Incomplete penetrance and variable expressivity: is there a microRNA connection?

Jasmine K. Ahluwalia,[1,2] *Manoj Hariharan,*[2] *Rhishikesh Bargaje,*[2] *Beena Pillai,*[2] *and Vani Brahmachari*[1]*

[1]Dr. B. R. Ambedkar Centre for Biomedical Research, Delhi University, Delhi, India
[2]Institute of Genomics and Integrative Biology (CSIR), Mall Road, Delhi, India

Incomplete penetrance and variable expressivity are non-Mendelian phenomena resulting in the lack of correlation between genotype and phenotype. Not withstanding the diversity in mechanisms, differential expression of homologous alleles within cells manifests as variations in penetrance and expressivity of mutations between individuals of the same genotype. These phenomena are seen most often in dominantly inherited diseases, implying that they are sensitive to concentration of the gene product. In this framework and the advances in understanding the role of microRNA (miRNA) in fine-tuning gene expression at translational level, we propose miRNA-mediated regulation as a mechanism for incomplete penetrance and variable expressivity. The presence of miRNA binding sites at 3′ UTR, co-expression of target gene–miRNA pairs for genes showing incomplete penetrance and variable expressivity derived from available data lend support to our hypothesis. Single nucleotide polymorphisms in the miRNA target site facilitate the implied differential targeting of the transcripts from homologous alleles.

Keywords: epigenetics; incomplete penetrance; MicroRNA; modifier; variable expressivity

Introduction

The lack of correlation between phenotype and genotype is observed in several phenomena showing non-Mendelian inheritance such as genomic imprinting, dynamic mutations, incomplete penetrance, variable expressivity, and maternal inheritance.[1,2] Coining of the terms, incomplete penetrance and variable expressivity (IP-VE) is attributed to Oskar Vogt, who recognized the inadequacy of the terminology ''dominance'' to describe inheritance and proposed the term ''penetrance'' and ''expressivity''.[3] He provided a distinction between the two as follows "When there is strong penetrance, which is a percentage increase in the manifestation of a trait (*in a population*), the character can occur in weak form (weak expressivity). Conversely, at low penetrance where there is a percentage decrease in the rare manifestation of this character, the character can occur in its very pronounced form (strong expressivity)''.[3] A condition is said to have ''complete penetrance'' if clinical symptoms are present in all individuals carrying the mutation and ''incomplete penetrance'' if clinical symptoms are not always present in individuals with the same genotype (Genetics Home Reference, NLM). The fact that incomplete penetrance is seen most often in dominantly inherited diseases indicates that these phenotypes are vulnerable to alteration in concentration of the gene product. On similar lines, variable expressivity has been described as ''variation'' in clinical features (type or severity) of a genetic disorder between individuals with the same genotype, even within the same family (Genetics Home Reference, NLM). These phenomena have been reported in several genetic disorders, *i.e.*, congenital heart disease,[4] type II diabetes,[5] cystic fibrosis,[6] Hirschsprung disease,[7] erythropoietic protoporphyria (EPP),[8] retinitis pigmentosa 11,[9] and retinoblastoma.[10,11] A large family segregating retinoblastoma was described by Dryja *et al.*, in which a germline deletion in the *RB* gene encompassing exon 4 was detected both in the affected children and one of their unaffected parents.[10]

The lack of satisfactory mechanistic explanations for the phenomena led Vogel and Motulsky to refer to IP-VE as "labels of our ignorance'' (Human Genetics: Principles and Approaches, 1997, third edition). However, in the recent years there has been considerable advances in the understanding of the basis of IP-VE in specific cases.[1,2] These studies reiterate that IP and VE are closely related phenomena forming a continuum and sharing common mechanisms. On the other hand, the molecular basis of differential expression of homologous alleles has been extensively investigated in the context of genomic imprinting and random inactivation of X chromosome in female mammals.[12–15]

Abbreviations: AS, Angelman syndrome; DAVID, Database for Annotation, Visualization and Integration Discovery; *DICER1*, dicer 1, ribonuclease type III; DMR, differential methylation region; *EHMT1*, euchromatic histone-lysine *N*-methyltransferase 1; EPP, erythropoietic protoporphyria; *EZH*, enhancer of zeste; HP1, heterochromatin protein 1; IP-VE, incomplete penetrance and variable expressivity; miRNA, microRNA; OMIM, Online Mendelian Inheritance in Man; *PRMD1*, PR domain containing 1, with ZNF domain; PWS, Prader Willi syndrome; SNP, single nucleotide polymorphism; TPM, transcripts per million; TRBP, transactivating response RNA-binding protein.

***Correspondence to:** V. Brahmachari, Dr. B. R. Ambedkar Center for Biomedical Research, Delhi University, Delhi 110007, India.
E-mails: vbrahmachari@acbr.du.ac.in; vani.brahmachari@gmail.com

Molecular mechanisms of differential expression of homologous alleles

The molecular correlates of differential expression of homologous alleles known so far for genomic imprinting and X chromosome inactivation illustrate the diversity of strategies and consequences of differential epigenetic marking that ultimately results in differential gene expression. A few examples are discussed here to represent the variety of mechanisms that underlie differential expression of homologous alleles.

The H19/Igf2 imprinted gene pair is one of the loci extensively studied for differential epigenetic modifications. Differential transcription at this locus is correlated with variation in DNA methylation at the differential methylation region (DMR) upstream of the H19 locus, differential DNase I hypersensitivity and replication asynchrony between imprinted alleles.[16–19] The consequence of these molecular correlates is that the enhancer present downstream of H19 gene is not accessible to Igf2 promoter on the maternal chromosome.[16]

Modification of histones is an epigenetic marking reported in several cases, including Prader Willi and Angelman syndrome (PWS and AS) region comprising of *SNRPN*, *IGF2R*, and *U2AF1-RS* genes.[20] Differential modifications of histones and/or DNA can be brought about by interaction of polycomb (PcG) and trithorax (trxG) complexes, which maintain, respectively a silenced or an active state of expression of developmental genes.[21] Members of PcG complex like enhancer of zeste (*EZH*), a SET domain containing protein, bring about histone H3K9 methylation, which is recognized by heterochromatin protein 1 (HP1), leading to inactivation.[22] The unification of mechanisms of epigenetic regulation is reflected by similar histone modifications through polycomb complexes to maintain X chromosome inactivation.[14] Non-coding RNA from *XIST* gene, the master regulator, initiates the process.[14] Variation in expression in different transgenic lines with the same transgene is often encountered. This discordance is correlated with copy number, the status of DNA, and histone methylation at the site-of-integration of the transgene.[23,24] These examples reiterate that in spite of diverse mechanisms, the reinforcement of differential expression appears to be principally brought about either by DNA methylation and/or histone modifications.

Molecular basis of IP-VE

The fact that IP-VE is seen most often in dominantly inherited diseases indicates that these phenotypes are sensitive to concentration of the gene product. The basis of incomplete penetrance and/or variable expressivity elucidated so far indicates that it varies between genes and that there may not be a common mechanism applicable to all cases.

IP-VE of triplet repeat-associated disorders, like the fragile X syndrome, is explained by the detection of premutation status in carriers and in patients; manifestation of the disease and its severity is directly correlated with CGG repeat number.[25] The length of CGG repeat in the 5′UTR of fragile X mental retardation gene (*FMR1*) is also tightly correlated with methylation of the promoter and hence transcription of the gene.[26]

Apart from epigenetic regulation, the number of genetic lesions and nature of functional domain altered by mutation or polymorphism may contribute to phenotypic variability in complex diseases. The loss of canonical splice site by a single base change in the *RET* gene, a receptor tyrosine kinase in Hirschsprung disease (aganglionic megacolon), results in the absence of expression of functional protein and the severity of the disease is related to the homozygosity of the mutation.[7,27] EPP inherited as an autosomal dominant disorder also shows incomplete penetrance since affected and normal individuals (both carrying the mutation) are observed in the same family.[28] The authors suggest that penetrance is modulated by an intronic single nucleotide polymorphism (SNP) at a cryptic splice site on the wild-type allele, leading to production of an aberrant transcript that is degraded and, therefore, results in near absence of the ferrochelatase enzyme. Individuals carrying the dominant mutation as well as the intronic SNP will, therefore, show a total absence of the ferrochelatase enzyme, while those with the mutation but no intronic SNP will have reduced levels. Thus, cryptic regulatory variations and allelic variations can lead to clinically indistinguishable phenotypes, which adequately explain the observed deviation in penetrance and expressivity.[29–31] However, such changes are not reported in several monogenic dominant disorders varying in penetrance and expressivity.[32–36] A comprehensive review of the mechanisms known for non-Mendelian inheritance is provided by Heyningen and Yeyati[1] and Zlotogora.[2]

The factors that affect gene product levels include gene-specific transcription[37,38] and translation regulators.[39] Therefore, a monogenic disease could be under the influence of gene-specific interacting partners that are referred to as modifier genes. Epigenetic regulators like DNA and histone methyltransferases are also included under this category.[40] Consequently, it follows that the wild-type or normal alleles of genes are also influenced by these interacting factors. However, when there are no mutations affecting the functional protein, the cell achieves a steady-state level of normal protein required for the function and, hence, manifests normal phenotype.

The role of *trans*-acting modifiers is implicated in IP-VE, although specific modifiers are identified in only a limited number of cases.[41,42] Promoter methylation is recognized as

a cancer risk modifier in BRCA1/BRCA2 mutation carriers.[42] Genes designated as modifiers can have functions other than catalytic activity leading to DNA and histone modifications. In case of breast cancer, tumor occurrence due to predisposing mutation in BRCA2 is affected by a SNP in Rad51, coding for a DNA repair enzyme.[43] However, in many cases the identity of modifier genes is not established.[44,45]

In addition to transcriptional regulation of gene expression, the impact of translational regulation by non-coding RNA on phenotype is increasingly being recognized as a critical step in disease manifestation.[46,47]

MicroRNAs as regulators of gene expression

MicroRNAs (miRNAs) are an important class of endogenously expressed non-coding RNAs that can bring about regulation at the post-transcriptional level, resulting in apparent lack of correlation between genotype and phenotype.[48] miRNAs are transcribed as long primary transcripts (pri-miRNA) from intergenic or intronic sequences of the genome (Fig. 1). The gene that gives rise to miRNA is called the source gene, while the target gene is one whose 3′UTR is recognized by the miRNA. The pri-miRNA is processed in the nucleus by an RNaseIII complex of Drosha and DiGeorge syndrome critical region 8 (DGCR8), to give ~70–100-nucleotide-long pre-miRNA that is transported to the cytoplasm by a RNA-binding, Ras-related nuclear protein (Ran)-GTP-dependent protein called Exportin5. In the cytoplasm, the pre-miRNA is further cleaved into a small 21–25-nucleotide duplex by the RNaseIII complex of Dicer and transactivating response RNA-binding protein (TRBP). Single strand from the duplex, *i.e.*, the mature miRNA is then incorporated into an miRNA-associated ribonucleoprotein (miRNP) complex containing argonaute and other proteins, which binds to 3′UTR region of target mRNA through partial or complete complementarity.[48] Thus, one miRNA can have multiple targets within the same cell. miRNAs can regulate translation either by degradation of the mRNA or by translational repression. The 2–8 nucleotides at the 5′end of miRNA that is completely complementary to the target region in the 3′UTR of target gene is called the "seed" region (Fig. 1).

There are numerous examples showing the effect of miRNA on gene expression.[49–55] In several instances, 3′UTR of a target gene may carry sites for multiple miRNAs, allowing for combinatorial regulation by these miRNAs. Recently, experimental evidence for the implied cooperativity of miRNAs acting on the same target has been reported.[56] One miRNA can directly bind to hundreds of targets, which entails competition between targets for the same miRNA.[57] Both cooperativity and competition are concentration-dependent phenomena. Furthermore, miRNA-target interaction does not require complete complementarity. Typically, complete complementarity at the seed sequence and additional interrupted interactions along the length, add to the affinity of the target for the miRNA and determine the stability of mRNA–miRNA interaction. *In vivo*, with several targets competing for the same miRNA, the number of sites, affinity, accessibility, and noise in transcription of miRNA and target gene are factors likely to influence miRNA-target interaction.[58,59] These attributes of miRNA-mediated regulation presumably introduce stochasticity to the regulatory network of gene expression. However, it is proposed that transcriptional noise could be reduced by miRNA regulation, leading to canalization of development.[60] In cases where the expression of target gene and miRNA are in mutually exclusive cells

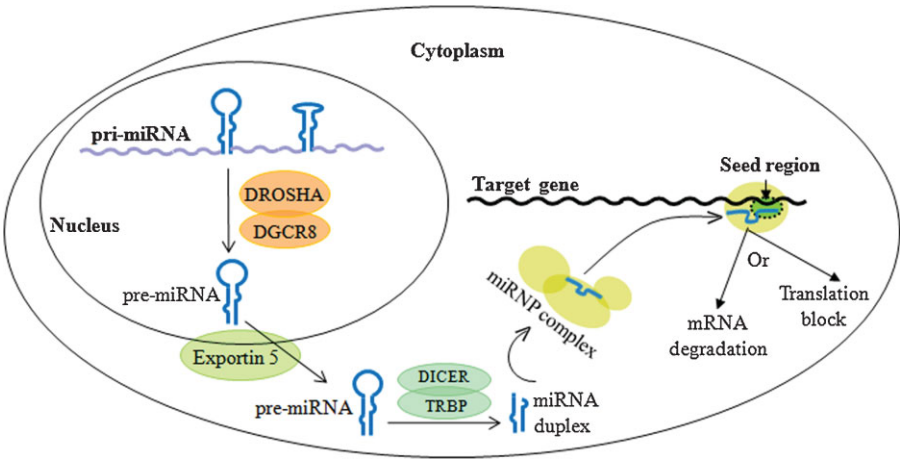

Figure 1. Outline of miRNA biogenesis and function. The functions of the proteins, DICER, DGCR8, Exportin, and TRBP are described in the text.

miRNA can abolish ectopic leaky expression of genes.[60] Accordingly, perturbations in interaction between miRNA and its target could affect the robustness of developmental processes.

A model for the role of miRNA in IP-VE

We propose that miRNA-mediated regulation of protein levels brings about non-equivalence of homologous alleles at the translational level within the same cell, and when one of the alleles is mutated as in monogenic diseases, the phenotypic outcome depends upon which of the two alleles is affected by miRNA. Based on this proposition, we suggest a model for miRNA-mediated regulation of IP-VE (Fig. 2). The miRNAs target the 3′UTR of mRNA and lead to either degradation of mRNA or to inhibition of translation, but given that IP-VE entails variability in the expression level of the two alleles rather than complete silencing of one of them, we consider translational block as the consequence of miRNA–mRNA interaction.

Under normal circumstances the homologous alleles are expected to have a nearly identical sequence and, therefore, any one of them may be targeted for translational regulation by the cognate miRNA. At this step the property of miRNA-mediated regulation to operate even in the absence of complete sequence complementarity is of significance. In the absence of any mutation in the target gene, the cellular level of the protein coded by the target gene is adequate to establish a normal phenotype. However, when there is a mutation in a heterozygous state, the phenotype is dictated by which of the two products of transcription from the homologous alleles is targeted by miRNA, thus resulting in IP-VE. miRNA-mediated regulation is now known to be a cell-autonomous process. Therefore, the outcome at the tissue level will depend upon which of the allelic products is suppressed in the majority of the cells that make up the tissue, thus bringing about variable expressivity between tissues/organ systems. The allele selection for miRNA targeting can be augmented by SNP at the target sequence in one of the alleles, as illustrated in Fig. 3. SNP in the 3′UTR could alter the stability of mRNA–miRNA duplex. In addition, the expression level of the miRNA source gene can also influence tissue level variation of regulatory miRNA and, hence, the severity and penetrance of the disease.

However, allelic selection can be speculated to be under the control of the organ-specific network of regulatory factors, which will include the regulators of target gene transcription,

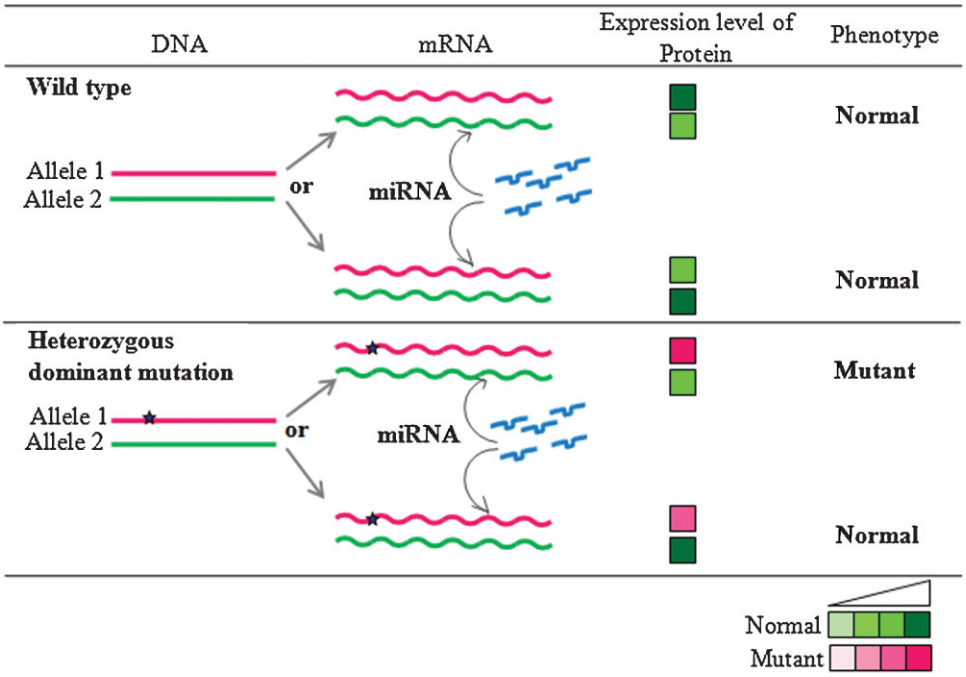

Figure 2. A model for miRNA-mediated regulation of penetrance and/or variable expressivity. The homologous alleles and their mRNA are shown as differently colored symbols. The levels of protein products from the genes are shown with graded colors: green for normal and pink for mutant protein. The mutation on the gene and the mRNA is shown with a blue star. The level of the protein product from the two alleles will vary depending on which of the allelic mRNAs is targeted by miRNA and when there is a mutation in heterozygous state in the gene, the differential targeting will result in mutant or normal phenotype depending on which allele is targeted. Further SNP in the target site will enhance the differential expression of the homologous alleles at the protein level and is described in Figure 3.

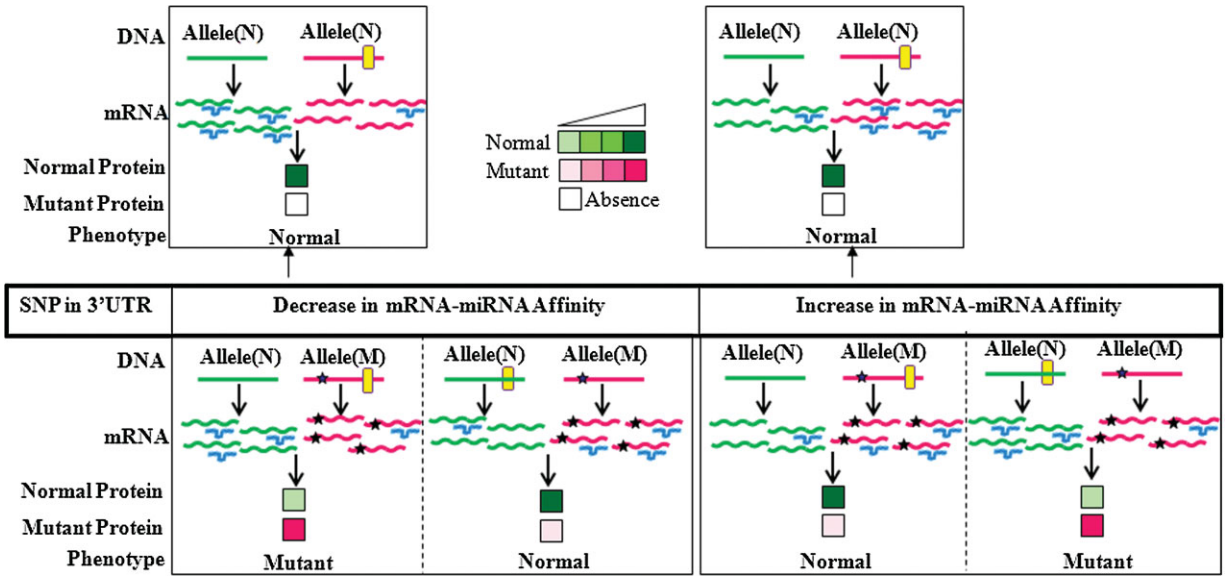

Figure 3. Consequence of the presence of SNP in the miRNA target site. SNP is shown as yellow box. The levels of protein products from the genes are shown with graded colors: green for normal and pink for mutant protein. The mutation on the gene and the mRNA is shown with a blue star. Occurrence of SNP in the absence of disease causing mutation is of no consequence on the phenotype (top panel). Various ways of strengthening differential regulation in mutation background are shown: SNP at target site can lead to either an increase or decrease in affinity between mRNA and miRNA, SNP can occur in the normal or the mutant allele. The consequence of each variable on the phenotype is depicted.

the number and transcriptional level of alternative targets of the same miRNA, transcriptional level of the miRNA source gene and its regulators.

As mentioned earlier, modifier genes can influence gene expression and, hence, penetrance and expressivity. We consider the possibility that miRNAs can also regulate the product level of modifier genes like DNA methyl transferase, which in turn can modulate methylation of upstream sequences of genes. By utilizing the data available in the public domain we have examined the possible targeting of IP-VE and modifier genes by miRNA and evaluated the correlations between the miRNA and its target mRNA predicted by our hypothesis.

Are IP-VE genes targeted by miRNA?

We retrieved genes implicated in autosomal single gene disorders showing incomplete or reduced penetrance and/or variable expressivity from literature available at Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/). The sequences of genes listed by Surani *et al.* as modifier genes were retrieved from ENSEMBL.[40]

A representative list of genes is given in Table 1 and the complete list is given in Supporting Information Table 1. We generated a list of 143 genes implicated in autosomal single gene disorders that show incomplete or reduced penetrance and/or variable expressivity and screened them

Table 1. A representative list of genes retrieved for analysis

Gene	Function	Disease	Phenomenon
BEST1	Ion transport	Macular dystrophy, vitelliform; VMD	IP
PAX3	Transcription factor	Waardenburg syndrome, type I; WS1	IP
FLT4	Signal transduction	Lymphedema, hereditary, I	IP
CRB1	Calcium Ion binding	Pigmented paravenous chorioretinal atrophy; PPCRA	VE
MSX1	Transcription factor	Orofacial cleft 5; Witkop syndrome	VE
NIPBL	Structural role in chromatin	Cornelia De Lange syndrome 1; CDLS1	VE
RB1	Transcription factor	Retinoblastoma; RB1	IP-VE
HOXD10	Transcription factor	Vertical talus, congenital; CVT	IP-VE
TOR1A	ATP binding	Torsion Dystonia 1, autosomal dominant; DYT1	IP
TTR	Signal transduction	Amyloidosis, senile systemic	IP
DNMT1	DNA methylase	–	Modifier
HDAC1	Histone deacetylase	–	Modifier

Table 2. Identification of miRNA targets

			miRNA target prediction		
Gene class	**Total genes**	**Genes with ENST[a]**	**Target genes**	**miRNA**	**Gene–miRNA pairs**
IP-VE	143[b]	128	88	167	233
Modifiers	29[c]	29	21	47	49
Control	516[d]	516	297	277	750

[a]Number of genes having Ensembl Transcript ID (ENST) annotation including 3′UTR sequence at Ensembl.
[b]Number of genes retrieved from OMIM showing IP-VE.
[c]Number of modifier genes taken from Surani *et al.*[40]
[d]Number of house-keeping genes used as control set[67] (p value$<$0.05 IP-VE to control).

for the presence of putative miRNA binding sites in their 3′UTR. The consensus of three miRNA target prediction tools, miRanda,[61] TargetScan,[62] and RNAhybrid,[63] was used for miRNA target prediction to reduce the false-positive predictions known to be associated with each of the tools.[64] The procedure followed was according to Hariharan *et al.* and is detailed under methods.[65] We found that amongst the 128 genes of the IP-VE set for which the 3′UTR sequences were retrieved, 70% of the genes (88 of 128) contained miRNA binding site(s) (Table 2, Supporting Information Table 2).

These 88 IP-VE genes were targeted by 167 human encoded miRNAs giving rise to 233 target gene–miRNA pairs, which indicated that some genes are targeted by multiple miRNA; 66% of the genes (58 of 88) contained multiple non-overlapping miRNA binding sites in their 3′UTR. For example, *TOR1A* (torsin family 1, member A; torsin A) contains six non-overlapping miRNA binding sites within 1 kb, while *TTR* (transthyretin, prealbumin, amyloidosis type I) has three such sites within 150 bases. This suggests a possible combinatorial regulation where varying combinations of miRNA binding to a target transcript could bring about variation in expression.

We used housekeeping genes as a control set to compare our list of IP-VE genes as potential targets of miRNA regulation, as expression of these essential genes is subjected to minimum stochasticity or noise as shown in *Saccharomyces cerevisiae*.[66] We retrieved the sequences of a set of housekeeping genes given by Eisenberg and Levanon.[67] In this comparison, IP-VE genes had a higher probability of being under miRNA-mediated regulation (p value$<$0.05, Fisher's exact test, 88 of 128 IP-VE and 297 of 516 housekeeping genes, Table 2).

Epigenetic modifiers are targeted by miRNA

The modifier gene list contains 31 epigenetic modifiers including histone and DNA methylases, histone acetylases, and deacetylases.[40] Two of these modifiers, *DNMT3B* (OMIM 242860) and *SMARCB1* (OMIM 601607) were also retrieved as IP-VE genes and were targets of miRNA. Of the 29 modifier genes, 21 were also targeted by miRNA (Supporting Information Table 2); 12 of these contained multiple non-overlapping miRNA-binding sites, including dicer 1, ribonuclease type III (*DICER1*) and PR domain containing 1, with ZNF domain (*PRMD1*) that contained five exclusive miRNA-binding sites each.

miRNAs can also arise from intronic sequences of genes.[68] We examined if any of the genes in our list was also a source of miRNA and found that euchromatic histone-lysine *N*-methyltransferase 1 (*EHMT1*), a modifier gene encodes an intronic miRNA, hsa-mir-602, which targets 30 other genes in the human genome, according to the consensus miRNA target prediction. Thus, the network effect of a target gene for one miRNA being the source of a different miRNA by itself can also bring in an additional layer of variability in miRNA-mediated regulation of IP-VE.

Functional classification of IP-VE genes

The functional classification of the 88 IP-VE target genes using Database for Annotation, Visualization and Integration Discovery (DAVID), http://david.niaid.nih.gov),[69] a tool that implements Gene Ontology (GO) term clustering, indicated that the IP-VE target genes are predominantly involved in biological processes of development and neurophysiology. In molecular function, the major category is transcription regulation (Table 3). The modifier target genes are enriched in cellular metabolism (biological process), transferase, and DNA binding activity (molecular function).

Co-expression of miRNA and the target gene

It has been suggested that miRNA can regulate translation either by degradation of mRNA or by translational repression.[55] In the former case, miRNA and its target mRNA

Table 3. GO functional classification of IP-VE genes targeted by miRNA

	Level	*p* value
Biological process		
Skeletal development	3	1.30E-08
Sensory perception of sound	4	1.51E-08
Sensory perception of mechanical stimulus	3	1.73E-08
Neurophysiological process	3	1.11E-04
Regulation of cellular metabolism	4	1.62E-04
Regulation of cellular physiological process	3	2.74E-04
Regulation of metabolism	3	3.49E-04
Organ morphogenesis	3	5.29E-04
Regulation of cell size	4	7.71E-04
Cell growth	3	8.30E-04
Visual perception	4	1.25E-03
Sensory perception of light stimulus	3	1.34E-03
Cellular morphogenesis	4	1.67E-03
Molecular function		
Transcription factor activity	4	9.68E-05
Sequence-specific DNA binding	4	1.49E-04
Alpha-type channel activity	3	5.49E-04
Adenyl nucleotide binding	4	1.37E-03
Purine nucleotide binding	3	2.59E-03
DNA binding	3	8.88E-03

are not expected to co-express.[70] However, preferential co-expression of target mRNA and miRNA has been experimentally demonstrated in some cases.[51–53] Since the description of IP-VE implies modulation rather than complete silencing of bi-allelic expression, we consider that for the miRNA to function as a post-transcriptional regulator, co-occurrence with target mRNA is required. Consequently, subtle changes in levels of the co-occurring target gene–miRNA pair, under different cellular conditions in different individuals, may lead to variable expressivity.

The data, we utilized, were generated from different tissues, and we have considered only those tissues where the data for both mRNA and the miRNA were available. mRNA expression data for tissues and developmental stages are available at Unigene, EST ProfileViewer (www.ncbi.nlm.nih.gov/UniGene), while miRNA data are from human tissues from commercial sources.[71] Based on this, we have considered co-expression of miRNA and mRNA in 27 adult tissues. However, the most relevant stage-specific co-expression would be at fetal stage for both mRNA and miRNA in the light of predominant representation of IP-VE genes in the developmental pathway. Secondly, the data refer to the tissue as a whole and not to any specific cell type within the tissue. In the light of these limitations, we have considered co-expression as minimal necessary criteria. Our rationale is that the lack of co-expression would negate the hypothesis of miRNA-mediated regulation within the framework of variable expressivity.

We carried out an analysis of the expression of miRNA and the target genes using data given by Liang *et al.*[71] and Unigene, EST ProfileViewer (www.ncbi.nlm.nih.gov/UniGene), respectively, and screened for co-expressed target gene–miRNA pairs. We considered Boolean criteria for expression/non-expression states and not the level of expression, as variable expression itself is the phenomenon under analysis. We obtained 120 and 36 target gene–miRNA pairs for IP-VE and modifier genes, respectively, co-expressed in at least one tissue (Table 4, Supporting Information Table 3).

We applied another filter to prioritize mRNA–miRNA pairs for experimental validation by examining whether the target gene–miRNA pair is co-expressed in the tissue affected by the disease attributed to target gene mutation. The data for this analysis were available at OMIM for 40 genes in our list, of which 24 genes accounting for 43 target gene–miRNA pairs co-occurred in the affected tissue. Sample data are shown in Table 5 and complete data in Supporting Information Table 3. The observations can be illustrated in the case of *JAG1*, involved in cell communication during early embryonic development, mainly in cardiovascular development and cell-fate determination during hematopoiesis. Defects in *JAG1* lead to Alagille syndrome (OMIM 118450), an autosomal dominant disorder with highly variable severity of heart or liver disease between individuals with the same genotype. The disease also affects eyes, skin, and other tissues. We observed that *JAG1* mRNA occurred in numerous tissues, but is co-expressed with the targeting miRNA, hsa-mir-124a, in brain and with hsa-mir-214 in brain, liver, heart, and kidney, which are generally affected tissues in Alagille syndrome. It is to be noted that one cannot expect the absence of the miRNA in the unaffected or non-target tissues since a few hundred miRNAs are expected to regulate thousands of targets in various tissues and tissue-specific expression is reported for only a limited number of miRNA.[72]

We carried out a comparison of expression for miRNA, target mRNA, and the protein product using the available data. Figure 4 shows a contrast in the expression levels of hsa-mir-520a*, -491, and -199a* *versus* RB1, FLT4, and TGFB1, respectively, suggesting miRNA-mediated post-transcriptional regulation of the mRNA.

Table 4. Summary of co-expression analysis. The number of miRNA and target genes contributing to co-expressed target pairs in the IP-VE and modifiers in different tissues are given. Co-expression of target gene–miRNA pairs in affected tissue is shown for IP-VE genes. The numbers in parenthesis are the total number of genes/miRNA/pairs for which expression data is available in the public domain

	Co-expression		
Gene class	**Genes**	**miRNA**	**Pairs**
IP-VE	62 (68)	90 (113)	120 (152)
IP-VE in affected tissue	24 (40)	34	43 (79)
Modifiers	19 (19)	34 (37)	36 (39)

Table 5. Examples of IP-VE genes co-expressed with their targeting miRNAs in disease-affected tissue

Gene	Disease	OMIM ID	Targeting miRNA	miRNA co-expressed with target gene (affected tissue)
JAG1	Alagille syndrome	118450	hsa-mir-124a, 214	hsa-mir-124a (brain), hsa-mir-214 (brain, liver, heart, kidney)
SPG3A	Spastic paraplegia 3	182600	hsa-mir-370	hsa-mir-370 (brain)
PTPN11	Noonan syndrome	163950	hsa-mir-302c, 187, 24, 646	hsa-mir-302c (heart), hsa-mir-187 (brain, heart), hsa-mir-24 (brain, heart, muscle)
SPAST	Spastic paraplegia 4, autosomal dominant	182601	hsa-mir-92, 367, 421, 453	hsa-mir-92 (brain)
SFTPC	Surfactant metabolism dysfunction, pulmonary, 2	610913	hsa-mir-320, 331	hsa-mir-320 (lung), hsa-mir -331 (lung)

Differential regulation of products of homologous alleles: polymorphism in miRNA target site

SNP in the miRNA binding site can alter the interaction between the target site and miRNA,[73,74] and SNP in pre-miRNA can alter the biogenesis or the structure of the resulting mature miRNA[75] and, hence, the phenotype in question. When the source of miRNA is an intron, sequence polymorphism within the miRNA *per se* does not alter the mRNA of the source gene. In target mRNA, SNP can be in the seed region (Fig. 3) or the region flanking the target site. Depending on the nature of the base altered, SNP can change the intramolecular structure of the target transcript and, thus, affect its accessibility to miRNA as well as the stability of miRNA and mRNA interaction. This has been analyzed and the data for the known SNP in the 3′UTR are available in the dbSMR database (http://miracle.igib.res.in/polyreg).[76] Amongst the list of IP-VE genes, we found 21 target gene–miRNA pairs (contributed by 16 genes and 20 miRNAs) that show either gain or loss of miRNA binding due to occurrence of validated SNP(s) within a 200-bp region spanning the target site (Table 6 and Supporting Information Table 4). For example, 3′UTR of *SPAST* (SPG4, OMIM 182601) and

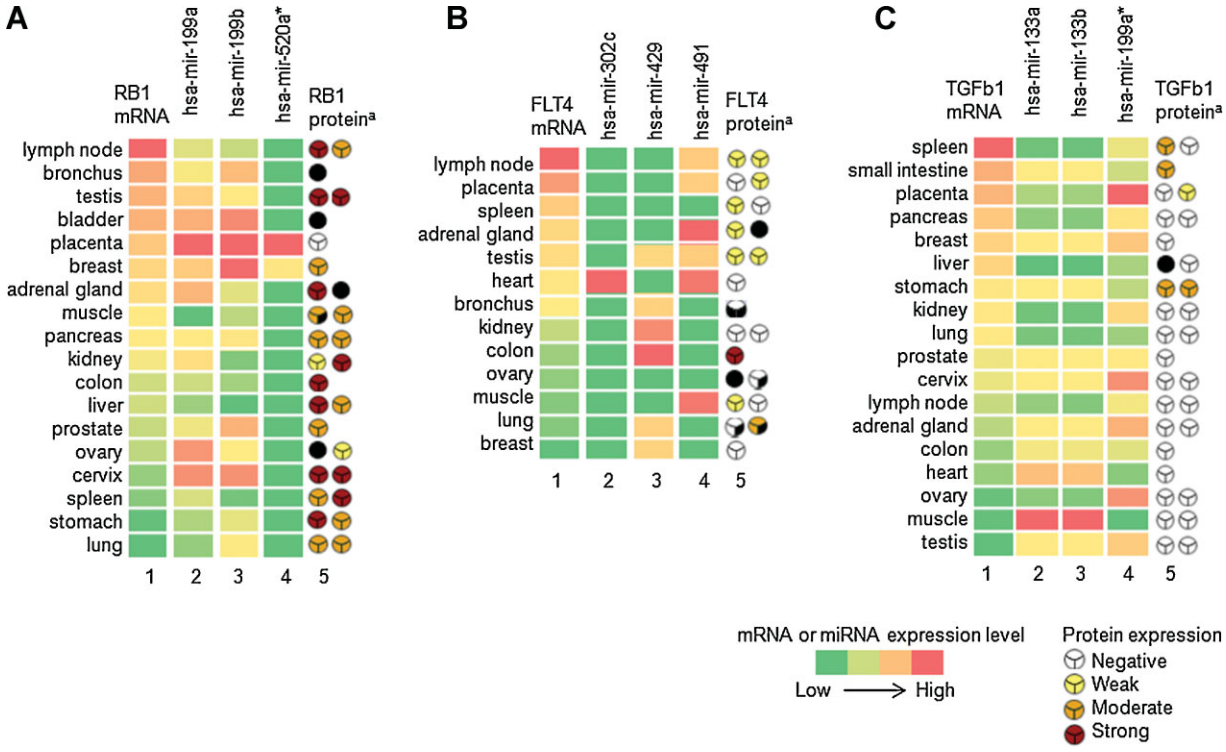

Figure 4. Comparison of expression of target mRNA (lane 1), miRNA (lanes 2–4), and the protein (lane 5) for (A) RB1, (B) FLT4, and (C) TGFb1. [a]Protein expression data were retrieved from www.proteinatlas.org. The protein level is inversely related to miRNA level, for instance, RB1 protein level is high in all cases except in placenta where the level of mir-520a* is high.

Table 6. Occurrence of SNP affecting miRNA binding to the targeted IP-VE genes

Gene	Disease	OMIM ID	miRNA	SNP ID	SNP position[a]	Effect of SNP
TRPM7	Amyotrophic lateral sclerosis-parkinsonism/ dementia complex 1	105500	hsa-mir-642	rs11070795	−20	Gain
TP73L	Ankyloblepharon-ectodermal defects-clefting syndrome, adult syndrome	106260, 103285	hsa-mir-491	rs35659283	−110	Loss
				rs35785527	−67	Loss
				rs36099321	−27	Loss
TOR1A	Torsion dystonia 1	128100	hsa-mir-133b	rs1182	74	Gain
TGFBR2	Aortic aneurysm, Familial Thoracic 3	608967	hsa-mir-642	rs11466534	15	Gain
			hsa-mir-34b	rs11466532	8	Loss
SPAST[b]	Spastic paraplegia 4, autosomal dominant	182601	hsa-mir-453	rs9789593	−67	Loss
			hsa-mir-92		−41	Loss
SMPD1	Niemann-Pick disease, type A	257200	hsa-mir-34a	rs1803160	−160	Loss
SFTPC[b]	Surfactant metabolism dysfunction, pulmonary, 2	610913	hsa-mir-331	rs8192331	20	Loss
PNKD	Paroxysmal nonkinesigenic dyskinesia 1	118800	hsa-mir-452*	rs921970	−56	Loss
PLN	Cardiomyopathy, dilated, 1P	609909	hsa-mir-449	rs12198461	89	Loss
NOS1	Pyloric stenosis, infantile hypertrophic 1	179010	hsa-mir-362	rs9658563	−117	Loss
LMBR1	Acheiropody	200500	hsa-mir-448	rs12530942	48	Loss
EYA4	Cardiomyopathy, dilated, 1J	605362	hsa-mir-136	rs3734279	−107	Gain
EXT1	Exostoses, multiple, type I	133700	hsa-mir-655	rs6988932	−26	Gain
EVC	Weyers acrofacial dysostosis	193530	hsa-mir-588	rs11939264	127	Gain
COL1A1	Caffey disease	114000	hsa-let-7d	rs1061947	12	Gain
			hsa-let-7g		12	Gain
			hsa-let-7a		13	Gain
			hsa-let-7f		13	Gain
CFTR	Cystic fibrosis transmembrane conductance regulator	602421	hsa-mir-645	rs17140308	15	Gain
SUV39H2 (m)		606503	hsa-mir-383	rs11594111	62	Loss

SNP in target gene within the 200-bp up- or downstream region of miR target site, causing gain or loss of miRNA binding is shown.
Disease refers to that caused by mutation in the respective gene.
[a]Position of SNP in the target site considering the first base pairing with miRNA as +1.
[b]Marks the target genes co-expressed with miRNA in affected tissue shown in Table 5.
(m), modifier gene.
hsa-mir-452*: nomenclature for miRNA given at http://microrna.sanger.ac.uk.

SFTPC (SMPD2, OMIM 610913) have SNP rs9789593 (G/T) and rs8192331(C/G), respectively, that can result in loss of binding of hsa-mir-92 (*SPAST*) and -331 (*SFTPC*). This alters the pairing pattern of bases in the mRNA–miRNA binding region, resulting in a structural change and hence loss of binding (http://miracle.igib.res.in/polyreg). As shown in Table 5, hsa-mir-92 and hsa-mir-331 are co-expressed with the respective genes in the affected tissue. Thus, in the background of a mutation in the gene, variable stability of miRNA–mRNA interaction due to SNP in the target region may result in variable expressivity seen in SPG4 and SMPD2. The predicted consequences of SNP-mediated allele distinction are depicted in Fig. 3.

We identified four SNPs in the miRNA binding site of modifier genes, three of which map to the ''seed region'' in the target site. We also examined the presence of SNP(s) in the pre-miRNA sequences that could possibly alter the biogenesis or the structure of the resulting mature miRNA and detected seven SNPs mapping to six pre-miRNA targeting twelve IP-VE genes (data not shown).

Concluding remarks

miRNAs have emerged as a new class of post-transcriptional regulators that participate in various processes, such as development, cell differentiation, apoptosis, and cancer. Our analysis of the possible role of miRNA in IP-VE, adds to the repertoire of regulatory circuits in which miRNA exert their influence. The stochasticity and affinity of multiple miRNAs binding to a transcript or the occurrence of polymorphism in the target or the miRNA-source sequence could add to the conditional expression of the product of one of the homologous alleles. Our results strongly suggest that miRNA-mediated regulation can underlie the mechanism of IP-VE of genes. Our analysis offers a theoretical framework for experimental validation and the need for network analysis of miRNA-mediated regulation of diseases showing IP-VE.

It is significant to note the predominance of developmental genes among the miRNA-targeted IP-VE genes. This could tender an explanation for the lasting effect of this mechanism

of differential expression of homologous alleles at early developmental stages, as a discernable effect on the phenotype can be brought about through regulatory cascades during development.

Methods

Dataset source

A list of 143 IP-VE genes was generated using OMIM (http://www.ncbi.nlm.nih.gov/Omim/) using the following keywords and criteria: single gene autosomal disorder, showing incomplete penetrance and/or "variable expressivity" of the disease phenotype, mutation, or gene. Thirty-one epigenetic modifiers reported by Surani *et al*. were taken;[40] and their human orthologs were taken from HomoloGene, NCBI. The list of housekeeping genes used as control was taken from Eisenberg and Levanon.[67]

miRNA target prediction

3′UTR sequences for the genes were downloaded from ENSEMBL release 42. Mature miRNA sequences were downloaded for the 470 miRNAs available at miRBase release 9.0 (http://microrna.sanger.ac.uk/).[77] miRNA target predictions were carried out by taking a consensus[65] of miRanda,[61] TargetScan,[62] and RNAhybrid.[63] Redundant target gene–miRNA pairs generated due to alternate transcripts of gene were removed to obtain unique pairs.

Analysis of expression data

Tissue expression profiles of miRNA were taken from Liang *et al*.,[71] and those of target mRNAs were retrieved from Unigene, EST ProfileViewer (www.ncbi.nlm.nih.gov/UniGene) as transcripts per million (TPM). We screened for co-expressed target gene–miRNA pairs in 27 tissues with the following criteria: copy number of miRNA as well as TPM count of its target gene in a tissue above "zero" signifies co-expression of the miRNA and its target gene. Information in context to disease phenotype or clinical symptom was obtained from OMIM. Protein expression data were retrieved from (www.proteinatlas.org) Protein Atlas version 3.0 – 2007.10.09. For the tissues having TPM above zero, target gene, miRNA, and protein expressions were compared.

SNP analysis

Information on SNPs occurring within 200 bp (up- and downstream) of target site and their effect on miRNA binding was retrieved from dbSMR database (http://miracle.igib.res.in/polyreg).[76] SNP(s) in pre- or mature miRNA were retrieved from UCSC genome browser.

Functional clustering

DAVID[69] was applied for functional classification and clustering of genes. Default cut-offs given at DAVID and p value ≤ 0.01 was considered for gene enrichment, at levels 3 and 4.

Acknowledgments: The authors thank Council of Scientific and Industrial Research (CSIR) India, NWP0036 grant for funding the project, and Shanti Chandrashekaran, Sridhar Sivasubbu, Vinod Scaria, and Anurag Aggarwal for critical reading of the paper. J. A., M. H., and R. B. thank Indian Council of Medical Research (ICMR) and CSIR for fellowships.

References

1. **Heyningen, V. V. and Yeyati, P. L.,** Mechanisms of non-Mendelian inheritance in genetic disease. *Hum Mol Genet* 2004. **13**: R225–R233.
2. **Zlotogora, J.,** Penetrance and expressivity in the molecular age. *Genet Med* 2003. **5**: 347–352.
3. **Stern, C.,** O. Vogt and the terms ''Penetrance'' and ''Expressivity''. *Am J Hum Genet* 1960. **12**: 141.
4. **Mitchell, M. E., Sander, T. L., Klinkner, D. B. and Tomita-Mitchell, A.,** The molecular basis of congenital heart disease. *Semin Thorac Cardiovasc Surg* 2007. **19**: 228–237.
5. **Accili, D., Kido, Y., Nakae, J., Lauro, D. and Park, B. C.,** Genetics of type 2 diabetes: insight from targeted mouse mutants. *Curr Mol Med* 2001. **1**: 9–23.
6. **Shin, D., Gilbert, F., Goldstein, M. and Schlegel, P. N.,** Congenital absence of the vas deferens: incomplete penetrance of cystic fibrosis gene mutations. *J Urol* 1997. **158**: 1794–1798.
7. **Amiel, J., Sproat-Emison, E., Garcia-Barcelo, M., Lantieri, F., Burzynski, G.,** *et al*. Hirschsprung disease, associated syndromes and genetics: a review. *J Med Genet* 2008. **45**: 1–14.
8. **Di, P. E., Brancaleoni, V., Moriondo, V., Besana, V. and Cappellini, M. D.,** Co-existence of two functional mutations on the same allele of the human ferrochelatase gene in erythropoietic protoporphyria. *Clin Genet* 2007. **71**: 84–88.
9. **Waseem, N. H., Vaclavik, V., Webster, A., Jenkins, S. A., Bird, A. C.,** *et al*. Mutations in the gene coding for the pre-mRNA splicing factor, PRPF31, in patients with autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci* 2007. **48**: 1330–1334.
10. **Dryja, T. P., Rapaport, J., McGee, T. L., Nork, T. M. and Schwartz, T. L.,** Molecular etiology of low-penetrance retinoblastoma in two pedigrees. *Am J Hum Genet* 1993. **52**: 1122–1128.
11. **Abouzeid, H., Schorderet, D. F., Balmer, A. and Munier, F. L.,** Germline mutations in retinoma patients: relevance to low-penetrance and low-expressivity molecular basis. *Mol Vis* 2009. **15**: 771–777.
12. **Delcuve, G. P., Rastegar, M. and Davie, J. R.,** Epigenetic control. *J Cell Physiol* 2009. **219**: 243–250.
13. **Surani, M. A.,** Reprogramming of genome function through epigenetic inheritance. *Nature* 2001. **414**: 122–128.
14. **Brockdorff, N.,** X-chromosome inactivation: closing in on proteins that bind Xist RNA. *Trends Genet* 2002. **18**: 352–358.
15. **Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T. and Wutz, A.,** A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. *PLoS Biol* 2004. **2**: E171.
16. **Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M.,** *et al*. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 2000. **405**: 486–489.
17. **Khosla, S., Aitchison, A., Gregory, R., Allen, N. D. and Feil, R.,** Parental allele-specific chromatin configuration in a boundary-imprinting-control

element upstream of the mouse H19 gene. *Mol Cell Biol* 1999. **19**: 2556–2566.
18. **Kitsberg, D., Selig, S., Brandeis, M., Simon, I., Keshet, I.,** *et al.* Allele-specific replication timing of imprinted gene regions. *Nature* 1993. **364**: 459–463.
19. **Gribnau, J., Hochedlinger, K., Hata, K., Li, E. and Jaenisch, R.,** Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes Dev* 2003. **17**: 759–773.
20. **Fournier, C., Goto, Y., Ballestar, E., Delaval, K., Hever, A. M.,** *et al.* Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO J* 2002. **21**: 6560–6570.
21. **Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. and Cavalli, G.,** Genome regulation by polycomb and trithorax proteins. *Cell* 2007. **128**: 735–745.
22. **Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U.,** *et al.* MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. *Nucleic Acids Res* 2007. **35**: 5402–5408.
23. **Garrick, D., Fiering, S., Martin, D. I. and Whitelaw, E.,** Repeat-induced gene silencing in mammals. *Nat Genet* 1998. **18**: 56–59.
24. **Mehta, A. K., Majumdar, S. S., Alam, P., Gulati, N. and Brahmachari, V.,** Epigenetic regulation of cytomegalovirus major immediate-early promoter activity in transgenic mice. *Gene* 2009. **428**: 20–24.
25. **Fu, Y. H., Kuhl, D. P., Pizzuti, A., Pieretti, M., Sutcliffe, J. S.,** *et al.* Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 1991. **67**: 1047–1058.
26. **Hagerman, R. J., Hull, C. E., Safanda, J. F., Carpenter, I., Staley, L. W.,** *et al.* High functioning fragile X males: demonstration of an unmethylated fully expanded FMR-1 mutation associated with protein expression. *Am J Med Genet* 1994. **51**: 298–308.
27. **Basel-Vanagaite, L., Pelet, A., Steiner, Z., Munnich, A., Rozenbach, Y.,** *et al.* Allele dosage-dependent penetrance of RET proto-oncogene in an Israeli-Arab inbred family segregating Hirschsprung disease. *Eur J Hum Genet* 2007. **15**: 242–245.
28. **Gouya, L., Puy, H., Robreau, A. M., Bourgeois, M., Lamoril, J.,** *et al.* The penetrance of dominant erythropoietic protoporphyria is modulated by expression of wildtype FECH. *Nat Genet* 2002. **30**: 27–28.
29. **Jezequel, P., Dubourg, C., Le, L. D., Odent, S., le Gall, J. Y.,** *et al.* Molecular screening of the CFTR gene in men with anomalies of the vas deferens: identification of three novel mutations. *Mol Hum Reprod* 2000. **6**: 1063–1067.
30. **Pastinen, T., Sladek, R., Gurd, S., Sammak, A., Ge, B.,** *et al.* A survey of genetic and epigenetic variation affecting human gene expression. *Physiol Genomics* 2004. **16**: 184–193.
31. **Cowles, C. R., Hirschhorn, J. N., Altshuler, D. and Lander, E. S.,** Detection of regulatory variation in mouse genes. *Nat Genet* 2002. **32**: 432–437.
32. **Krantz, I. D., Piccoli, D. A. and Spinner, N. B.,** Clinical and molecular genetics of Alagille syndrome. *Curr Opin Pediatr* 1999. **11**: 558–564.
33. **Campos-Xavier, B., Saraiva, J. M., Savarirayan, R., Verloes, A., Feingold, J.,** *et al.* Phenotypic variability at the TGF-beta1 locus in Camurati-Engelmann disease. *Hum Genet* 2001. **109**: 653–658.
34. **Caspi, M., Coquelle, F. M., Koifman, C., Levy, T., Arai, H.,** *et al.* LIS1 missense mutations: variable phenotypes result from unpredictable alterations in biochemical and cellular properties. *J Biol Chem* 2003. **278**: 38740–38748.
35. **Vithana, E. N., bu-Safieh, L., Pelosini, L., Winchester, E., Hornan, D.,** *et al.* Expression of PRPF31 mRNA in patients with autosomal dominant retinitis pigmentosa: a molecular clue for incomplete penetrance? *Invest Ophthalmol Vis Sci* 2003. **44**: 4204–4209.
36. **Imel, E. A., Hui, S. L. and Econs, M. J.,** FGF23 concentrations vary with disease status in autosomal dominant hypophosphatemic rickets. *J Bone Miner Res* 2007. **22**: 520–526.
37. **So, A. Y., Chaivorapol, C., Bolton, E. C., Li, H. and Yamamoto, K. R.,** Determinants of cell- and gene-specific transcriptional regulation by the glucocorticoid receptor. *PLoS Genet* 2007. **3**: e94.
38. **Ktistaki, E., Lacorte, J. M., Katrakili, N., Zannis, V. I. and Talianidis, I.,** Transcriptional regulation of the apolipoprotein A-IV gene involves synergism between a proximal orphan receptor response element and a distant enhancer located in the upstream promoter region of the apolipoprotein C-III gene. *Nucleic Acids Res* 1994. **22**: 4689–4696.
39. **Calkhoven, C. F., Muller, C. and Leutz, A.,** Translational control of gene expression and disease. *Trends Mol Med* 2002. **8**: 577–583.
40. **Surani, M. A., Hayashi, K. and Hajkova, P.,** Genetic and epigenetic regulators of pluripotency. *Cell* 2007. **128**: 747–762.
41. **Shankar, S. P., Fingert, J. H., Carelli, V., Valentino, M. L., King, T. M.,** *et al.* Evidence for a novel x-linked modifier locus for leber hereditary optic neuropathy. *Ophthalmic Genet* 2008. **29**: 17–24.
42. **Kontorovich, T., Cohen, Y., Nir, U. and Friedman, E.,** Promoter methylation patterns of ATM, ATR, BRCA1, BRCA2 and P53 as putative cancer risk modifiers in Jewish BRCA1/BRCA2 mutation carriers. *Breast Cancer Res Treat* 2008. 19.
43. **Levy-Lahad, E., Lahad, A., Eisenberg, S., Dagan, E., Paperna, T.,** *et al.* A single nucleotide polymorphism in the RAD51 gene modifies cancer risk in BRCA2 but not BRCA1 carriers. *Proc Natl Acad Sci USA* 2001. **98**: 3232–3236.
44. **Bergren, S. K., Chen, S., Galecki, A. and Kearney, J. A.,** Genetic modifiers affecting severity of epilepsy caused by mutation of sodium channel Scn2a. *Mamm Genome* 2005. **16**: 683–690.
45. **Owens, S. E., Broman, K. W., Wiltshire, T., Elmore, J. B., Bradley, K. M.,** *et al.* Genome-wide linkage identifies novel modifier loci of aganglionosis in the Sox10Dom model of Hirschsprung disease. *Hum Mol Genet* 2005. **14**: 1549–1558.
46. **Mattick, J. S. and Makunin, I. V.,** Small regulatory RNAs in mammals. *Hum Mol Genet* 2005. **14**: R121–R132.
47. **Xie, W., Ted, B. W. and Denman, R. B.,** Translational regulation by non-protein-coding RNAs: different targets, common themes. *Biochem Biophys Res Commun* 2008. **373**: 462–466.
48. **He, L. and Hannon, G. J.,** MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 2004. **5**: 522–531.
49. **Bartel, D. P. and Chen, C. Z.,** Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat Rev Genet* 2004. **5**: 396–400.
50. **Sevignani, C., Calin, G. A., Siracusa, L. D. and Croce, C. M.,** Mammalian microRNAs: a small world for fine-tuning gene expression. *Mamm Genome* 2006. **17**: 189–202.
51. **Lee, H. J., Palkovits, M. and Young, W. S., III,** miR-7b, a microRNA up-regulated in the hypothalamus after chronic hyperosmolar stimulation, inhibits Fos translation. *Proc Natl Acad Sci USA* 2006. **103**: 15669–15674.
52. **Zhu, S., Si, M. L., Wu, H. and Mo, Y. Y.,** MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). *J Biol Chem* 2007. **282**: 14328–14336.
53. **Karres, J. S., Hilgers, V., Carrera, I., Treisman, J. and Cohen, S. M.,** The conserved microRNA miR-8 tunes atrophin levels to prevent neurodegeneration in *Drosophila*. *Cell* 2007. **131**: 136–145.
54. **Visone, R., Russo, L., Pallante, P., De, M., Ferraro, A.,** *et al.* MicroRNAs (miR)-221 and miR-222, both overexpressed in human thyroid papillary carcinomas, regulate p27Kip1 protein levels and cell cycle. *Endocr Relat Cancer* 2007. **14**: 791–798.
55. **Jackson, R. J. and Standart, N.,** How do microRNAs regulate gene expression? *Sci STKE* 2007. **2007**: re1.
56. **Saetrom, P., Heale, B. S., Snove, O., Jr., Aagaard, L., Alluin, J.,** *et al.* Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic Acids Res* 2007. **35**: 2333–2342.
57. **Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L.,** *et al.* Combinatorial microRNA target predictions. *Nat Genet* 2005. **37**: 495–500.
58. **Hobert, O.,** Gene regulation by transcription factors and microRNAs. *Science* 2008. **319**: 1785–1786.
59. **Hobert, O.,** Common logic of transcription factor and microRNA action. *Trends Biochem Sci* 2004. **29**: 462–468.
60. **Hornstein, E. and Shomron, N.,** Canalization of development by microRNAs. *Nat Genet* 2006. **38**: S20–S24.
61. **John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C.,** *et al.* Human MicroRNA targets. *PLoS Biol* 2004. **2**: e363.
62. **Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. and Burge, C. B.,** Prediction of mammalian microRNA targets. *Cell* 2003. **115**: 787–798.

63. **Rehmsmeier, M., Steffen, P., Hochsmann, M. and Giegerich, R.,** Fast and effective prediction of microRNA/target duplexes. *RNA* 2004. **10**: 1507–1517.
64. **Martin, G., Schouest, K., Kovvuru, P. and Spillane, C.,** Prediction and validation of microRNA targets in animal genomes. *J Biosci* 2007. **32**: 1049–1052.
65. **Hariharan, M., Scaria, V., Pillai, B. and Brahmachari, S. K.,** Targets for human encoded microRNAs in HIV genes. *Biochem Biophys Res Commun* 2005. **337**: 1214–1218.
66. **Fraser, H. B., Hirsh, A. E., Giaever, G., Kumm, J. and Eisen, M. B.,** Noise minimization in eukaryotic gene expression. *PLoS Biol* 2004. **2**: e137.
67. **Eisenberg, E. and Levanon, E. Y.,** Human housekeeping genes are compact. *Trends Genet* 2003. **19**: 362–365.
68. **Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L. and Bradley, A.,** Identification of mammalian microRNA host genes and transcription units. *Genome Res* 2004. **14**: 1902–1910.
69. **Dennis, G., Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W.,** *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 2003. **4**: 3.
70. **Stark, A., Brennecke, J., Bushati, N., Russell, R. B. and Cohen, S. M.,** Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3′UTR evolution. *Cell* 2005. **123**: 1133–1146.
71. **Liang, Y., Ridzon, D., Wong, L. and Chen, C.,** Characterization of microRNA expression profiles in normal human tissues. *BMC Genomics* 2007. **8**: 166.
72. **Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N.,** *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 2007. **129**: 1401–1414.
73. **Georges, M., Coppieters, W. and Charlier, C.,** Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease. *Curr Opin Genet Dev* 2007. **17**: 166–176.
74. **Martin, M. M., Buckenberger, J. A., Jiang, J., Malana, G. E., Nuovo, G. J.,** *et al.* The human angiotensin II type 1 receptor +1166 A/C polymorphism attenuates microrna-155 binding. *J Biol Chem* 2007. **282**: 24262–24269.
75. **Iwai, N. and Naraba, H.,** Polymorphisms in human pre-miRNAs. *Biochem Biophys Res Commun* 2005. **331**: 1439–1444.
76. **Hariharan, M., Scaria, V. and Brahmachari, S. K.,** dbSMR: a novel resource of genome-wide SNPs affecting microRNA mediated regulation. *BMC Bioinformatics* 2009. **10**: 108.
77. **Griffiths-Jones, S., Grocock, R. J., Van DS, Bateman, A. and Enright, A. J.,** miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 2006. **34**: D140–D144.