore, R.; Casale, F.P.; Jan Bonder, M.; Horta, D.; BIOS Consortium; Franke, L.; Barroso, I.; Stegle, O. A Linear Mixed-Model Approach to Study Multivariate Gene-Environment Interactions. *Nat. Genet.* **2019**, *51*, 180–186. [[CrossRef](#)] [[PubMed](#)]
14. Ryu, J.; Lee, C. Regulatory Nucleotide Sequence Signals for Expression of the Genes Encoding Ribosomal Proteins. *Front. Genet.* **2020**, *11*, 501. [[CrossRef](#)] [[PubMed](#)]
15. Ryu, J.; Lee, C. Underestimation of Heritability across the Molecular Layers of the Gene Expression Process. *Processes* **2021**, *9*, 2144. [[CrossRef](#)]
16. Lee, C. Bayesian Inference for Mixed Model-Based Genome-Wide Analysis of Expression Quantitative Trait Loci by Gibbs Sampling. *Front. Genet.* **2019**, *10*, 199. [[CrossRef](#)]
17. Zeng, H.; Gifford, D.K. Predicting the Impact of Non-coding Variants on DNA Methylation. *Nucleic Acids Res.* **2017**, *45*, e99. [[CrossRef](#)]
18. Meng, X.H.; Xiao, H.M.; Deng, H.W. Combining Artificial Intelligence: Deep Learning with Hi-C Data to Predict the Functional Effects of Non-coding variants. *Bioinformatics* **2021**, *37*, 1339–1344. [[CrossRef](#)]
19. Zou, J.; Huss, M.; Abid, A.; Mohammadi, P.; Torkamani, A.; Telenti, A. A Primer on Deep Learning in Genomics. *Nat. Genet.* **2019**, *51*, 12–18. [[CrossRef](#)]
20. Shin, J.; Lee, C. A Mixed Model Reduces Spurious Genetic Associations Produced by Population Stratification in Genome-Wide Association Studies. *Genomics* **2015**, *105*, 191–196. [[CrossRef](#)]
21. Zeng, B.; Gibson, G. PolyQTL: Bayesian Multiple eQTL Detection With Control for Population Structure and Sample Relatedness. *Bioinformatics* **2019**, *35*, 1061–1063. [[CrossRef](#)] [[PubMed](#)]
22. Cembrowski, M.S.; Bachman, J.L.; Wang, L.; Sugino, K.; Shields, B.C.; Spruston, N. Spatial Gene-Expression Gradients Underlie Prominent Heterogeneity of CA1 Pyramidal Neurons. *Neuron* **2016**, *89*, 351–368. [[CrossRef](#)] [[PubMed](#)]
23. Lokki, M.L.; Paakkanen, R. The Complexity and Diversity of Major Histocompatibility Complex Challenge Disease Association Studies. *HLA* **2019**, *93*, 3–15. [[PubMed](#)]
24. GTEx Consortium. Human Genomics. The Genotype-Tissue Expression (GTEx) Pilot Analysis: Multitissue Gene Regulation in Humans. *Science* **2015**, *348*, 648–660. [[CrossRef](#)]

25. Degner, J.F.; Pai, A.A.; Pique-Regi, R.; Veyrieras, J.B.; Gaffney, D.J.; Pickrell, J.K.; De Leon, S.; Michelini, K.; Lewellen, N.; Crawford, G.E.; et al. DNase I Sensitivity QTLs Are a Major Determinant of Human Expression Variation. *Nature* **2012**, *482*, 390–394. [[CrossRef](#)]
26. Tehrani, A.; Hie, B.; Dacre, M.; Kaplow, I.; Pettie, K.; Combs, P.; Fraser, H.B. Fine-Mapping cis-Regulatory Variants in Diverse Human Populations. *eLife* **2019**, *8*, e39595. [[CrossRef](#)]
27. Kilpinen, H.; Waszak, S.M.; Gschwind, A.R.; Raghav, S.K.; Witwicki, R.M.; Orioli, A.; Migliavacca, E.; Wiederkehr, M.; Gutierrez-Arcelus, M.; Panousis, N.I.; et al. Coordinated Effects of Sequence Variation on DNA Binding, Chromatin Structure and Transcription. *Science* **2013**, *342*, 744–747. [[CrossRef](#)]
28. Bonder, M.J.; Luijk, R.; Zhernakova, D.V.; Moed, M.; Deelen, P.; Vermaat, M.; van Iterson, M.; van Dijk, F.; van Galen, M.; Bot, J.; et al. Disease Variants Alter Transcription Factor Levels and Methylation of Their Binding Sites. *Nat. Genet.* **2017**, *49*, 131–138. [[CrossRef](#)]
29. Grubert, F.; Zaugg, J.B.; Kasowski, M.; Ursu, O.; Spacek, D.V.; Martin, A.R.; Greenside, P.; Srivas, R.; Phanstiel, D.H.; Pekowska, A.; et al. Genetic Control of Chromatin States in Humans Involves Local and Distal Chromosomal Interactions. *Cell* **2015**, *162*, 1051–1065. [[CrossRef](#)]
30. Waszak, S.M.; Delaneau, O.; Gschwind, A.R.; Kilpinen, H.; Raghav, S.K.; Witwicki, R.M.; Orioli, A.; Wiederkehr, M.; Panousis, N.I.; Yurovsky, A.; et al. Population Variation and Genetic Control of Modular Chromatin Architecture in Humans. *Cell* **2015**, *162*, 1039–1050. [[CrossRef](#)]
31. Tehrani, A.K.; Myrthil, M.; Martin, T.; Hie, B.L.; Golan, D.; Fraser, H.B. Pooled ChIP-seq Links Variation in Transcription Factor Binding to Complex Disease Risk. *Cell* **2016**, *165*, 730–741. [[CrossRef](#)] [[PubMed](#)]
32. Chandra, V.; Bhattacharyya, S.; Schmiedel, B.J.; Madrigal, A.; Gonzalez-Colin, C.; Fotsing, S.; Crinklaw, A.; Seumois, G.; Mohammadi, P.; Kronenberg, M.; et al. Promoter-Interacting Expression Quantitative Trait Loci Are Enriched for Functional Genetic Variants. *Nat. Genet.* **2021**, *53*, 110–119. [[CrossRef](#)] [[PubMed](#)]
33. Tang, Z.; Luo, O.J.; Li, X.; Zheng, M.; Zhu, J.J.; Szalaj, P.; Trzaskoma, P.; Magalska, A.; Wlodarczyk, J.; Rusczycki, B.; et al. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell* **2015**, *163*, 1611–1627. [[CrossRef](#)] [[PubMed](#)]
34. Zhernakova, D.V.; Deelen, P.; Vermaat, M.; van Iterson, M.; van Galen, M.; Arindrarto, W.; van't Hof, P.; Mei, H.; van Dijk, F.; Westra, H.J.; et al. Identification of Context-Dependent Expression Quantitative Trait Loci in Whole Blood. *Nat. Genet.* **2017**, *49*, 139–145. [[CrossRef](#)]
35. Lappalainen, T.; Sammeth, M.; Friedländer, M.R.; 't Hoen, P.A.; Monlong, J.; Rivas, M.A.; González-Porta, M.; Kurbatova, N.; Griebel, T.; Ferreira, P.G.; et al. Transcriptome and Genome Sequencing Uncovers Functional Variation in Humans. *Nature* **2013**, *501*, 506–511. [[CrossRef](#)]
36. Huan, T.; Rong, J.; Liu, C.; Zhang, X.; Tanriverdi, K.; Joehanes, R.; Chen, B.H.; Murabito, J.M.; Yao, C.; Courchesne, P.; et al. Genome-Wide Identification of microRNA Expression Quantitative Trait Loci. *Nat. Commun.* **2015**, *6*, 6601. [[CrossRef](#)]
37. Tan, J.Y.; Smith, A.; Ferreira da Silva, M.; Matthey-Doret, C.; Rueedi, R.; Sönmez, R.; Ding, D.; Kutalik, Z.; Bergmann, S.; Marques, A.C. cis-Acting Complex-Trait-Associated lincRNA Expression Correlates with Modulation of Chromosomal Architecture. *Cell Rep.* **2017**, *18*, 2280–2288. [[CrossRef](#)]
38. Han, Z.; Xue, W.; Tao, L.; Lou, Y.; Qiu, Y.; Zhu, F. Genome-Wide Identification and Analysis of the eQTL lincRNAs in Multiple Sclerosis Based on RNA-seq Data. *Brief. Bioinform.* **2020**, *21*, 1023–1037. [[CrossRef](#)]
39. Ahmed, I.; Karedath, T.; Al-Dasim, F.M.; Malek, J.A. Identification of Human Genetic Variants Controlling Circular RNA Expression. *RNA* **2019**, *25*, 1765–1778. [[CrossRef](#)]
40. Liu, Z.; Ran, Y.; Tao, C.; Li, S.; Chen, J.; Yang, E. Detection of Circular RNA Expression and Related Quantitative Trait Loci in the Human Dorsolateral Prefrontal Cortex. *Genome Biol.* **2019**, *20*, 1–16. [[CrossRef](#)]
41. Knowles, D.A.; Burrows, C.K.; Blischak, J.D.; Patterson, K.M.; Serie, D.J.; Norton, N.; Ober, C.; Pritchard, J.K.; Gilad, Y. Determining the Genetic Basis of Anthracycline-Cardiotoxicity by Molecular Response QTL Mapping in Induced Cardiomyocytes. *eLife* **2018**, *7*, e33480. [[CrossRef](#)] [[PubMed](#)]
42. Fairfax, B.P.; Humburg, P.; Makino, S.; Naranbhai, V.; Wong, D.; Lau, E.; Jostins, L.; Plant, K.; Andrews, R.; McGee, C.; et al. Innate Immune Activity Conditions the Effect of Regulatory Variants upon Monocyte Gene Expression. *Science* **2014**, *343*, 1246949. [[CrossRef](#)] [[PubMed](#)]
43. Garrido-Martín, D.; Borsari, B.; Calvo, M.; Reverter, F.; Guigó, R. Identification and Analysis of Splicing Quantitative Trait Loci Across Multiple Tissues in the Human Genome. *Nat. Commun.* **2021**, *12*, 1–16. [[CrossRef](#)] [[PubMed](#)]
44. Mittleman, B.E.; Pott, S.; Warland, S.; Zeng, T.; Mu, Z.; Kaur, M.; Gilad, Y.; Li, Y. Alternative Polyadenylation Mediates Genetic Regulation of Gene Expression. *eLife* **2020**, *9*, e57492. [[CrossRef](#)]
45. Park, E.; Guo, J.; Shen, S.; Demirdjian, L.; Wu, Y.N.; Lin, L.; Xing, Y. Population and Allelic Variation of A-to-I RNA Editing in Human Transcriptomes. *Genome Biol.* **2017**, *18*, 1–15. [[CrossRef](#)]
46. Zhang, Z.; Luo, K.; Zou, Z.; Qiu, M.; Tian, J.; Sieh, L.; Shi, H.; Zou, Y.; Wang, G.; Morrison, J.; et al. Genetic Analyses Support the Contribution of mRNA N6-Methyladenosine (m6A) Modification to Human Disease Heritability. *Nat. Genet.* **2020**, *52*, 939–949. [[CrossRef](#)]
47. Li, Y.I.; van de Geijn, B.; Raj, A.; Knowles, D.A.; Petti, A.A.; Golan, D.; Gilad, Y.; Pritchard, J.K. RNA Splicing Is a Primary Link Between Genetic Variation and Disease. *Science* **2016**, *352*, 600–604. [[CrossRef](#)]

48. Pai, A.A.; Cain, C.E.; Mizrahi-Man, O.; De Leon, S.; Lewellen, N.; Veyrieras, J.B.; Degner, J.F.; Gaffney, D.J.; Pickrell, J.K.; Stephens, M.; et al. The Contribution of RNA Decay Quantitative Trait Loci to Inter-Individual Variation in Steady-State Gene Expression Levels. *PLoS Genet.* **2012**, *8*, e1003000. [[CrossRef](#)]
49. Kristjánsson, K.; Dziubek, A.; Kang, H.M.; Kwak, H. Population-Scale Study of eRNA Transcription Reveals Bipartite Functional Enhancer Architecture. *Nat. Commun.* **2020**, *11*, 1–12. [[CrossRef](#)]
50. Battle, A.; Khan, Z.; Wang, S.H.; Mitrano, A.; Ford, M.J.; Pritchard, J.K.; Gilad, Y. Genomic Variation. Impact of Regulatory Variation From RNA to Protein. *Science* **2015**, *347*, 664–667. [[CrossRef](#)]
51. Demirkan, A.; Henneman, P.; Verhoeven, A.; Dharuri, H.; Amin, N.; van Klinken, J.B.; Karssen, L.C.; de Vries, B.; Meissner, A.; Göraler, S.; et al. Insight in Genome-Wide Association of Metabolite Quantitative Traits by Exome Sequence Analyses. *PLoS Genet.* **2015**, *11*, e1004835. [[CrossRef](#)] [[PubMed](#)]
52. Suhre, K.; Wallaschofski, H.; Raffler, J.; Friedrich, N.; Haring, R.; Michael, K.; Wasner, C.; Krebs, A.; Kronenberg, F.; Chang, D.; et al. A Genome-Wide Association Study of Metabolic Traits in Human Urine. *Nat. Genet.* **2011**, *43*, 565–569. [[CrossRef](#)] [[PubMed](#)]
53. Bonder, M.J.; Kurilshikov, A.; Tigchelaar, E.F.; Mujagic, Z.; Imhann, F.; Vila, A.V.; Deelen, P.; Vatanen, T.; Schirmer, M.; Smeekens, S.P.; et al. The Effect of Host Genetics on the Gut Microbiome. *Nat. Genet.* **2016**, *48*, 1407–1412. [[CrossRef](#)]
54. Turpin, W.; Espin-Garcia, O.; Xu, W.; Silverberg, M.S.; Kevans, D.; Smith, M.I.; Guttman, D.S.; Griffiths, A.; Panaccione, R.; Otley, A.; et al. Association of Host Genome with Intestinal Microbial Composition in a Large Healthy Cohort. *Nat. Genet.* **2016**, *48*, 1413–1417. [[CrossRef](#)] [[PubMed](#)]
55. Lee, C.; Kim, Y. Optimal Designs for Estimating and Testing Interaction Among Multiple Loci in Complex Traits by a Gibbs Sampler. *Genomics* **2008**, *92*, 446–451. [[CrossRef](#)] [[PubMed](#)]
56. Widmer, C.; Lippert, C.; Weissbrod, O.; Fusi, N.; Kadie, C.; Davidson, R.; Listgarten, J.; Heckerman, D. Further Improvements to Linear Mixed Models for Genome-Wide Association Studies. *Sci. Rep.* **2014**, *4*, 6874. [[CrossRef](#)] [[PubMed](#)]
57. Segura, V.; Vilhjálmsson, B.J.; Platt, A.; Korte, A.; Seren, Ü.; Long, Q.; Nordborg, M. An Efficient Multi-Locus Mixed-Model Approach for Genome-Wide Association Studies in Structured Populations. *Nat. Genet.* **2012**, *44*, 825–830. [[CrossRef](#)]
58. Wainberg, M.; Sinnott-Armstrong, N.; Mancuso, N.; Barbeira, A.N.; Knowles, D.A.; Golan, D.; Ermel, R.; Ruusalepp, A.; Quertermous, T.; Hao, K.; et al. Opportunities and Challenges for Transcriptome-Wide Association Studies. *Nat. Genet.* **2019**, *51*, 592–599. [[CrossRef](#)]
59. Quaas, R.L.; Pollak, E.J. Mixed Model Methodology for Farm and Ranch Beef Cattle Testing Programs. *J. Anim. Sci.* **1980**, *51*, 1277–1287. [[CrossRef](#)]
60. Neal, R.M. MCMC Using Hamiltonian Dynamics. In *Handbook of Markov Chain Monte Carlo*; Brooks, S., Gelman, A., Jones, G.L., Meng, X.-L., Eds.; CRC Press: New York, NY, USA, 2011.
61. Richardson, S.; Tseng, G.C.; Sun, W. Statistical Methods in Integrative Genomics. *Annu. Rev. Stat. Appl.* **2016**, *3*, 181–209. [[CrossRef](#)]
62. Wu, M.; Yi, H.; Ma, S. Vertical Integration Methods for Gene Expression Data Analysis. *Brief. Bioinform.* **2021**, *22*, bbaa169. [[CrossRef](#)] [[PubMed](#)]
63. Li, G.; Shabalin, A.A.; Rusyn, I.; Wright, F.A.; Nobel, A.B. An Empirical Bayes Approach for Multiple Tissue eQTL Analysis. *Biostatistics* **2018**, *19*, 391–406. [[CrossRef](#)] [[PubMed](#)]
64. Gusev, A.; Ko, A.; Shi, H.; Bhatia, G.; Chung, W.; Penninx, B.W.; Jansen, R.; de Geus, E.J.; Boomsma, D.I.; Wright, F.A.; et al. Integrative Approaches for Large-scale Transcriptome-wide Association Studies. *Nat. Genet.* **2016**, *48*, 245–252. [[CrossRef](#)] [[PubMed](#)]
65. Li, L.; Wang, X.; Xiao, G.; Gazdar, A. Integrative Gene Set Enrichment Analysis Utilizing Isoform-specific Expression. *Genet. Epidemiol.* **2017**, *41*, 498–510. [[CrossRef](#)] [[PubMed](#)]
66. Wu, C.; Pan, W. Integrating eQTL Data with GWAS Summary Statistics in Pathway-based Analysis with Application to Schizophrenia. *Genet. Epidemiol.* **2018**, *42*, 303–316. [[CrossRef](#)]
67. Golov, A.K.; Ulianov, S.V.; Luzhin, A.V.; Kalabusheva, E.P.; Kantidze, O.L.; Flyamer, I.M.; Razin, S.V.; Gavrilov, A.A. C-TALE, a New Cost-effective Method for Targeted Enrichment of Hi-C/3C-seq Libraries. *Methods* **2020**, *170*, 48–60. [[CrossRef](#)]
68. Vösa, U.; Claringbould, A.; Westra, H.J.; Bonder, M.J.; Deelen, P.; Zeng, B.; Kirsten, H.; Saha, A.; Kreuzhuber, R.; Yazar, S.; et al. Large-Scale cis- and Trans-eQTL Analyses Identify Thousands of Genetic Loci and Polygenic Scores That Regulate Blood Gene Expression. *Nat. Genet.* **2021**, *53*, 1300–1310. [[CrossRef](#)]
69. Westra, H.J.; Peters, M.J.; Esko, T.; Yaghootkar, H.; Schurmann, C.; Kettunen, J.; Christiansen, M.W.; Fairfax, B.P.; Schramm, K.; Powell, J.E.; et al. Systematic Identification of Trans eQTLs as Putative Drivers of Known Disease Associations. *Nat. Genet.* **2013**, *45*, 1238–1243. [[CrossRef](#)]