Full text provided by www.sciencedirect.com

Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease

Michel Georges, Wouter Coppieters and Carole Charlier

The expression of at least a third of mammalian genes is post-transcriptionally fine-tuned by ~1000 microRNAs (miRNAs), assisted by the RNA silencing machinery, comprising tens of components. Polymorphisms and mutations in the corresponding sequence space (machinery, miRNA precursors and target sites) are likely to make a significant contribution to phenotypic variation, including disease susceptibility. Here we review basic miRNA biology in animals, survey the available evidence for DNA sequence polymorphisms affecting miRNA-mediated gene regulation and thus phenotype, and discuss their possible importance in the determination of complex traits.

Addresses
Unit of Animal Genomics, GIGA-R and Faculty of Veterinary Medicine, University of Liège, (B34), 1 Avenue de l'Hôpital, 4000-Liège, Belgium

Corresponding author: Georges, Michel (michel.georges@ulg.ac.be)

Current Opinion in Genetics & Development 2007, **17**:1–11

This review comes from a themed issue on
Genetics of disease
Edited by Robert Nussbaum and Leena Peltonen

0959-437X/$ – see front matter
© 2007 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.gde.2007.04.005

Introduction

Identifying the genes and mutations underlying phenotypic variation is one of the primary objectives of modern genetics, especially for traits of medical or agronomic importance. The vast majority of causal mutations identified to date alter the primary sequence and hence the structure of proteins. They are either missense or nonsense mutations, insertion–deletions in the open reading frame, or mutations causing splicing errors. This over-representation of protein-altering variants amongst known causative mutations does not necessarily equate with their true contribution to common inherited variation. It is more likely to reflect ascertainment bias due to the fact that it is easier to predict the effect of protein-altering variants on gene function, and possibly to the fact that they cause larger allele-substitution effects that are consequently easier to study.

It is increasingly recognized that regulatory mutations could make a significant contribution to genetic variation, especially for complex traits, including disease susceptibility. Along those lines, efforts for the systematic identification of regulatory variants (rSNPs) affecting transcript levels *in cis* are ongoing (for example [1]). The most common interpretation of such *cis* effects is that the corresponding variants are modulating the activity of regulatory elements, including promotors and enhancers. Recently, a new layer of post-transcriptional miRNA-mediated gene regulation has been discovered and shown to control the expression levels of a large proportion of our genes. Moreover, at least two studies have demonstrated that genetic variants might influence organismal phenotypes by perturbing miRNA-mediated gene regulation [2••,3••]. We herein review and discuss the presently available evidence for polymorphic miRNA-mediated gene regulation and its importance for phenotypic variation and disease.

A quick tour of mammalian miRNA biology

Genomics

MicroRNAs (miRNA) are a novel class of single-stranded regulatory RNAs of ~22 nucleotides found in metazoans. miRNAs are the final product of a multistep maturation process that starts with the generation of a transcript — referred to as primary miRNA (pri-miRNA) — that hosts one or more miRNA precursors with a characteristic hairpin structure. Most pri-miRNAs are regular RNA polymerase II transcripts that undergo capping, splicing and polyadenylation. At least half of them are protein-encoding mRNAs, whereas the remainder are so-called 'mRNA-like non-coding RNAs'. The miRNA precursor hairpins are usually embedded in the introns of their host genes (~80%), but can be found in exons or across exon–intron boundaries. Tandem clusters of co-transcribed precursors account for more than half of the known miRNAs (reviewed in [4]). Note that some SINE-associated miRNAs have recently been shown to be pol III- rather than pol II-dependent [5].

Although the total number of miRNA precursors in the genome of *Caenorhabditis elegans* and *Drosophila melanogaster* is unlikely to exceed much more than 130 and 160, respectively (for example [6]), the number of miRNA precursors in the typical mammalian genome is estimated to be of the order of 1000 (reviewed in [4]). Some authors even dare to suggest numbers as high as 25 000 on the basis of bioinformatics prediction [7]. Among the 439 human miRNAs known at present, 9%, 22%, 29% and 34% are conserved across invertebrates, vertebrates, mammals and primates, respectively, and 5% are specific to humans (for example [8]). These numbers highlight the fact that miRNA-mediated gene regulation is an

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

ancient process, but that it is characterized by considerable evolutionary plasticity. In general, recently acquired miRNAs seem to be characterized by lower or more restricted expression than ancient miRNAs. This could be an obligatory path towards acquisition of novel functions without perturbing fitness [9].

Cropping and dicing

miRNA precursor hairpins are defined by (i) flanking single-stranded basal segments, and (ii) an imperfect three-helical turn stem, surmounted by (iii) a terminal loop of 10–20 nt. The future mature miRNA occupies either the ascending (5′ donors) or descending (3′ donors) branch of the upper stem, or in exceptional cases both. miRNA precursor hairpins do not have specific consensus sequences. It is their secondary structure that positions the double-stranded RNA-binding protein DGCR8 (known as Pasha in *Drosophila*) such that its ribonuclease III partner Drosha cleaves the stem at approximately one helical turn from the base. This cleavage releases the so-called 'pre-miRNA' — a free hairpin characterized by a staggered 2-nt 3′ overhang defining either the 5′ (5′ donors) or the 3′ (3′ donors) end of the mature miRNA [10]. Cropping of intronic miRNAs occurs co-transcriptionally, prior to completion of the splicing reaction. This slows the splicing process down but otherwise does not interfere with it, because flanking exons are physically bound in a splicing commitment complex [11]. Drosha-dependent cropping of pri-miRNA is temporally regulated during early mouse development and might be inhibited during cancer development [12]. Some pre-miRNAs have been shown to undergo A-to-I editing mediated by ADAR deaminases. This might affect the rate of cropping, as shown for pri-miR-142 [13], or the sequence of the mature miRNA and hence the spectrum of target genes, as shown for the miR-376 cluster [14].

Exportin-5 recognizes pre-miRNAs by their typical 2-nt 3′ overhangs and moves them to the cytoplasm by means of the Ran-GDP/Ran-GTP transport system. Once in the cytoplasm, pre-miRNAs are further processed by Dicer. This protein is thought to bind the 3′ overhang end of a pre-miRNA with its PAZ domain, while positioning the pre-miRNA upper stem in the dual RNase III center with the help of its double-stranded RNA-binding domain. Dicer then inflicts a staggered cut akin to Drosha, generating a double-stranded miRNA–miRNA* duplex that is freed from the terminal loop [15].

The mature miRNA is then ferried to a RNA-induced silencing complex (miRISC or miRNP), while the complementary passenger mRNA strand is eliminated. The miRNA is distinguished from the miRNA* by virtue of the lower stability of the duplex at the 5′ end of the miRNA. This processing step requires the assistance of the double-stranded RNA-binding protein TRBP (the HIV *trans*-activating response RNA-binding protein) [16]. Notably, as many as 10^3 to 10^4 molecules might accumulate per cell for a given miRNA species, which is one to two orders of magnitude more than for most mRNAs [17].

Slicing

The best characterized components of miRISCs are the Argonaute proteins with characteristic PAZ and PIWI domains, of which there are four in mammals (Ago1–Ago4). Structural studies of prokaryotic Argonaute-like proteins strongly suggest that the phosphorylated 5′ residue of the miRNA is docked in a tilted position in a highly conserved basic pocket of the PIWI domain, while contacts with the sugar–phosphate backbone of residues 2 to 5 force these residues into a quasi-helical configuration promoting base-pairing. The 3′-OH end of the miRNA is anchored in a pocket of the PAZ domain (reviewed in [18]). Other components of miRISC include the Vasa intronic gene (VIG) protein, the Tudor-SN protein, Fragile-X-related protein, the putative RNA helicase Dmp68, and Gemin3 and Gemin4, whose precise roles in RNA interference remain to be determined (reviewed in [19]).

The miRNA guides the miRISC complex to target mRNAs, which it recognizes by virtue of complementary target sequences. In animals, target sites are preferentially located in the 3′ untranslated region (UTR), whereas in plants they are mostly in the open reading frame. If the complementarity between the miRNA and its target is close to perfect, the two will engage in a two-turn double helix that places the target strand facing residues 10 and 11 of the miRNA in front of the RNase H-like DDH catalytic triad of the PIWI domain, resulting in target cleavage or 'slicing'. In *D. melanogaster*, 5′ and 3′ mRNA slicing products are further degraded by the exosome–SKI complex and by XRN1, respectively (reviewed in [20]). Note that Ago2 is the only Argonaute endowed with slicer activity in mammals. Moreover, although perfect complementarity between miRNAs and their target is in essence the rule in plants, it is exceptional in animals having been demonstrated only for *HOXB8* and *PEG11* [21,22].

Slicing-independent silencing

The typical miRNA–target interaction in animals is based on imperfect 'fuzzy' complementarity. It is mainly dependent on perfect Watson–Crick (WC) base-pairing occurring in the region between bases 2 and 8 of the miRNA — referred to as the miRNA 'seed' — which corresponds closely to the residues adopting a quasi-helical conformation when bound to Argonaute. All known target sites show perfect WC pairing over at least four consecutive residues in the seed. This basic four-residue helix needs to be further stabilized by 2–3 additional perfect WC interactions in the seed, creating so-called '5′ dominant' (5D) sites, or by extensive complementarity on the 3′ side of the miRNA, creating so-

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

called '3′ compensatory' (3C) sites. 5D sites can be subdivided in 'seed' sites (5DS) that base-pair exclusively with the miRNA seed, and 'canonical' sites (5DC) that show base-pairing in the 3′ end in addition to a 7- or 8-nt seed match [23]. Whether 5D or 3C, base-pairing seems to be avoided in the middle of the miRNA — in other words, at the expected slice site. This observation explains why miRNA regulation is usually not accompanied by target slicing in animals.

The most striking effect of miRNA–target interaction in mammals is a reduction in the amount of detectable protein. It was initially suggested that this reduction was due to inhibition of the translation process after initiation, because miRNAs and their targets were shown to co-sediment with translating polysomes. That protein production was nevertheless reduced was first attributed to premature ribosome drop-off and more recently to interference — by an unknown mechanism — with the accumulation of growing polypeptides on otherwise normally translating polysomes. Other groups have produced evidence supporting inhibition of cap-dependent translation initiation (reviewed in [24]).

The effect on protein production was initially thought to occur without an appreciable effect on target mRNA levels. However, it has recently become apparent that in many cases target mRNA levels are in fact reduced, albeit to a lesser degree than the protein levels - thus certainly not fully accounting for it. miRNAs were shown to accelerate deadenylation of maternal targets during early embryogenesis in zebrafish and reporter targets in 293T cells. miRISC components including Argonaute, miRNA and their targets were shown to localize to P-bodies, known sites of mRNA decapping and degradation (reviewed in [19,20,24]). Subsequently, it has been shown that effective miRNA-mediated silencing requires several P-body components including GW182, the DCP1–DCP2 decapping complex, the RCK/p54 decapping coactivator, the XRN1 5′ to 3′ exonuclease and the CCR4/NOT deadenylase. Note that individuals suffering from autoimmune forms of motor and sensory ataxic polyneuropathy and primary biliary cirrhosis might have antibodies against P-body components. Moreover, miR-16 was shown to function with RISC and the sequence-specific RNA binding protein TTP to target an mRNA containing an AU-rich element (ARE) for degradation (reviewed in [19,20,24]).

Interestingly, miRNA-mediated downregulation and recruitment to P-bodies of the *CAT-1* mRNA has been recently shown to be reversible in cells subjected to different stress conditions, and this derepression requires binding of HuR (an AU-rich-element binding protein) to the 3′ UTR of *CAT-1* [25]. Along the same lines, BDNF treatment has been shown to relieve *miR-134*-dependent repression of Limk1 translation in cortical neurons [26]. Reversibility is clearly not compatible with slicer-mediated silencing and might thus explain the preference for slicer-independent silencing in animal cells.

Recently, *miR-29b* was shown in HeLa cells to be preferentially located in the nucleus [24]. The hexanucleotide sequence at the 3′ end of *miR-29b* seems to be necessary and sufficient for nuclear localization. Nuclear localization apparently accelerates the turnover of *miR-29b*, but this was not observed for other miRNAs appended with the hexanucleotide and consequently directed to the nucleus. The precise meaning of these findings remains to be determined, but they suggest that miRNAs might have nuclear functions such as transcriptional or splicing control, in addition to canonical translational control in the cytoplasm [27].

Thus, miRNA-mediated gene silencing in animals seems to be a complex activity involving many mechanisms that interact in ways that largely remain to be determined.

miRNA targets in animals

In plants, miRNA-mediated gene regulation is based on a quasi one-to-one relationship between an miRNA and its nearly perfectly complementary target. Although the number of experimentally validated miRNA targets remains limited in animals (compiled in TarBase [28]), bioinformatics points towards a crowded and intricate miRNA-based regulatory network in animals. In mammals, for example, the number of evolutionary conserved matches to 7-nt seeds in 3′ UTRs exceeds that for control heptamers by a factor of ~2.2, enabling the prediction of more than 15 000 5D target sites with a false discovery rate (FDR) of 0.45. This corresponds to more than 200 target genes per miRNA and at least 30% of our genes being under miRNA-mediated 5D control. The specificity of the prediction can be improved — at the cost of sensitivity — by increasing the number of species defining conservation, by using 8-nt seeds, by requiring an 'A anchor' to face the first base of the miRNA (irrespective of its identity), by accounting for 3′-pairing stability or by demanding co-occurrence of multiple target sites in the same 3′ UTR (for example [23,29,30]).

In *Drosophila*, the number of sites including a conserved seed of less than 7 nt with 3′-pairing energy above specific thresholds exceeds the number obtained with appropriate controls, thus pointing towards the detection of 3C sites (albeit an order of magnitude less than 5D sites) [23]. Distinct miRNAs might share the same seed sequence. It is thought that 3C sites evolved to allow target recognition by individual members of a seed-sharing miRNA family, which in part explains why the interspecies conservation of mature miRNAs usually extends over their whole length rather than just the seed sequence.

Most studies have exploited sequence information from known miRNAs to identify cognate target sites. By con-

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

trast, Xie *et al.* [31] have been able to predict putative miRNA target sites by identifying octamer motifs in 3'UTRs characterized by unusually high motif conservation scores (i.e. proportion of conserved occurrences amongst all occurrences of a given motif). Highly conserved octamer motifs that didn't match seeds of known miRNA have in turn allowed prediction of novel miRNAs.

Functions of animal miRNAs

The first miRNAs identified by forward genetic analysis of developmental defects in *C. elegans* and *D. melanogaster* highlighted their role as developmental switches (reviewed in [32]). The subsequent identification of a majority of developmental genes amongst miRNA targets in plants corroborated this key role in development and differentiation.

Studying the distribution of target sites across the animal gene collection indicates, however, that animal miRNAs are endowed with additional functions. Paradoxically, target site avoidance has been as informative as target site content. Quite logically, given that they have to operate in all cell types, the 3′ UTRs of housekeeping genes are depleted in target sites in general and (presumably as a means to avoid target sites) are shorter than the 3′ UTRs of other genes [33]. Some genes are under evolutionary pressure to avoid target sites for specific rather than all miRNAs. Again quite logically, because they have to cohabitate peacefully, these are the genes that are strongly expressed in the same cells as the corresponding miRNA. In some cases, the signature of 5D target site avoidance is strong enough to allow prediction of the tissue where a given miRNA is expressed. Genes that are under selective pressure to avoid target sites are referred to as 'anti-targets'. Interestingly, signatures of target site avoidance are not restricted to 3′ UTRs [33,34].

Some genes are under selective pressure to maintain target sites. Conserved target sites have been found to be enriched among genes that are strongly expressed in tissues that are spatially or temporally adjacent to the domains where the cognate miRNA is preferentially expressed. This observation suggests that miRNAs accentuate tissue transition by dampening residual expression of genes expressed in the precursor tissue or dampening precocious expression of genes specifying the derived tissue [33,34]. This role as differentiation guards is compatible with the exquisite tissue specificity of the expression profile of miRNAs, as observed in zebrafish [35], and by the fact that signatures of miRNA expression seem remarkably effective in predicting the developmental lineage and differentiation stage of tumors [36].

The number of non-conserved target sites exceeds that of conserved ones by a factor of approximately ten. Most of the non-conserved target sites have the potential to mediate downregulation by the cognate miRNA, as demonstrated in reporter silencer assays [34]. Although most non-conserved sites are probably non-functional *in vivo*, because miRNA and target have non-overlapping expression domains, evidence from population genetics (see below) indicates that at least some of the non-conserved sites are selectively constrained and thus functional. Thus, target site acquisition also seems to be an ongoing and dynamic evolutionary process.

Genes often contain multiple target sites for either the same or distinct miRNAs. The combination of multiple target sites is an effective means to fine-tune expression according to the precise needs of distinct tissues, especially at low, otherwise noisy levels. A nice illustration of this 'denoising function' has been recently provided by the demonstration that *miR-9a* loss in *D. melanogaster* results in the expression levels of the *Senseless* gene to trespass a critical threshold in an increased number of neuroectodermal progenitor cells, thereby generating surplus sensory organ precursor cells per proneural cluster [37]. Fine-tuning expression levels and buffering of genetic noise might thus be another important function of miRNAs [38,39].

Inherited variation in miRNA-mediated gene regulation

miRNA-mediated gene silencing thus emerges as a key regulator of cellular differentiation and homeostasis to which metazoans devote a considerable amount of sequence space. This sequence space is bound to suffer its toll of mutations, of which some will be selectively neutral and others will be advantageous or more often at least slightly deleterious. DNA sequence polymorphisms (DSPs) occurring in this sequence space certainly contribute to phenotypic variation, including disease susceptibility and agronomically important traits. An important issue is how important their contribution actually is. DSPs might affect miRNA-mediated gene regulation by perturbing core components of the silencing machinery, by affecting the structure or expression level of miRNAs, or by altering target sites (Table 1).

DSP in core components of the silencing machinery might affect the overall efficacy of silencing. Mutations that drastically perturb RNA silencing will obviously be rare, given their predictable highly deleterious consequences. Nevertheless, DSPs with subtle effects on gene function might occur. Because distinct targets might be more or less sensitive to variations in miRNA concentration or silencing efficiency, such DSPs might affect some pathways more than others.

Specific miRNA–target interactions might be influenced by mutations affecting either the miRNA or its target. On the miRNA side of the equation: (i) the sequence of the mature miRNA might be altered, thereby stabilizing or destabiliz-

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

Table 1

Categories of DSPs affecting miRNA-mediated gene regulation.

Target	miRNA	Silencing machinery
DSPs altering miRNA recognition sites in the target Altering existing target sites *Stabilizing or destabilizing the interaction with the miRNA* Creating illegitimate target sites	**DSPs altering the sequence of the miRNA** Stabilizing or destabilizing the interaction with the target	**DSPs altering the amino acid sequence of silencing components** **DSPs altering the concentration of silencing components**
DSPs altering the 3′ UTR of the target (e.g. Polymorphic polyadenylation)	**DSPs altering the concentration of the miRNA** CNVs encompassing the pri-miRNA DSPs altering the transcription rate of the pri-miRNA *Cis* or *trans*-acting DSPs affecting the processing efficiency of the pri- or pre-miRNA	**Copy number variants encompassing silencing components**

ing its interaction with targets; (ii) mutations in the pri- or pre-miRNA might affect stability or processing efficiency; (iii) mutations acting in *cis* or *trans* on the pri-miRNA promoter might influence the transcription rate; and (iv) copy number variants (CNVs) might affect the number of copies of the miRNA or the integrity of the pri-miRNA host. On the target side of the equation: (i) mutations might affect functional target sites, thereby destabilizing or stabilizing the interaction with the miRNA; (ii) mutations might create illegitimate miRNA target sites (in the 3′ UTR or even in other segments of the transcript) that will be particularly relevant if occurring in anti-targets; and (iii) mutations causing polymorphic alternative polyadenylation might affect the content of a gene in target sites.

It has become apparent that somatic mutations or epimutations affecting miRNA-mediated silencing contribute to cancer. Dicer dysregulation and deletions of Argonaute genes have been reported in several types of tumor. miRNA levels have been found to be generally decreased in tumors, possibly marking de-differentiation. The downregulation or deletion of specific tumor-suppressing miRNAs (e.g. *miR-15a*, *miR-16-1*, *miR-143*, *miR-145* and *let-7*) and overexpression of putative oncogenic miRNAs (e.g. *miR-21* and *miR-155*) might directly contribute to cancer progression. Oncogene tenors, including *RAS*, *BCL2* and *KIT*, are targets of dysregulated miRNAs, whereas the *miR-17-19b-1* cluster is involved in a complex interplay with *MYC*. Moreover, somatic mutations altering binding sites for *miR-221*, *miR-222* and *miR-146* have been observed in the 3′ UTR of the *KIT* oncogene in papillary thyroid carcinoma. The contribution of somatic perturbations of RNA silencing to cancer has been reviewed recently [40] and will not be considered further here.

DSPs affecting the silencing machinery

Mutations underlying two well-known inherited diseases, DiGeorge syndrome and Fragile X syndrome, involve components of the RNA silencing machinery.

DGCR8, Drosha's cropping partner, was named for 'DiGeorge syndrome critical region gene 8' [41]. It is indeed located within the 3-Mb or 1.5-Mb microdeletions observed in 98% of individuals with DiGeorge, or more precisely, del22q11 syndrome [42]. del22q11 is the most common microdeletion syndrome, afflicting 1 in 4000 live births. Symptoms include pharyngeal (congenital cardiovascular defects, craniofacial anomalies and aplasia of the thymus and parathyroid) and neurobehavioral (learning, cognitive, attention deficits and psychiatric disorders) phenotypes. The syndrome is typically characterized by variable expressivity and incomplete penetrance. Loss of function of the *Tbx1* and *Crkl* homologs in the mouse recapitulate human symptomatology, at least in part, but haploinsufficiency of other genes in the region might contribute to the complex clinical pattern. Could this apply to *DGCR8*, which is located in the 1.5-Mb critical region? *Dgcr8* knockout mice are not viable and the phenotypic characterization of hemizygous mice awaits further characterization. However, while homozygous *Dgcr8* knockout ES cells suffer anomalous miRNA biogenesis as well as proliferation and differentiation defects, heterozygous ES cells are in essence not distinguishable from wild type ES cells [43]. This doesn't support *DGCR8* haploinsufficiency and hence an important contribution to *del22q11* symptomatology.

Fragile X is another of the most common inherited mental retardations, afflicting 1 in 4000 men and 1 in 8000 women. About 99% of cases result from madumnal inheritance of an expanded (from pre- to full mutation length) CGG repeat in the 5′ UTR of the X-linked gene *FMR1*, resulting in epigenetic gene silencing that causes a fully penetrant disease in men and a 50% penetrant disease in women. *FMR1* encodes the RNA-binding protein FMRP, which is thought to regulate mRNA transport and localized translation of a selected group of mRNAs. The *Drosophila* ortholog *dfmr1* has been shown to associate with components of RISC complexes, suggesting that FMRP-mediated translational suppression might

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

operate through the RNA silencing machinery (reviewed in [44]). *FMR1* has two paralogs in vertebrates: *FXR1* and *FXR2*.

Screening the dbSNP database for DSPs causing non-synonymous substitutions revealed 59 coding single nucleotide polymoprhisms (cSNPs) in the 14 proteins involved in RNA silencing mentioned above. In addition, consulting a recent catalogue of 1447 human CNVs found among 270 healthy Africans, Asians and Europeans [45] identified DSPs spanning or interrupting *DGCR8* (CNV id: 1363; 283 kb), *DROSHA* (CNV id 427; 174 kb) and *XRN1* (CNV id: 273; 401 kb). The corresponding DSPs are all compiled in the Patrocles database (www.patrocles.org).

DSPs affecting miRNA structure or expression

Humans

Iwai and Naraba [46] sequenced 173 pre-miRNAs in 96 Japanese individuals and identified ten SNPs. One of these mapped to nucleotide 10 of the mature *miR-30c-2* and could potentially have an impact on its interaction with targets, although this was not verified experimentally. According to the authors, the nine other SNPs, of which three mapped to the loop and seven to the stem, are unlikely to affect miRNA processing.

Calin *et al.* [47] identified a C-to-T transition at a position 7-bp 3′ of the *miR-16-1* precursor in two individuals with chronic lymphocytic leukemia. This SNP was not observed in 160 control subjects. The affected individuals were found to be heterozygous in both cancerous and normal cells, suggesting that this is a rare SNP. The expression level of *miR-16-1* was found to be considerably reduced in the tumors. The notion that this reduced expression might be caused by the C-to-T SNP was suggested by the lower levels of pre- and mature *miR-15a* and *miR-16-1* obtained after transfecting 293 cells with vectors expressing the mutant versus the wild-type sequences. Deletions and downregulation of the closely linked *miR-15a* and *miR-16-1* pair on 13q13.4 are common findings in chronic lymphocytic leukemia and are thought to promote cell proliferation by releasing downregulation of their anti-apoptotic *BCL2* target.

We (M Georges *et al.*, unpublished) have screened dbSNP to identify additional DSPs that might affect the structure or processing of human miRNAs. Ninety SNPs were found within pre-miRNAs. The mature miRNA is affected by 18 of these SNPS, of which six are located within the seed (http://www.patrocles.org). We also searched for miRNA host genes with evidence for inherited variation in expression levels (http://www.patrocles.org). As an example, Spielman *et al.* [48] reported 1097 genes with expression levels (in lymphoblastoid cell lines) that differ significantly between Caucasians and Asians. Eight of these genes are host genes of one miRNA each. Previously, Morley *et al.* [49] used linkage analysis to search for *cis*- and *trans*-acting expression quantitative trait loci (eQTLs) affecting the expression level of 3554 genes in lymphoblastoid cell lines. They reported 984 eQTLs with a genome-wide *P* value of 0.05, corresponding to an FDR of ~0.18. Five of the corresponding genes corresponded to known miRNA host genes. In addition, we searched for miRNA hosts among genes showing allelic expression imbalance in heterozygous individuals. Pant *et al.* [50] identified 732 genes showing allelic differential expression in white blood cells with an estimated FDR of 0.116. Six of these genes corresponded to known miRNA host genes. Likewise, Chakravarti *et al.* have identified 193 genes showing allelic imbalance at the mRNA level, of which one proved to be a miRNA host (A. Chakravarti *et al.*, personal communication). It clearly remains to be established whether variations in host gene expression level translate into variation in miRNA concentration, but such an assertion seems reasonable. Lastly, we identified 43 miRNAs mapping to the previously mentioned catalogue of 1447 human CNVs [45]. Along similar lines, Wong *et al.* [51] have recently reported 14 CNVs encompassing 21 known human miRNAs.

Non-humans

Obviously, mutations affecting miRNAs underlie the developmental defects that led to the discovery of miRNAs in the first place, including *lin-4*, *let-7* and *lsy-6* in *C. elegans* and *bantam* in *D. melanogaster* (reviewed in [32]). But what about non-model organisms? A sequence comparison of miRNAs in the genome of the Kaposi's-sarcoma-associated herpesvirus (KSHV) in latently infected cell lines revealed a G-to-A transition in the passenger branch of the *miR-K5* stem that disrupts base-pairing and DROSHA-mediated processing [52]. As a consequence, the mature *miR-K5* is not expressed in cells containing KSHV with the A genotype. Neither the phenotypic consequences of this miRNA loss nor the target genes of the 11 KSHV miRNAs, however, are known. The same study reported DSPs in three other KSHV miRNAs including two polymorphisms that alter the sequence of the mature *miR-K10* at positions 2 and 16. Assuming that *miR-K10* targets a 5D site, the first of these DSPs has a high likelihood of affecting recognition of the target.

An interesting example of how perturbed miRNA expression might contribute to phenotypic variation is provided by the callipyge phenotype in sheep. This muscular hypertrophy results from ectopic expression of the DLK1 protein in skeletal muscle, caused by a mutation (*CLPG*) in a silencer regulating the imprinted *DLK1–GTL2* gene cluster. Individuals inheriting the *CLPG* mutation from their father have increased transcript levels of the paternally expressed *DLK1* and *PEG11* genes, whereas those receiving the mutation from their mother have higher levels of maternally expressed mRNA-like non-coding RNA genes (*GTL2*, *MEG8* and *MIRG*). The

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

maternally expressed non-coding RNA genes are host to a large cluster of C/D small nucleolar RNAs and miRNAs [53,54]. Surprisingly, DLK1 protein and consequently the phenotype are only observed in $+^{\mathrm{Mat}}/CLPG^{\mathrm{Pat}}$ sheep ('polar overdominance'). This observation is intriguing because *CLPG/CLPG* sheep have received the *CLPG* mutation from their father and show the expected increase in *DLK1* mRNA. There is strong evidence that the lack of DLK1 protein and thus phenotype in homozygous mutant sheep is due to translational inhibition of the *DLK1* transcripts by miRNAs that are hosted by the maternally expressed non-coding RNA genes. Such miRNA-mediated *trans*-inhibition has indeed been formally demonstrated in the same locus for the paternally expressed *PEG11* gene [55•,56•].

DSPs affecting miRNA target sites

Humans

The first direct evidence that mutations in target sites might affect phenotype came from studies of the gene *SLITRK1* (Slit and Trk-like family member 1) [2••]. *SLITRK1* was considered to be a strong candidate for Tourette's syndrome because of its proximity to a *de novo* inversion observed in an affected child. Mutation scanning in 174 unrelated individuals with Tourette's syndrome revealed a frameshift mutation generating a grossly truncated protein and, remarkably, two independent occurrences of the same G-to-A transition in the 3′ UTR of the gene (var321) that was not encountered in 1800 healthy controls. var321 is predicted to stabilize the interaction of *SLITRK1* with *miR-189* by changing a G:U wobble pair to an A:U WC pair at position nine of the miRNA. This assertion is supported by the observations that the 3′ UTR of *SLITRK1* indeed facilitates *miR-189*-mediated downregulation of a luciferase reporter, by the finding that the var321 mutation significantly accentuates this downregulation, and by the overlapping expression domains of *SLITRK1* and *miR-189* in the central nervous system. Taken together, these observations support the idea that var321 has a direct role in causing Tourette's syndrome [2••].

Recently, Sethupathy *et al.* [57•] searched the databases for SNPs mapping to experimentally confirmed miRNA target sites compiled in TarBase [28]. They identified one SNP (rs5186) located in a non-conserved *miR-155* 5D site in the 3′ UTR of the gene encoding the angiotensin II type-1 receptor (*AGTR1*). The A-to-C transversion replaces an A:U WC match with a C:U mismatch at position 5 of the miRNA seed, and can thus confidently be predicted to suppress the miRNA–target interaction. Consistently, *miR-155*-dependent downregulation of reporter vectors endowed with the *AGTR1* 3′ UTR was reduced by the A-to-C transversion. Interestingly, in several studies, the rs5186 C-allele of *AGTR1* has been associated with hypertension, which might result from *AGTR1* overexpression. Moreover, because *miR-155* maps to chromosome 21, Setupathy *et al.* speculate that *miR-155* overexpression in trisomics — which they confirmed experimentally — accounts for the lower levels of diastolic and systolic blood pressure associated with Down syndrome [57•].

Züchner *et al.* [58] have identified two point mutations in the 3′ UTR of the gene encoding receptor expression-enhancing protein 1 (*REEP1*) in individuals suffering from hereditary spastic paraplegia. *REEP1* is a strong candidate gene because protein-disrupting mutations have been reported among individuals affected with this disorder. The two 3′ UTR mutations are located 7-bp apart in a highly conserved segment that shows perfect WC complementarity to the seed sequence (nt 2–8) of *miR-140*, which is known to be expressed in the central nervous system. The c.606+50G→A mutation replaces a G:U wobble match with a A:U WC match at position 9 of *miR-140* and could thus stabilize the interaction. The c.606+43G→T replaces a G:U wobble match with a U:U mismatch at position 15 of *miR-140*. Züchner *et al.* [55•] surmise that both mutations might stabilize the miRNA–target interaction, thereby downregulating *REEP1*, which would cause the disease. The hypothesis seems reasonable, especially for the c.606+50G→A mutation, but further experimental support is certainly needed.

To facilitate the identification of additional DSPs that might modulate miRNA–target interactions, we and others have searched the databases for SNPs that alter the 3′ UTR content in 5D miRNA target sites [3••,59••]. We defined 5D target sites either as one of the 540 octamer motifs identified by Xie *et al.* [28] on the basis of their unusual motif conservation score or by their complementarity to known miRNA seeds in accordance with Lewis *et al.* [26]. Among the ~120 000 known 3′ UTR SNPs, 19 913 modified the 3′ UTR content in putative targets. We refer to these SNPs as 'Patrocles SNPs' or pSNPs (http://www.patrocles.org). The corresponding derived alleles (determined from the ortholog in chimpanzee) destroy 785 conserved target sites, destroy 9470 non-conserved target sites and create 10 283 novel sites. pSNPs destroying conserved target sites have the highest likelihood of affecting gene function and thus are prime candidates for further studies. To assist in the identification of functionally relevant pSNPs affecting non-conserved sites, we generated coexpression scores for all miRNA–target pairs. Indeed, for a pSNP to have a chance to be functionally relevant, miRNA and target need to have overlapping expression profiles.

Signatures of selection

What is the evidence that any of the candidate pSNPs listed above truly affect gene function and thus phenotype? Indirect evidence that a significant proportion of them are functional can be obtained from population genetics. Indeed, pSNPs without an appreciable effect

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

on gene function will evolve neutrally, subject only to the vagaries of random genetic drift, whereas pSNPs affecting gene function might undergo positive, negative or balancing selection by means of their effect on phenotype. Selection might leave distinct signatures on the level of inter-species divergence, intra-species variability, allelic distribution and linkage disequilibrium (reviewed in [60]).

Evidence for strong signatures of selection has been recently obtained for pSNPs [59••]. This work first showed that conserved target sites are more severely depleted in SNPs than are other conserved sequences in the 3′ UTR and that this effect is strongest at the 3′ end (matching the miRNA seed). These results are to be expected if it assumed that these sites are subject to selective constraints causing both reduced polymorphisms and divergence. Less expected is the observation that pSNPs that either destroy or create non-conserved target site are also underrepresented [61]. This finding strongly suggests that a significant proportion of human target sites have the potential to be functional even if they are not conserved across mammals. It is important to stress that by 'functional' we do not mean that the target sites have the potential to respond to the presence of a cognate miRNA if artificially exposed to it, but that they truly interact with the miRNA *in vivo*.

An under-representation of pSNPs in miRNA target sites indicates that selection has wiped out several functional pSNPs. Does this mean that the pSNPs that are still segregating in human populations are devoid of function? To address this idea, Chen and Rajewski [59••] have compared the distributions of derived allele frequencies of pSNPs with those of synonymous and non-synonymous coding SNPs and other 3′ UTR SNPs. pSNPs are characterized by an excess of low derived allele frequencies when compared with all other classes, consistent with the notion that at least a fraction of the derived alleles are deleterious. This has been found to be the case for pSNPs that destroy conserved target sites, but also (albeit to a lesser extent) for pSNPs that destroy or create non-conserved target sites for coexpressed miRNAs. Remarkably, conserved target sites seem to be more constrained than non-synonymous coding SNPs or SNPs affecting other conserved 3′ UTR sequences, whereas non-conserved target sites for coexpressed miRNAs are at least as constrained as other conserved 3′ UTR sequences. Using the so-called 'Poisson random field framework', Chen and Rajewski [59••] estimate that as many as 85% of conserved sites and 30–50% of non-conserved sites (for coexpressed miRNAs) are under negative selection. It can thus be concluded that a significant fraction of pSNPs has an effect on phenotype, probably including disease susceptibility.

The argument developed above is restricted to derived alleles with deleterious effects. Some neo-mutations must be selectively advantageous. Chen and Rajewski [59••] looked for signatures of positive selection affecting pSNPs in conserved target sites, by means of the Fst parameter measuring population differentiation. They found that one SNP (rs1054528) in a conserved binding site for *miR-204* and *miR-211* in the *Map1lc3b* gene is present in 87% of Africans but is almost absent in Europeans and Asians, and noted that post-transcriptional misregulation of *Map1lc3b* has been implicated in giant axonal neuropathy and Fragile X syndrome. An alternative approach for the identification of pSNPs undergoing positive selection is the 'long-range haplotype test' (reviewed in [60]), and genotyping of the HapMap population with pSNPs is progressing towards that goal.

Non-humans

In mammals, the most convincing example of a phenotype resulting from the perturbation of miRNA–target interaction is probably the muscular hypertrophy of Texel sheep [3••]. A QTL scan was undertaken to identify the genes and mutations underlying the polygenic hypermuscularity of Texel sheep. A QTL accounting for a quarter of the muscular hypertrophy was mapped to chromosome 2 and then fine-mapped to a chromosome segment containing the *GDF8* gene encoding a muscle-specific chalone known as myostatin.

Subsequent detailed genetic analysis pinpointed a G→A substitution in a poorly conserved part of the *GDF8* 3′ UTR that showed perfect association with QTL genotype. This point mutation has been shown to create (by substituting an A:U WC match for a G:U wobble match at seed position 6) an illegitimate 5D target site for *miR-1* and *miR-206*, two seed-sharing miRNAs highly expressed in skeletal muscle. The G→A substitution has been found to be associated with a threefold reduction in circulating myostatin levels and a 1.5-fold reduction in *GDF8* transcript levels, and to enhance *miR-1*- and *miR-206*-dependent downregulation of luciferase reporters.

Conclusions

Given the pervasiveness of miRNA-mediated gene regulation, in the future we will almost certainly witness a flurry of reports describing mutations that affect disease susceptibility and other phenotypes by modulating miRNA-mediated gene regulation. Whereas loss of specific miRNAs has been associated with severe developmental defects in model organisms such as *C. elegans* and *D. melanogaster*, highly penetrant mutations seem more and more the exception rather than the rule in animals. As a matter of fact, the systematic knock down of *C. elegans* miRNAs does not yield an obvious phenotype in most cases, which probably highlights both the high degree of redundancy built into the miRNA regulatory network and the fine-tuning rather than switching function of most miRNAs.

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

Mutations affecting miRNA–target interactions seem more likely to have subtle phenotypic effects, creating hyper- or hypomorphic alleles as opposed to causing simple loss or gain of function — akin to the findings for Tourette's syndrome, hypertension and the Texel muscular hypertrophy. A significant contribution to the inheritance of traits and defects showing Mendelian inheritance can be more or less excluded. DSPs perturbing miRNA-mediated silencing might, however, have a more significant impact on complex traits, which — after all — encompass most inherited phenotypes. It can indeed be argued, on evolutionary grounds (for example [62]), that mutations with modest effects account for most of the standing variation underlying adaptive traits. Polymorphisms affecting the interaction between miRNAs and their targets are unlikely to cause drastic perturbations, but rather they probably underlie modest alterations of gene output and hence phenotype, providing a more effective gradual path for adaptation and maybe even speciation (for example [63]).

Update

A database, PolymiRTS (http://compbio.utmem.edu/miRSNP/), has been established that compiles pSNPs in human and mice and provides information about co-localization with reported *cis*-acting eQTLs as well as physiological QTLs [64]. Also, an additional compilation of human SNPs affecting miRNA precursors and target sites has been reported, and includes preliminary evidence based on the long-range haplotype test for positive selection acting on three pSNPs [65].

Acknowledgements

We are grateful to S Antonarakis, A Chakravarti and V Cheung for sharing unpublished information, and to H Takeda for comments on the manuscript. We acknowledge support from the European Union (FW6 Callimir STREP and Epigenome NoE), the Belgian Science Policy organization (SSTC Genefunc and Biomagnet PAI), the Belgian FNRS, from the Communauté Française de Belgique (Game and Biomod ARC), and the University of Liège.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Pastinen T, Ge B, Hudson TJ: **Influence of human genome polymorphism on gene expression**. *Hum Mol Genet* 2006, **15**:R9-R16.

2. •• Abelson JF, Kwan KY, O'Roak BJ, Baek DY, Stillman AA, Morgan TM, Mathews CA, Pauls DL, Rasin MR, Gunel M *et al.*: **Sequence variants in SLITRK1 are associated with Tourette's syndrome**. *Science* 2005, **310**:317-320.
The first identification of a 3′UTR mutation that by stabilizing the interaction with *miR-189* creates a hypomorphic *SLTRK1* allele that predisposes to Tourette's syndrome.

3. •• Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibe B, Bouix J, Caiment F, Elsen JM, Eychenne F *et al.*: **A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep**. *Nat Genet* 2006, **38**:813-818.
The identification of a point mutation that creates an illegitimate miR-1/miR-206 target site in the 3′UTR of the ovine *GDF8* gene, thereby creating a hypomorphe promoting increased muscle mass.

4. Kim VN, Nam JW: **Genomics of microRNA**. *Trends Genet* 2006, **22**:165-173.

5. Borchert GM, Lanier W, Davidson BL: **RNA polymerase III transcribes human microRNAs**. *Nat Struct Mol Biol* 2006, **13**:1097-1101.

6. Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP: **Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans***. *Cell* 2006, **127**:1193-1207.

7. Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I: **A pattern-based method for the identification of microRNA binding sites and their corresponding heteroduplexes**. *Cell* 2006, **126**:1203-1217.

8. Berezikov E, Thuemmler F, van Laake LW, Kondova I, Bontrop R, Cuppen E, Plasterk RH: **Diversity of microRNAs in human and chimpanzee brain**. *Nat Genet* 2006, **38**:1375-1377.

9. Chen K, Rajewsky N: **The evolution of gene regulation by transcription factors and microRNAs**. *Nat Rev Genet* 2007, **8**:93-103.

10. Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, Sohn SY, Cho Y, Zhang BT, Kim VN: **Molecular basis for the recognition of primary microRNAs by the Drosha–DGCR8 complex**. *Cell* 2006, **125**:887-901.

11. Kim YK, Kim VN: **Processing of intronic microRNAs**. *EMBO J* 2007, **26**:775-783.

12. Thomson JM, Newman M, Parker JS, Morin-Kensicki EM, Wright T, Hammond SM: **Extensive post-transcriptional regulation of microRNAs and its implications for cancer**. *Genes Dev* 2006, **20**:2202-2207.

13. Yang W, Chendrimada TP, Wang Q, Higuchi M, Seeburg PH, Shiekhattar R, Nishikura K: **Modulation of microRNA processing and expression through RNA editing by ADAR deaminases**. *Nat Struct Mol Biol* 2006, **13**:13-21.

14. Kawara Y, Zinshteyn B, Sethupathy P, Iizasa H, Hatzigeorgiou AG, Nishikura K: **Redirection of silencing targets by adenosine-to-inosine editing of miRNAs**. *Science* 2007, **315**:1137-1140.

15. Filipowicz W, Jaskiewicz L, Kolb FA, Pillai RS: **Post-transcriptional gene silencing by siRNAs and miRNAs**. *Curr Opin Struct Biol* 2005, **15**:331-341.

16. Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R: **Human RISC couples microRNA biogenesis and posttranscriptional gene silencing**. *Cell* 2005, **123**:631-640.

17. Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, Burge CB, Bartel DP: **The microRNAs of *Caenorhabditis elegans***. *Genes Dev* 2003, **17**:991-1008.

18. Parker JS, Barford D: **Argonaute: a scaffold for the function of short regulatory RNAs**. *Trends Biochem Sci* 2006, **31**:622-630.

19. Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R: **Control of translation and mRNA degradation by miRNAs and siRNAs**. *Genes Dev* 2006, **20**:515-524.

20. Eulalio A, Behm-Ansmant I, Izaurralde E: **P bodies: at the crossroads of post-transcriptional pathways**. *Nat Rev Mol Cell Biol* 2007, **8**:9-22.

21. Yekta S, Shih IH, Bartel DP: **MicroRNA-directed cleavage of HOXB8 mRNA**. *Science* 2004, **304**:594-596.

22. Davis E, Caiment F, Tordoir X, Cavaille J, Ferguson-Smith A, Cockett N, Georges M, Charlier C: **RNAi-mediated allelic *trans*-interaction at the imprinted *Rtl1/Peg11* locus**. *Curr Biol* 2005, **15**:743-749.

23. Brennecke J, Stark A, Russell RB, Cohen SM: **Principles of microRNA–target recognition**. *PLoS Biol* 2005, **3**:e85.

24. Pillai RS, Bhattacharyya SN, Filipowicz W: **Repression of protein synthesis by miRNAs: how many mechanisms?** *Trends Cell Biol* 2007, **17**:118-126.

25. Bhattacharyya SN, Habermacher R, Martine U, Closs EI, Filipowicz W: **Relief of microRNA-mediated translational**

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

repression in human cells subjected to stress. *Cell* 2006, **125**:1111-1124.

26. Schratt GM, Tuebing F, Nigh EA, Kane CG, Sabatini ME, Kiebler M, Greenberg ME: **A brain-specific microRNA regulates dendritic spine development**. *Nature* 2006, **439**:283-289.

27. Hwang HW, Wentzel EA, Mendell JT: **A hexanucleotide element directs microRNA nuclear import**. *Science* 2007, **315**:97-100.

28. Sethupathy P, Corda B, Hatziegeorgiou AG: **TarBase: a comprehensive database of experimentally supported animal microRNA targets**. *RNA* 2006, **12**:192-197.

29. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets**. *Cell* 2005, **120**:15-20.

30. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N: **Combinatorial microRNA target predictions**. *Nat Genet* 2005, **37**:495-500.

31. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M: **Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals**. *Nature* 2005, **434**:338-345.

32. Alvarez-Garcia I, Miska EA: **MicroRNA functions in animal development and human disease**. *Development* 2005, **132**:4653-4662.

33. Stark A, Brennecke J, Bushati N, Russell RB, Cohen SM: **Animal microRNAs confer robustness to gene expression and have a significant impact on 3′ UTR evolution**. *Cell* 2005, **123**:1133-1146.

34. Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, Burge CB, Bartel DP: **The widespread impact of mammalian microRNAs on mRNA repression and evolution**. *Science* 2005, **310**:1817-1821.

35. Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, de Bruijn E, Horvitz HR, Kauppinen S, Plasterk RH: **MicroRNA expression in zebrafish embryonic development**. *Science* 2005, **309**:310-311.

36. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA *et al.*: **MicroRNA expression profiles classify human cancers**. *Nature* 2005, **435**:834-838.

37. Li Y, Wang F, Lee JA, Gao FB: **MicroRNA-9a ensures the precise specification of sensory organ precursors in *Drosophila***. *Genes Dev* 2006, **20**:2793-2805.

38. Bartel DP, Chen CZ: **Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs**. *Nat Rev Genet* 2004, **5**:396-400.

39. Hornstein E, Shomron N: **Canalization of development by microRNAs**. *Nat Genet* 2006, **38(Suppl)**:S20-S24.

40. Esquela-Kerscher A, Slack FJ: **Oncomirs — microRNAs with a role in cancer**. *Nat Rev Cancer* 2006, **6**:259-269.

41. Shiohama A, Sasaki T, Noda S, Minoshima S, Shimizu N: **Molecular cloning and expression analysis of a novel gene DGCR8 located in the DiGeorge syndrome chromosomal region**. *Biochem Biophys Res Commun* 2003, **304**:184-190.

42. Lindsay EA: **Chromosomal microdeletions: dissecting del22q11 syndrome**. *Nat Rev Genet* 2001, **2**:858-868.

43. Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R: **DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal**. *Nat Genet* 2007, **39**:380-385.

44. Jin P, Alisch RS, Warren ST: **RNA and microRNAs in fragile X mental retardation**. *Nat Cell Biol* 2004, **6**:1048-1053.

45. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W *et al.*: **Global variation in copy number in the human genome**. *Nature* 2006, **444**:444-454.

46. Iwai N, Naraba H: **Polymorphisms in human pre-miRNAs**. *Biochem Biophys Res Commun* 2005, **331**:1439-1444.

47. Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M *et al.*: **A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia**. *N Engl J Med* 2005, **353**:1793-1801.

48. Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG: **Common genetic variants account for differences in gene expression among ethnic groups**. *Nat Genet* 2007, **39**:226-231.

49. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG: **Genetic analysis of genome-wide variation in human gene expression**. *Nature* 2004, **430**:74374-74377.

50. Pant PV, Tao H, Beilharz EJ, Ballinger DG, Cox DR, Frazer KA: **Analysis of allelic differential expression in human white blood cells**. *Genome Res* 2006, **16**:331-339.

51. Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE, Lam WL: **A comprehensive analysis of common copy-number variations in the human genome**. *Am J Hum Genet* 2007, **80**:91-104.

52. Gottwein E, Cai X, Cullen BR: **A novel assay for viral microRNA function identifies a single nucleotide polymorphism that affects Drosha processing**. *J Virol* 2006, **80**:5321-5326.

53. Seitz H, Youngson N, Lin SP, Dalbert S, Paulsen M, Bachellerie JP, Ferguson-Smith AC, Cavaille J: **Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene**. *Nat Genet* 2003, **34**:261-262.

54. Seitz H, Royo H, Bortolin ML, Lin SP, Ferguson-Smith AC, Cavaille J: **A large imprinted microRNA gene cluster at the mouse Dlk1–Gtl2 domain**. *Genome Res* 2004, **14**:1741-1748.

55. • Davis E, Jensen CH, Schroder HD, Farnir F, Shay-Hadfield T, Kliem A, Cockett N, Georges M, Charlier C: **Ectopic expression of DLK1 protein in skeletal muscle of padumnal heterozygotes causes the callipyge phenotype**. *Curr Biol* 2004, **14**: 1858-1862.
See annotation [56•].

56. • Davis E, Caiment F, Tordoir X, Cavaille J, Ferguson-Smith A, Cockett N, Georges M, Charlier C: **RNAi-mediated allelic trans-interaction at the imprinted *Rtl1/Peg11* locus**. *Curr Biol* 2005, **15**:743-749.
These two publications by Davis *et al.* [55•,56•] summarize recent advances in the molecular understanding of the non-Mendelian mode of inheritance (polar overdominance) of the callipyge muscular hypertrophy in sheep, which involves ectopic expression of imprinted genes including miRNAs, and miRNA-mediated *trans* interaction between the maternal and paternal chromosomes.

57. • Sethupathy P., Borel C, Gagnebin M, Grant GR, Hatzigeorgiou AG, Antonorakis, SE: **Human *miR-155* on chromosome 21 differentially interacts with its polymorphic target in the *AGTR1* 3'UTR – a mechanism for functional SNPs related to phenotypes**. *Am J Hum Genet* 2007, in press.
The identification of an SNP that destroys a conserved *miR-155* site in the 3′UTR of the *AGTR1* gene thus generating a hypermorphic allele associated with hypertension.

58. Zuchner S, Wang G, Tran-Viet KN, Nance MA, Gaskell PC, Vance JM, Ashley-Koch AE, Pericak-Vance MA: **Mutations in the novel mitochondrial protein REEP1 cause hereditary spastic paraplegia type 31**. *Am J Hum Genet* 2006, **79**:365-369.

59. •• Chen K, Rajewsky N: **Natural selection on human microRNA binding sites inferred from SNP data**. *Nat Genet* 2006, **38**:1452-1456.
The demonstration — using a population genetics approach — that miRNA target sites are severely constraint selectively, that the derived alleles of an important fraction of pSNPs (both conserved and non-conserved) are slightly deleterious, and that at least one pSNP might have undergone selective sweeps in the African population.

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005

60. Bamshad M, Wooding SP: **Signatures of natural selection in the human genome**. *Nat Rev Genet* 2003, **4**:99-111.

61. Georges M, Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Hiard S, Bibé B, Bouix J, Caiment F *et al.*: **Polymorphic miRNA–target interactions**. In *Regulatory RNAs. Cold Spring Harbor Symposium on Quantitative Biology, Vol. 71*. Edited by Stillman B, Stewart D. Cold Spring Harbor: Cold Spring Harbor Laboratory; 2006, **71**:343-350.

62. Orr HA: **The genetic theory of adaptation: a brief history**. *Nat Rev Genet* 2005, **6**:119-127.

63. Plasterk RH: **Micro RNAs in animal development**. *Cell* 2006, **124**:877-881.

64. Bao L, Zhou M, Wu L, Lu L, Goldowitz D, Williams RW, Cui Y: **PolymiRTS Database: linking polymorphisms in microRNA target sites with complex traits**. *Nucleic Acids Res* 2007, **35**:D51-D54.

65. Saunders MA, Liang H, Li WH: **Human polymorphism at microRNAs and microRNA target sites**. *Proc Natl Acad Sci USA* 2007, **104**:3300-3305.

Please cite this article in press as: Georges M, et al., Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease, Curr Opin Genet Dev (2007), doi:10.1016/j.gde.2007.04.005