J Biosci Vol. 43, No. 5, December 2018, pp. 1069–1075
DOI: 10.1007/s12038-018-9817-7

Review

Regulatory SNPs and their widespread effects on the transcriptome

Vasily M Merkulov[1], Elena Yu Leberfarb[1] and Tatiana I Merkulova[1,2]*

[1]*Laboratory of Gene Expression Regulation, Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Sciences, Novosibirsk, Russia*

[2]*Novosibirsk State University, Novosibirsk, Russia*

**Corresponding author (Email, merktativ@gmail.com)*

MS received 4 June 2018; accepted 25 September 2018; published online 3 October 2018

Currently, it is generally accepted that the *cis*-acting effects of noncoding variants on gene expression are a major factor for phenotypic variation in complex traits and disease susceptibility. Meanwhile, the protein products of many target genes for the identified *cis*-regulatory variants (rSNPs) are regulatory molecules themselves (transcription factors, effectors, components of signal transduction pathways, etc.), which implies dramatic downstream effects of these variations on complex gene networks. Here, we brief the results of recent most comprehensive studies on the role of rSNPs in transcriptional regulation across the genome.

Keywords. Downstream effects; gene expression; regulatory SNPs (rSNPs); transcriptome

A central goal of human genetics is to understand how genetic variation leads to phenotypic differences and complex diseases. Recent genome-wide association studies (GWAS) have identified more than 24 000 SNP–trait associations (MacArthur *et al*. 2017). The vast majority of GWAS variants map to noncoding part of genome and are enriched in regulatory regions (promoters, enhancers, etc.) suggesting many causal variants may affect gene expression (Maurano *et al*. 2012; Bryzgalov *et al*. 2013; Pickrell 2014; Farh *et al*. 2015). Indeed, quite a few SNPs that influence gene regulation (rSNPs) and are associated with various diseases (Edwards *et al*. 2013; Huang *et al*. 2015; Diederichs *et al.* 2016) have been found so far. It is highly important to keep in mind that the products of many target genes for identified *cis*-regulatory variants (transcription factors, effectors, components of signal transduction pathways, etc.) are regulatory molecules themselves, which implies further dramatic effects of these variations on complex gene networks. It seems evident that our understanding of the role of genetic variation in a trait or a disease remains limited without clarifying the effects of the SNPs located in transcriptional regulatory regions on the associated downstream genes and pathways. However, the studies into the functional consequences of rSNPs with regards to widespread effects on the transcriptome are yet very few. The experimental design of such studies (see Table 1) usually consists in (1) overexpression or knockdown of a gene under the control of an rSNP (the key regulator in the context of this paper); (2) CRISPR/Cas9 mediated deletion of a regulatory element or a single nucleotide substitution; and (3) comparison of the transcriptomes of the cells obtained from carriers of different genotypes.

In particular, overexpression of the key regulator was used when examining the downstream effects of rs7132434, located in 2p12.1 multiple renal cell carcinoma (RCC) susceptibility locus. This locus is an enhancer for *BHLHE41* (Li *et al*. 2014; Bigot *et al.* 2016), which encodes a transcription factor involved in the control of circadian rhythm, apoptosis, and cell differentiation (Kato *et al.* 2014). Competitive binding and super shift EMSAs have shown that the G → A substitution (rs7132434) leads to formation of an AP-1 binding site, agreeing well with the luciferase assay data that demonstrate a significantly higher transcriptional activity of the risk allele A as compared with the alternative allele (Bigot *et al.* 2016). The *BHLHE41* overexpression has changed the expression levels of 142 genes, 107 of which were downregulated (Bigot *et al.* 2016) consistent with that BHLHE41 was shown to usually act as a transcriptional repressor (Kato *et al.* 2014). In a xenograft model, the tumors overexpressing BHLHE41 demonstrated a higher growth rate and a larger mass. Since IL-11 is one of the genes most significantly induced by *BHLHE41*

Table 1. The results of recent studies on transcriptomic effects of SNPs, located in genome regulatory regions

Report	SNP	Location	Experimental design	RNA-seq/microarray data	Key regulator
Bigot *et al.* 2016	rs7132434	12p12.1 (enhancer for *BHLHE41)*	BHLHE41 overexpression	35 upregulated genes and 107 downregulated genes	Transcription factor BHLHE41
Schartner *et al.* 2017	rs4500567	−7,8 kb from *TSPAN* TSS	knockdown of *TSPAN8*	231 differentially expressed genes*	TSPAN8
Hill *et al.* 2012	rs1344706	Intron 2 of *ZNF804A*	knockdown of *ZNF804A*	65 upregulated genes and 89 downregulated genes*	Transcription factor ZNF804A
Gupta *et al.* 2017	rs9349379	6p24 regulatory element	1.CRISPR/cas generated isogenic stem cell lines homozygous major (A/A) or minor (G/G) allele 2. CRISPR/cas generated 88 bp deletion	1. 279 differentially expressed genes 2. 423 differentially expressed genes	ET-1, the protein product of the *EDN1* gene
Luo *et al.* 2017	rs10486567 rs67152137 rs7808935	7p15.2 4kb (regulatory region repressive for' *HOXA13)*	1.CRISPR mediated 4kb deletion 2.HOXA13 overexpression	1. 281 upregulated genes and 557 downregulated genes 2. 173 upregulated genes and 933 downregulated genes	Transcription factor HOXA13
Tak *et al.* 2016	rs6983267	E7 enhancer	CRISPR mediated E7 (2.6 kb) deletion	590 upregulated genes and 565 downregulated genes	Transcription factor c-MYC
Ho *et al.* 2018	rs11849538 rs7160302 rs7359033	3′ flank of *TCL1A*	Lymphoblastoid cell lines from homozygous wild-type (C,T,T) and homozygous variant (G, C, C) genotypes	357 genes differentially induced by estradiol	Novel transcription factor TCL1A
Szpakowicz *et al.* 2015	rs12526453	Intron 4 of the *PHACTR1*	peripheral blood mononuclear cells from CC homozygotes and G-allele carriers	31 differentially expressed genes* (patients on admission) 44 differentially expressed genes* (patients on discharge)	unknown
Piras *et al.* 2017	rs17070145	intron 9 of the *WWC1*	Post-mortal hippocampus samples from individuals with T/T and C/C genotypes	8 differentially expressed genes in hippocampus 68 differentially expressed genes in dentate gyrus 20 differentially expressed genes in pyramidal cells	unknown

*Microarray data

overexpression, Bigot *et al.* (2016) assume that the activation of STAT3 transcription factor by IL-11 plays an important role in these processes (as is known, STAT3 enhances proliferation, promotes tumorigenesis, and inhibits apoptosis in RCC (Horiguchi *et al.* 2002; Kaminska *et al.* 2015).

Knockdown of key regulator gene was used to study the transcriptome effects of rs4500567, associated with a bipolar disorder (Schartner *et al.* 2017). The rs4500567 is located in the 5′ flanking region of tetraspanin 8 gene (*TSPAN8*), the product of which is a cell surface protein and a 'molecular facilitator' of cell adhesion, cell–cell interactions, and cell motility (Hemler 2005). The minor C allele of rs4500567 has been found much more frequently in healthy individuals, suggesting its role in protection from the disease (Scholz *et al.* 2010). As has been shown, the expression of luciferase

reporter is significantly decreased in the presence of C allele. RNA interference–mediated knockdown of *TSPAN8* was used to mirror the effects of putatively reduced *TSPAN8* expression in minor allele carriers. Transcriptome analysis of knockdown SH-SY5Y neuroblastoma cells has detected 231 differentially expressed genes. The mechanism underlying the effect of *TSPAN8* repression on the consequent transcriptomic changes is still unknown. However, the available data (Wei *et al*. 2015) suggest involvement of the ERK MAPK signaling pathway. Functional categorization of differentially expressed genes has revealed enrichments of pathways relevant for psychiatric disorders, neuronal interaction, or development and brain function (Schartner *et al*. 2017). The expression of neurotrophic receptor tyrosine kinase 2 (NTRK2), which plays an important role in the CNS development, synaptic plasticity, and, learning (Minichiello 2009), was most significantly changed (Schartner *et al.* 2017).

Similarly, the effects of RNAi-mediated knockdown of key regulator gene *ZNF804A* on human neural progenitor cell transcriptome were assessed (Hill *et al.* 2012). T-allele of rs1344706, located within the intron 2 of the *ZNF804A*, is associated with the broad phenotype of psychosis, including schizophrenia and bipolar disorder (Steinberg *et al*. 2011; Harrison 2017). T- and G- alleles of rs1344706 differentially bind nuclear proteins from several neural cell lines (Hill and Bray, 2011) and more important, it was found that the risk T allele significantly reduces *ZNF804A* expression in the fetal human brain (Hill and Bray 2012). ZNF804A is a multifunctional regulatory protein (Zhou *et al*. 2018) that can act as a transcription factor (Girgenti *et al*. 2012). As a result of transcriptomic study 65 upregulated genes and 89 downregulated genes in knockdown cells were identified. GO analysis of them revealed a significant effect of ZNF804A knockdown on the expression of genes participating in cell adhesion, which suggests a role for ZNF804A in processes such as neural migration, neurite outgrowth and synapse formation (Hill *et al.* 2012).

CRISPR/Cas genome editing is a powerful tool to study the effects of genetic risk variants on gene expression (Beaudoin *et al*. 2015; Soldner *et al*. 2016; Jin *et al*. 2016; Liu *et al*. 2017), including widespread transcriptomic changes. In particular, Gupta *et al.* (2017) used this technology to detect the genome-wide downstream effects of rs9349379, residing in putative enhancer and associated with a number of vascular diseases. First, an 88-bp region containing rs9349379 was deleted in an induced pluripotent stem cell using CRISPR/Cas9 technology. Then clones with a biallelic 88-bp deletion (Δ88) were differentiated into endothelial cells (ECs) to assay six WT (wild-type) and six Δ88 clones by RNA-seq. As a result, 423 differentially regulated genes were identified that displayed enrichments of the pathways relevant for the blood vessel development and morphogenesis. Homozygous major (A/A) or minor (G/G) allele isogenic stem cell lines were generated and differentiated into ECs also. RNA-seq assay of the corresponding clones revealed 279 differentially expressed genes, thereby demonstrating a considerable overlap in the gene set enrichment analysis with those from RNA-seq of the deletion series ECs. Note that both the loss of the 88-bp fragment and replacement of major allele A by minor G resulted in higher expression of the endothelin-1 (*EDN1*) gene. This gene is located over 600 kb distal to rs9349379 and they are likely to interact via a common contact site that lies ∼300 kb from both the EDN1 promoter and rs9349379 regulatory element (Gupta *et al*. 2017). The gene encodes endothelin-1 (ET-1), a 21-residue peptide, presumably influencing the transcriptome via multiple signaling pathways (Sandoval *et al*. 2014).

Luo *et al*. (2017) used CRISPR/Cas9 editing technology to analyze the effect of 7p15.2 noncoding region (one of the 100 GWAS-identified susceptibility loci for prostate cancer) on the transcriptome. This locus contains three SNPs (rs10486567, rs67152137, and rs7808935) associated with an increased risk of disease (Han *et al*. 2015) and, according to Hi-C analysis, has long range interactions with the HOXA locus, located ∼873 kb away (Luo *et al.* 2017). Using the CRISPR/Cas9 system, the authors deleted a 4-kb region encompassing the risk SNPs in RWPE-1 prostate cells. Genome-wide transcriptomic analysis identified 281 upregulated and 557 downregulated genes in the cells lacking this region. Expression of some genes within the HOXA locus (namely *HOXA13* and the adjacent noncoding RNA *HOTTIP*) was also changed. Furthermore, *HOXA13* and *HOTTIP* belong to the most significantly upregulated genes in the whole transcriptome. Then the authors tested if any of the genes with the expression changed as a result of deletion were regulated by HOXA13. According to RNA-seq data, HOXA13 overexpression has led to downregulation of 933 genes and upregulation of 173 genes. Comparison of the obtained datasets detected 21 genes upregulated in the deleted cells and by HOXA13 overexpression and 27 genes downregulated in both cases. Several cancer-related transcription factors were present among the products of these genes. In particular, GATA2 is upregulated in both deletion cells and the cells overexpressing HOXA13. As is known, this factor is associated with prostate cancer and its overexpression increases invasiveness and tumorigenicity (Chiang *et al*. 2014; Xiao *et al*. 2016).

Selfsame, relatively large targeted deletion was generated using CRISPR/Cas9 technology to study trancriptomic effects of E7 enhancer, which resides in an intron of the noncoding RNA CASC8 (Tak *et al*. 2016). E7 harbors rs6983267 associated with an increased risk for colorectal cancer (Yao *et al*. 2014). It was shown that deletion of the region containing rs6983267 leads to large changes in the transcriptome, including upregulation of 590 genes and downregulation of 565 genes, most of which were on other

chromosomes. The authors suggest that one of the main causes of alterations observed is the five-fold reduction of *c-MYC* expression, followed by changes in the MYC regulatory network (Tak *et al*. 2016). The suggestion is strongly supported by the results of previous studies demonstrating the interaction of E7 enhancer region with the *c-MYC* promoter in colon cancer cells (Pomerantz *et al*. 2009; Wright *et al*. 2010)

A comparative analysis of the cell transcriptomes of the subjects with different genotypes is also used to study the genome-wide effects of rSNPs on transcriptional regulation. For example, Ho *et al.* (2018) have used this approach to study a group of three SNPs (rs11849538, rs7160302, and rs7359033) located in the *TCL1A* gene 3′ flanking region and associated with aromatase inhibitor–induced musculoskeletal pain in the women receiving aromatase inhibitors (Ingle *et al*. 2010). It was shown that minor allele G (rs11849538) creates a functional estrogen response element (ERE) and that all three minor variants (G, rs11849538; T, rs7160302; and T, rs7359033) act in concert to increase the *TCL1A* induction by estradiol (Ingle *et al*. 2010; Ho *et al.* 2016). By using ChIP-seq analysis and EMSA TCL1A was identified as a novel transcription factor. RNA-seq assays with two lymphoblastoid cell lines (LCLs) from homozygous wild-type (C, T, and T) and three LCLs with homozygous variant (G, C, and C) genotypes identified 357 genes differentially induced by estradiol (Ho *et al*. 2018). Functional categorization of differentially expressed genes showed their clustering within the pathways involved in transcriptional regulation and T-cell activation, which confirms and supplements the earlier data on the role of TCL1A in the transcriptional regulation of immune mediators (Liu *et al*. 2012; Ho *et al*. 2016).

Another similar example is the study of rs12526453, which is associated with an increased risk of coronary artery disease (van Setten *et al*. 2013; Szpakowicz *et al.* 2015). For transcriptomic analysis, the peripheral blood mononuclear cells from CC homozygotes and G-allele carriers (23 in each group) were used. The clinical samples were collected from patients on admission and discharge (0–24 h); 31 and 44 differentially expressed genes, respectively, were identified. One gene, *NLRP2A,* was upregulated in the group of CC homozygotes (as compared with the carriers of G allele) at both time points. NLRP2 belongs to the nucleotide oligomerization and binding domain (NOD)-like receptor (NLR) family (Lupfer and Kanneganti 2013) and is involved in the inflammatory response (Bruey *et al*. 2004). Presumably, the longest-term survival of rs12526453 CC homozygotes after myocardial infarction is associated with activation of this particular gene (Szpakowicz *et al*. 2015).

In this approach, interesting data were also obtained by Piras *et al*. (2017) who studied genome-wide effects of rs17070145, located in intron 9 of the *WWC1* gene, in human brain postmortal samples. *WWC1* encodes for KIBRA, a scaffold or adaptor-like protein, implicated in a variety of cellular processes. It is known, that rs17070145-T variant is associated with increased episodic memory performance and lowered risk for late onset Alzheimer's disease (Schneider *et al*. 2010; Rodríguez-Rodríguez *et al*. 2009). The authors used laser capture microdissection and RNA-sequencing to investigate the effect of rs17070145 genotypes on transcriptome in the hippocampus (HP), with a specific focus on the dentate gyrus (DG) and at the pyramidal cells (PC) of CA1 and CA3 sub-regions. The number of samples with successful RNA-sequencing was 10 with C/C and 6 with T/T genotypes (HP), 12 with C/C and 10 with T/T genotypes (DG), and 7 with C/C and 3 with T/T genotypes (PC). In the HP, 8 differentially expressed genes were identified, 4 upregulated and 4 downregulated in T/T carriers. In the DG, a total of 66 genes were differentially expressed, specifically 57 genes upregulated and 6 downregulated in T/T carriers. In the PC 1 underexpressed gene and 19 overexpressed genes in T/T carriers were found. The enrichment analysis revealed the T/T genotype to be associated with an overexpression of genes involved in the MAPK signaling pathway, known to be involved in memory and learning (Ciccarelli and Giustetto 2014; Peng *et al*. 2010).

All described approaches have their own advantages and limitations. In particular, overexpression or knockdown of key regulator gene excludes from consideration the effects of other genes that may be under a *cis*-control of the same rSNP, which is a rather frequent situation (Antontseva *et al*. 2015; Fadason *et al.* 2017; Luo *et al*. 2017; Korbolina *et al.* 2018; Baxter *et al.* 2018). A targeting deletion of the regulatory region encompassing rSNP by means of CRISPR/Cas9 editing is aimed at determine the functional role of the region as a whole rather than analyzing the effect of a single SNP, although it is very useful for hypothesizing and further studies. It seems obvious that the introduction of a point mutation using CRISPR/Cas-based technology is the most efficient and objective approach to study genome-wide downstream effects of rSNP; however, the trend of cleaving DNA in a nonspecific manner off the target sites is yet to be overcome (Christie *et al.* 2017; Wang and Ui-Tei 2017). The use of cell lines is concurrently an advantage because it allows the rSNP effects to be examined on the background of the same genotype and a disadvantage since the same cells in culture and within the organism are rather different (Wong and Chen 2012; Soldatow *et al.* 2013; Dvorak 2016). Taking into account this shortcoming, the use of clinical material displays evident advantages; however, genetic heterogeneity of the human populations presents another problem and requires that rather large samples are formed.

Nonetheless, the findings described in the paper significantly contribute to the interpretation of how genotypic variation corresponds to phenotypic outcome, while further development and improvement of the methodical approaches

to the research into the rSNP effects on the associated downstream genes and pathways will evidently give a deeper insight into the basic molecular mechanisms of disease development.

Acknowledgements

The work was supported by the Grants 18-29-09041 from Russian Foundation for Basic Research and State Budget Project 0324-2018-0019.

References

Antontseva EV, Matveeva MY, Bondar NP, Kashina EV, Leberfarb EY, Bryzgalov LO, Gervas PA, Ponomareva AA, Cherdyntseva NV, Orlov YL and Merkulova TI 2015 Regulatory single nucleotide polymorphisms at the beginning of intron 2 of the human KRAS gene. *J. Biosci.* **40** 873–883

Baxter JS, Leavy OC, Dryden NH, Maguire S, Johnson N, Fedele V, Simigdala N, *et al*. 2018 Capture Hi-C identifies putative target genes at 33 breast cancer risk loci. *Nat. Commun.* **9** 1028

Beaudoin M, Gupta RM, Won HH, Lo KS, Do R, Henderson CA, Lavoie-St-Amour C, Langlois S, Rivas D, Lehoux S, Kathiresan S, Tardif JC, Musunuru K and Lettre G 2015 Myocardial Infarction-Associated SNP at 6p24 Interferes With MEF2 Binding and Associates With PHACTR1 Expression Levels in Human Coronary Arteries. *Arterioscler. Thromb. Vasc. Biol*. **35** 1472–1479

Bigot P, Colli LM, Machiela MJ, Jessop L, Myers TA, Carrouget J, Wagner S, Roberson D, Eymerit C, Henrion D and Chanock SJ 2016 Functional characterization of the 12p12.1 renal cancer-susceptibility locus implicates BHLHE41. *Nat. Commun* **7** 12098

Bryzgalov LO, Antontseva EV, Matveeva MY, Shilov AG, Kashina EV, Mordvinov VA and Merkulova TI 2013 Detection of regulatory SNPs in human genome using ChIP-seq ENCODE Data. *PLoS One* **8** e78833

Bruey JM, Bruey-Sedano N, Newman R, Chandler S, Stehlik C and Reed JC 2004 PAN1/NALP2/PYPAF2, an inducible inflammatory mediator that regulates NF-K B and caspase-1 activation in macrophages. *J. Biol. Chem*. **279** 51897–51907

Ciccarelli A and Giustetto M 2014 Role of ERK signaling in activity-dependent modifications of histone proteins. *Neuropharmacology* **80** 34–44

Chiang YT, Wang K, Fazli L, Qi RZ, Gleave ME, Collins CC, Gout PW and Wang Y 2014 GATA2 as a potential metastasis-driving gene in prostate cancer. *Oncotarget* **5** 451–461

Christie KA, Courtney DG, DeDionisio LA, Shern CC, De Majumdar S, Mairs LC, Nesbit MA and Tare Moore CB 2017 Towards personalised allele-specific CRISPR gene editing to treat autosomal dominant disorders. *Sci. Rep*. **23** 16174

Diederichs S, Bartsch L, Berkmann JC, Fröse K, Heitmann J, Hoppe C, Iggena D, Jazmati D, *et al.* 2016 The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations. *EMBO Mol. Med*. **2** 442–457

Dvorak Z 2016 Opportunities and challenges in using human hepatocytes in cytochromes P450 induction assays. *Expert Opin. Drug Metab. Toxicol*. **12** 169–174

Edwards SL, Beesley J, French JD and Dunning AM 2013 Beyond GWASs: illuminating the dark road from association to function. *Am. J. Hum. Genet*. **93** 779–797

Fadason T, Ekblad C, Ingram JR, Schierding WS and O'Sullivan JM 2017 Physical Interactions and Expression Quantitative Traits Loci Identify Regulatory Connections for Obesity and Type 2 Diabetes Associated SNPs. *Front. Genet*. **13** 150

Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shoresh N, Whitton H, *et al.* 2015 Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **19** 337–343

Gupta RM, Hadaya J, Trehan A, Zekavat SM, Roselli C, Klarin D, Emdin CA, Hilvering CRE, *et al.* 2017 Genetic variant associated with five vascular diseases is a distal regulator of endothelin-1 gene expression. *Cell* **27** 522–533

Girgenti MJ, LoTurco JJ and Maher BJ 2012 ZNF804a regulates expression of the schizophrenia-associated genes PRSS16, COMT, PDE4B, and DRD2. *PloS One* **7** e32404

Han Y, Hazelett DJ, Wiklund F, Schumacher FR, Stram DO, Berndt SI, Wang Z, Rand KA, Hoover RN, Machiela MJ, *et al.* 2015 Integration of multiethnic finemapping and genomic annotation to prioritize candidate functional SNPs at prostate cancer susceptibility regions. *Hum. Mol. Genet*. **24** 5603–5618

Harrison PJ 2017 znf804a: insights from the first genome-wide significant schizophrenia gene. *Biol. Psychiatry* **82** 6–7

Hemler ME 2005 Tetraspanin functions and associated microdomains. *Nat. Rev. Mol. Cell Biol.* **6** 801–811

Hill MJ and Bray NJ 2011 Allelic differences in nuclear protein binding at a genome-wide significant risk variant for schizophrenia in ZNF804A. *Mol. Psychiatry* **16** 787–789

Hill MJ and Bray NJ 2012 Evidence that schizophrenia risk variation in the ZNF804A gene exerts its effects during fetal brain development. *Am. J. Psychiatry* **169** 1301–1308

Hill MJ, Jeffries AR, Dobson RJ, Price J and Bray NJ 2012 Knockdown of the psychosis susceptibility gene ZNF804A alters expression of genes involved in cell adhesion. *Hum. Mol. Genet*. **21** 1018–1024

Ho MF, Bongartz T, Liu M, Kalari KR, Goss PE, Shepherd LE, Goetz MP, Kubo M, Ingle JN, Wang L and Weinshilboum RM 2016 Estrogen, SNP-dependent Chemokine expression and selective estrogen receptor modulator regulation. *Mol. Endocrinol.* **30** 382–398

Ho MF, Lummertz da Rocha E, Zhang C, Ingle J, Goss PE, Shepherd LE, Kubo M, Wang L, Li H and Weinshilboum RM 2018 TCL1A, a novel transcription factor and a co-regulator of NF-κB p65: SNP and estrogen-dependence. *J. Pharmacol. Exp. Ther.* **28** 700–710

Horiguchi A, Oya M, Shimada T, Uchida A and Marumo K, Murai M 2002 Activation of signal transducer and activator of transcription 3 in renal cell carcinoma: a study of incidence and its association with pathological features and clinical outcome. *J. Urol.* **168** 762–765

Huang Q 2015 Genetic study of complex diseases in the post-GWAS era. *J. Genet. Genomics* **20** 87–98

Jin HJ, Jung S, DebRoy AR and Davuluri RV 2016 Identification and validation of regulatory SNPs that modulate transcription factor chromatin binding and gene expression in prostate cancer. *Oncotarget*. **23** 54616–54626

Ingle JN, Schaid DJ, Goss PE, Liu M, Mushiroda T, Chapman J-AW, Kubo M, Jenkins GD, *et al.* 2010 Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. *J. Clin. Oncol.* **28** 4674–4682

Kaminska K, Czarnecka AM, Escudier B, Lian F and Szczylik C 2015 Interleukin-6 as an emerging regulator of renal cell cancer. *Urol. Oncol.* **33** 476–485

Kato Y, Kawamoto T, Fujimoto K and Noshiro M 2014 DEC1/STRA13/SHARP2 and DEC2/SHARP1 coordinate physiological processes, including circadian rhythms in response to environmental stimuli. *Curr. Top. Dev. Biol*. **110** 339–372

Korbolina EE, Brusentsov II, Bryzgalov LO, Leberfarb EY, Degtyareva AO and Merkulova TI 2018 Novel approach to functional SNPs discovery from genome-wide data reveals promising variants for colon cancer risk. *Hum. Mutat*. **23** 851–859

Li Q, Stram A, Chen C, Kar S, Gayther S, Pharoah P, Haiman C, Stranger B, Kraft P and Freedman ML 2014 Expression QTL-based analyses reveal candidate causal genes and loci across five tumor types. *Hum. Mol. Genet.* **23** 5294–5302

Liu M, Wang L, Bongartz T, Hawse J, Markovic S, Schaid D, Mushiroda T, Kubo M, Nakamura Y, Kamatani N, Goss P, Ingle JN and Weinshilboum RM 2012 Aromatase inhibitors, estrogens and musculoskeletal pain: estrogen-dependent T-cell leukemia 1A (TCL1A) gene-mediated regulation of cytokine expression. *Breast Cancer Res.* **14** R41

Liu S, Liu Y, Zhang Q, Wu J, Liang J, Yu S, Wei GH, White KP and Wang X 2017 Systematic identification of regulatory variants associated with cancer risk. *Genome Biol.* **23** 194

Luo Z, Rhie SK, Lay FD and Farnham PJ 2017 A Prostate Cancer Risk Element Functions as a Repressive Loop that Regulates HOXA13. *Cell Rep.* **21** 1411–1417

Lupfer C and Kanneganti TD 2013 Unsolved Mysteries in NLR Biology. *Front. Immunol*. **4** 285

MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, Junkins H, McMahon A, *et al.* 2017 The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res*. **4** 896–901

Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, *et al.* 2012 Systematic localization of common disease-associated variation in regulatory DNA. *Science* **7** 1190–1195

Minichiello L 2009 TrkB signalling pathways in LTP and learning. *Nat. Rev. Neurosci.* **10** 850–860

Peng S, Zhang Y, Zhang J, Wang H and Ren B 2010 ERK in learning and memory: a review of recent research. *Int. J. Mol. Sci*. **11** 222–232

Piras IS, Krate J, Schrauwen I, Corneveaux JJ, Serrano GE, Sue L, Beach TG and Huentelman MJ 2017 Whole transcriptome profiling of the human hippocampus suggests an involvement of the KIBRA rs17070145 polymorphism in differential activation of the MAPK signaling pathway. *Hippocampus* **27** 784–793

Pickrell JK 2014 Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am. J. Hum. Genet*. **3** 559–573

Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, Beckwith CA, Chan JA, Hills A, Davis M, *et al*. 2009 The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat. Genet*. **41** 882–884

Rodríguez-Rodríguez E, Infante J, Llorca J, Mateo I, Sánchez-Quintana C, García-Gorostiaga I, Sánchez-Juan P, Berciano J and Combarros O 2009 Age-dependent association of KIBRA genetic variation and Alzheimer's disease risk. *Neurobiol. Aging.* **30** 322–324

Sandoval YH, Atef ME, Levesque LO, Li Y and Anand-Srivastava MB 2014 Endothelin-1 signaling in vascular physiology and pathophysiology. *Curr. Vasc. Pharmacol*. **12** 202–214

Schartner C, Scholz CJ, Weber H, Post A, Freudenberg F, Grünewald L and Reif A 2017 The regulation of tetraspanin 8 gene expression-A potential new mechanism in the pathogenesis of bipolar disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **174** 740–750

Scholz CJ, Jacob CP, Buttenschon HN, Kittel-Schneider S, Boreatti-Hümmer A, Zimmer M, Walter U, Lesch KP, Mors O, Kneitz S, Deckert J and Reif A 2010 Functional variants of TSPAN8 are associated with bipolar disorder and schizophrenia. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **153** 967–972

Schneider A, Huentelman MJ, Kremerskothen J, Duning K, Spoelgen R and Nikolich K 2010 KIBRA: A new gateway to learning and memory? *Front. Aging Neurosci*. **2** 4

Soldatow VY, Lecluyse EL, Griffith LG and Rusyn I 2013 *In vitro* models for liver toxicity testing. *Toxicol. Res.* **2** 23–39

Soldner F, Stelzer Y, Shivalila CS, Abraham BJ, Latourelle JC, Barrasa MI, Goldmann J, Myers RH, Young RA and Jaenisch R 2016 Parkinson-associated risk variant in distal enhancer of α-synuclein modulates target gene expression. *Nature* **5** 95–99

Steinberg S, Mors O, Børglum AD, Gustafsson O, Werge T, Mortensen PB, Andreassen OA, Sigurdsson E, *et al*. 2011 Expanding the range of ZNF804A variants conferring risk of psychosis. *Mol. Psychiatry* **16** 59–66

Szpakowicz A, Kiliszek M, Pepinski W, Waszkiewicz E, Franaszczyk M, Skawronska M, Ploski R., Niemcunowicz-Janica A, *et al*. 2015 The rs12526453 Polymorphism in an intron of the PHACTR1 gene and its association with 5-year mortality of patients with myocardial infarction. *PLoS One* **18** e0129820

Tak YG, Hung Y, Yao L, Grimmer MR, Do A, Bhakta MS, O'Geen H, Segal DJ and Farnham PJ 2016 Effects on the transcriptome upon deletion of a distal element cannot be predicted by the size of the H3K27Ac peak in human cells. *Nucleic Acids Res*. **44** 4123–4133

van Setten J, Isgum I, Smolonska J, Ripke S, de Jong PA, Oudkerk M, *et al.* 2013 Genome-wide association study of coronary and aortic calcification implicates risk loci for coronary artery disease and myocardial infarction. *Atherosclerosis* **228** 400–405

Wang Q and Ui-Tei K 2017 computational prediction of CRISPR/Cas9 target sites reveals potential off-target risks in human and mouse. *Methods Mol. Biol*. **1630** 43–53

Wei L, Li Y and Suo Z 2015 TSPAN8 promotes gastric cancer growth and metastasis via ERK MAPK pathway. *Int. J. Clin. Exp. Med.* **8** 8599–8607

Wong C and Chen S 2012 The development, application and limitations of breast cancer cell lines to study tamoxifen and aromatase inhibitor resistance. *J. Steroid Biochem. Mol. Biol.* **131** 83–92

Wright JB, Brown SJ and Cole MD 2010 Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. *Mol. Cell. Biol.* **30** 1411–1420

Xiao L, Feng Q, Zhang Z, Wang F, Lydon JP, Ittmann MM, Xin L, Mitsiades N and He B 2016 The essential role of GATA transcription factors in adult murine prostate. *Oncotarget* **7** 47891–47903

Yao L, Tak YG, Berman BP and Farnham PJ 2014 Functional annotation of colon cancer risk SNPs. *Nat. Commun*. **5** 5114

Zhou Y, Dong F, Lanz TA, Reinhart V, Li M, Liu L, Zou J, Xi HS and Mao Y 2018 Interactome analysis reveals ZNF804A, a schizophrenia risk gene, as a novel component of protein translational machinery critical for embryonic neurodevelopment. *Mol. Psychiatry.* **23** 952–962

Corresponding editor: PARTHA P MAJUMDER