BMC
Genomics

RESEARCH ARTICLE　　　　　　　　　　　　　　　　　　　**Open Access**

Small RNA expression and strain specificity in the rat

Sam EV Linsen, Elzo de Wit, Ewart de Bruijn and Edwin Cuppen*

Abstract

Background: Digital gene expression (DGE) profiling has become an established tool to study RNA expression. Here, we provide an in-depth analysis of small RNA DGE profiles from two different rat strains (BN-Lx and SHR) from six different rat tissues (spleen, liver, brain, testis, heart, kidney). We describe the expression patterns of known and novel micro (mi)RNAs and *piwi*-interacting (pi)RNAs.

Results: We confirmed the expression of 588 known miRNAs (54 in antisense orientation) and identified 56 miRNAs homologous to known human or mouse miRNAs, as well as 45 new rat miRNAs. Furthermore, we confirmed specific A to I editing in brain for *mir-376a/b/c* and identified *mir-377* as a novel editing target. In accordance with earlier findings, we observed a highly tissue-specific expression pattern for all tissues analyzed. The brain was found to express the highest number of tissue-specific miRNAs, followed by testis. Notably, our experiments also revealed robust strain-specific differential miRNA expression in the liver that is caused by genetic variation between the strains. Finally, we identified two types of germline-specific piRNAs in testis, mapping either to transposons or in strand-specific clusters.

Conclusions: Taken together, the small RNA compendium described here advances the annotation of small RNAs in the rat genome. Strain and tissue-specific expression patterns furthermore provide a strong basis for studying the role of small RNAs in regulatory networks as well as biological process like physiology and neurobiology that are extensively studied in this model system.

Background

miRNAs are ~22 nt-long, single stranded RNA molecules that mediate post-transcriptional regulation of gene expression by directing the RNA-induced Silencing Complex (RISC) to the 3' untranslated region (UTR) of target mRNAs [1,2]. As a result, translation is inhibited and/or the mRNA degraded [3,4]. The target spectrum of a miRNA is mostly defined by the seed, i.e. the 1st or 2nd 7 nt, which hybridizes to the target mRNA [5,6]. miRNAs can both act as developmental switches [7-9] or subtly tune expression, when tight regulation of a gene is required [10,11]. Thousands of mRNAs are expected to be under regulatory control of miRNAs [12,13] and the presence or absence of a single miRNA has been shown to affect, albeit modestly, the level of thousands of proteins [14,15]. Thus, miRNAs form a complex regulatory network affecting the majority of genes.

A second developmentally vital class of small RNAs are the *piwi*-interacting (pi)RNAs [16,17], which play a role in the formation of the germ line. In mammals, these ~27 nt ssRNAs are expressed in the reproductive organs, mainly the testis [18,19], where two types can be distinguished. The pre-pachytene piRNAs, which are repeat- and transposon-derived, likely play a role in guiding DNA methylation to repeats, thereby silencing transposons [20] and preventing genome instability. Conversely, the pachytene piRNAs are mostly derived from a selected set of genomic clusters that show a very strong strand bias. The function of these genomic clusters, however, remains elusive [18-21].

To a certain extent, development and homeostasis of organ systems depend on miRNAs [22,23] and piRNAs [24]. The laboratory rat (*Rattus norvegicus*) is a model organism in which organ-systems physiology has been studied for decades [25]. Recent advances in techniques to genetically modify the rat [26-30] enables detailed analyses of rat physiology at molecular levels. Furthermore, well-established genetic systems, such as congenic,

* Correspondence: e.cuppen@hubrecht.eu
1 Hubrecht Institute-KNAW & University Medical Center Utrecht, Cancer Genomics Center, Utrecht, The Netherlands
Full list of author information is available at the end of the article

consomic and recombinant inbred lines are versatile tools for studying the effect of genetic variation on quantitative traits such as blood pressure [31] or gene expression [32] (corresponding to quantitative trait loci (QTLs) and expression (e)QTLs, respectively). Comprehensive small RNA inventories and profiles are instrumental in such genetical genomics and systems biological approaches, as they serve as a resource for annotation of the genome. Small RNAs are important players in many regulatory processes and are thus important for understanding disease etiology. The rat small RNA inventory described here will also be important for understanding human disease, since many rat models were selected to reflect clinical symptoms [33].

Conserved expression specificity of miRNAs has been described for a number of organ systems or cell cultures, based on (deep) sequencing approaches [34-36]. Simultaneously, species-specific miRNAs have been identified in closely related species [34,36-38], indicating that miRNAs are evolutionary dynamic. The availability of comprehensive species-specific miRNA profiles of different tissues and organ systems is an important requirement for elucidating the biological roles that miRNAs fulfill. More exhaustive profiling will likely improve existing profiles and increase insight in the basis of quantitative and qualitative variations in miRNA expression. We therefore performed digital gene expression (DGE) profiling of small RNAs from six tissues, i.e. brain, liver, spleen, heart, testis and kidney of the BN-Lx and SHR rat inbred strains, the founder strains of the BXH/HXB recombinant inbred panel [39]. We identified 588 known miRNAs (54 in antisense orientation) and 101 new rat miRNAs, originating from 276 and 61 precursor (pre) miRNA loci, respectively. Thirty-one of these pre-miRNAs were not previously characterized in rat, but were found to be homologous to mouse or human loci; 30 novel candidate pre-miRNA loci do not have an apparent homologue in these species. By generating DGE profiles for liver from three individuals from each strain, we observed strain-specific differential miRNA expression for 4 miRNAs. Finally, we identified thousands of piRNAs in the testis samples. The dataset described here greatly contributes to our understanding of miRNA divergence, variation and expression and may be a valuable resource in evolutionary analyses as well as in the interpretation of regulatory networks and functional genomics experiments in the rat.

Results and discussion
miRNA identification
We collected the small RNA fraction and prepared small RNA sequencing libraries from 6 different tissues (i.e. whole brain, liver, spleen, heart, testis and kidney) from two rat inbred strains (polydactyly-luxate syndrome brown Norway (BN-Lx) and spontaneous hypertensive rats (SHR), adult males). The libraries were sequenced on the SOLiD platform version 2 (ABI), generating 115 million small RNA sequence reads. Of all raw reads, 41.9 million could be mapped to the rat genome (see Additional file 1, **Table S1 for individual libraries**). The length distribution of the vast majority of small RNA reads was between 18 and 23 nt, corresponding with the expected size for miRNAs. Exceptions to this observation were testis small RNAs, which are generally longer (Additional file 2, **Figure S1A**). In all tissues except testis, the majority of the reads mapped to known miRNAs [40]; small RNAs from testis were mostly derived from repeats or small RNA classes other than miRNAs most likely corresponding to (pre-)pachytene piRNAs [18-21] (Additional file 2, **Figure S1B and see below**). We aimed to address global profiling as well as RNA editing and for profiling we only included perfectly mapping reads (6.4 million). Each tissue dataset consisted out of approximately 400,000 miRNA reads.

We identified 688 miRNAs, derived from 276 of the known miRNA loci (miRbase v12 [40]), 31 miRNA loci that were homologous to known miRbase miRNAs from either mouse or human and 30 confident novel pre-miRNA loci (Additional files 3, 4, 5 and 6, **Files S1-4**). We assume that our global expression profiles allowed us to detect most miRNAs. Only eleven miRBase miRNAs (4%), i.e. rno-mir-7b, 147, 291a, 297, 327, 335, 347, 340, 349, 352 and 671, were not detected. This is likely due to the absence of the miRNA in the investigated tissues, or highly specific expression in a limited number of cells. Eight of the novel miRNAs reside in close proximity (< 100 kb) to known miRNAs in the rat. We obtained reads from both the 5' and 3' arm for 73.4% of the pre-miRNAs. Although the arm with the lowest read frequency is likely to represent the miRNA* sequence, it is also clear that the read frequency of small RNAs is subject to capture bias [41] and read frequencies per arm may be misleading. Therefore, we treated miRNAs derived from either hairpin arm as individual entities for further analysis steps. Seed (nt 1-7 or 2-8) overlaps with known miRBase miRNAs showed that the majority of miRNAs is conserved; 515 are conserved in rodents, whereas comparison to all categorized vertebrates shows that 610 seeds are conserved. Likely, the higher number of seed conservation in all vertebrates is due to the poor miRNA annotation in most rodents. Seventy-seven miRNAs (Additional file 7, **File S5**) contain seeds without known homologues in vertebrates, of which twenty-four are derived from the opposite strand of known miRNAs. These may represent rat-specific miRNAs, although it cannot be excluded that these remain to be detected in other species.

Tissue-specific miRNA profiles
miRNAs may act redundantly [42] or in concert [43,44]. In order to understand the function of miRNAs in specific tissues, it may therefore be helpful to identify global

miRNA profiles. To be able to compare miRNA profiles, we normalized the miRNA reads per tissue and determined the relative expression of miRNAs among the six tissues. Comparisons between strains showed highly similar expression profiles, with the Spearman's ρ ranging from 0.88 (testis) to 0.95 (brain). The ρ among different tissue profiles was considerably lower, as expected [45], ranging from 0.62 (brain vs. Testis) to 0.87 (*e.g.* liver vs. spleen) (Additional file 8, **Figure S2**). We determined the number of reads that covered individual known miRNAs (Figure 1A, B, C and Additional file 9, **Figure S3**) and observed that 38% of the known miRNAs are ubiquitously expressed among all tissues that were assayed, whereas only 14% of the homologous and 2% of the novel miRNAs were identified in all examined tissues.

To systematically detect miRNAs that are differentially expressed between tissues, we applied a strict (5-fold) expression level difference that had to apply for each of the two strains. By applying this conservative cut-off we excluded subtle expression differences and expected to only include miRNAs that exert tissue specific regulatory roles. By demanding that this threshold is met in both strains, it is guaranteed that these differences are not specific for a specific genetic background. At this threshold, we identified a set of 212 tissue-specific miRNAs. The number of tissue-specific miRNAs varies considerably between tissues, ranging from 9 in liver to 133 in brain (Figure 1D and Additional file 10, **table S2**). Notably, we find 35 tissue-specific miRNAs in testis, the tissue from which we obtained least miRNA reads. This excludes coverage as a confounding factor in this analysis. By applying a 5-fold lower expression threshold in a specific tissue, we identified miRNAs that specifically lower expressed (or avoidant) in a given tissue. Five miRNAs are specifically depleted in brain, liver or heart (Figure 1D). These miRNAs are medium to highly expressed in all other tissues analyzed (Additional file 9, **Figure S3**).

Some of these tissue-specific expression patterns are evolutionary conserved, for instance, the testis-specific *mir-202* was previously reported to be testis-specific in chicken as well [46]. Other specifically expressed miRNAs that we verified are *mir-208*, which is known to be restricted to the heart [47] and *mir-122*, the most prominent miRNA in liver [48]. Brain, the most complex tissue in our dataset, also contains most tissue-specific and avoidant miRNAs. Eighteen miRNAs that have no known homologs are expressed in a tissue-specific manner, of which 5 are present in testis (the least covered datasets), 1 in liver, 1 in spleen and 11 in brain. The enrichment for both testis and brain-specific miRNAs in this dataset partially excludes lack of sequencing depth as a confounding factor. These results suggest that the emergence of expression of novel miRNAs occurs in a tissue-specific fashion. Furthermore, these novel tissue-specific miR-

NAs are found primarily in the brain, confirming earlier observations that miRNAs might play an important role in the evolutionary plasticity of the brain [49].

Tissue-specific RNA editing

An auxiliary pathway to regulate miRNA expression and targeting is via the deamination of adenosines into inosines, a process known as RNA editing [50-52]. This process is mediated by adenosine deaminases acting on RNA (ADAR) [50]. During sequencing, inosines are read as guanines. To identify RNA editing we specifically identified A to G changes in otherwise perfectly mapped miRNA sequences and selected those miRNAs that are edited in at least 10% of the reads. After this selection we found that six miRNAs were subject to editing, including five positions in the miRNA seed region (Figure 2A). The most abundantly edited miRNA is *mir-376b*, in the brain (Figure 2B), fitting with the observation that in murine and human brains, *mir-376a/b/c* are targets of the RNA editing machinery [51,52]. We find the *mir-376a/c* isoforms to be edited as well, albeit to a lesser extent (Figure 2B). Importantly, in the case of *mir-376a-5p*, *mir-376b-5p*, *mir-376a-3p* and *mir-376-3p* the edited nucleoside resides in the seed region, which is instrumental in targeting the miRNA to its target and could therefore lead to an altered target spectrum.

In addition to known editing targets, such as *mir-376b*, we also identified the brain-specific *mir-377* that was hitherto unknown to be edited. To quantify the degree of editing we verified editing of *mir-377* by qRT-PCR analysis (Figure 2B, C, Additional file 11, **Table S3**). Based on dilution series we confirmed that the level of editing was ~37%, which is comparable to the level of editing as determined by DGE profiling.

Strain specificity in the liver

The two strains from which we prepared small RNA libraries are the founder strains of the BXH/HXB recombinant inbred panel [39]. This is a well-established panel of rat strain that has been used for defining expression quantitative trait loci (eQTLs). As an initial indexation, we determined the liver miRNAs that are differentially expressed between the two strains. To estimate the biological variation between individuals, we performed DGE profiling on liver samples of three BN-Lx and three SHR rats. By estimating the variance in similarly expressed miRNAs within one strain (see methods for details) we found four miRNAs that were differentially expressed between strains *(p < 0.05)* in the liver, *i.e.* two increased in BN-Lx and two increased in SHR liver (Figure 3A). The two miRNAs that were more prominent in BN-Lx, *mir-742-3p* and *rno-mir-741-3p*, reside in the same genomic cluster and may thus originate from the same primary transcript. We found that the *mir-293-5p* homologue,
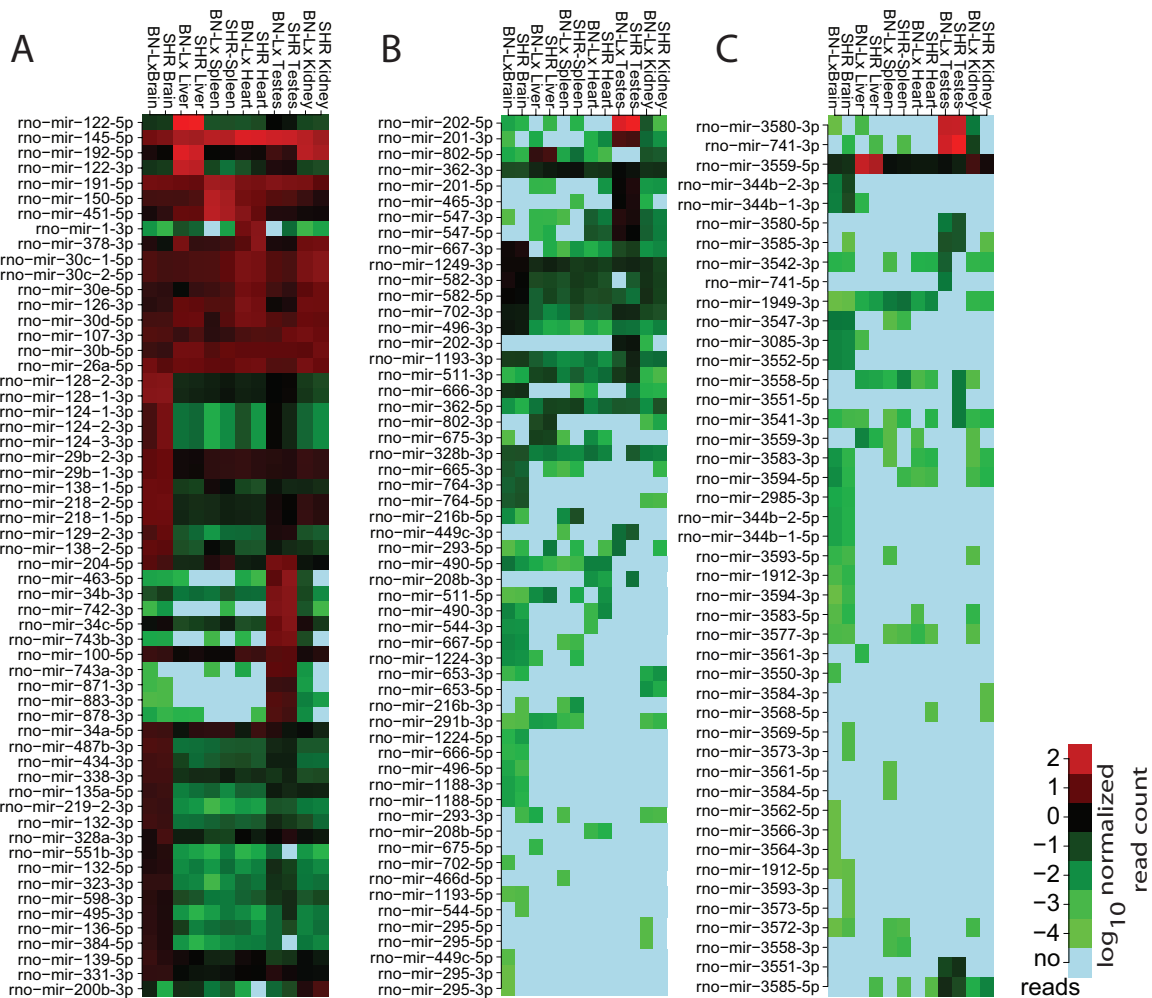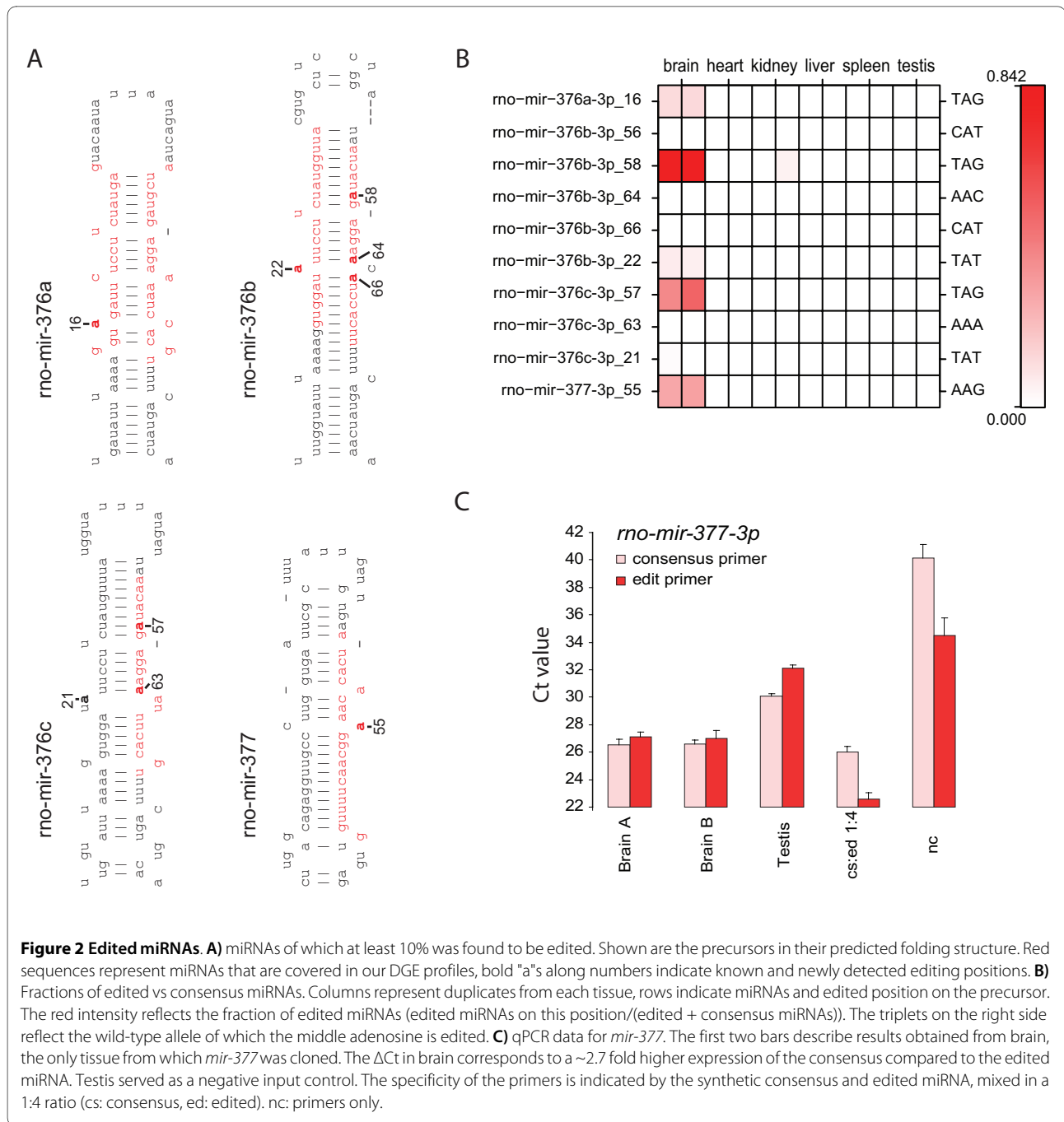
Figure 1 miRNA tissue expression profiles. **A-C)** Heatmaps show the expression (normalized, $\log_{10}$) of miRNAs. Each arm is represented separately, as are miRNAs in the reverse orientation (if applicable). In **(A)**, the first 10% of the heatmap is shown. miRNAs derived from loci that are homologous to miRBase human or mouse miRNA loci are shown in **(B)**. Completely novel miRNAs are shown in **(C)**. The legend is valid for the 3 heatmaps. **D)** Each tissue expresses a distinct set of miRNAs (tissue specific miRNAs). Also, some miRNAs are specifically not expressed in certain tissues (tissue avoidant miRNAs). miRNAs are specific in both replicates and the arbitrary, but strict, cutoff of 5-fold difference in expression was applied to determine specificity and avoidance.

which was not previously identified in the rat, was upregulated in SHR livers. Finally, we found that *mir-34a-5'*, which is known to be a transcriptional target of the pro-apoptotic p53 protein [53], was robustly increased in the SHR liver. We confirmed the differential expression of

this miRNA by qRT-PCR (Figure 3B). The observed ΔCt of ~1.5 approximates the observed difference in our sequencing datasets (Figure 3A, B).
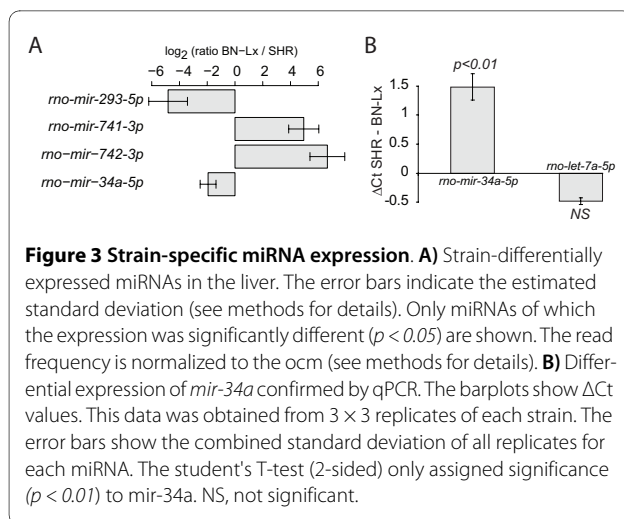
Expression of *rno-mir-741-3p* and *mir-742-3p*, which reside in the *mir-463* genomic cluster and is thus likely to

Figure 2 Edited miRNAs. **A)** miRNAs of which at least 10% was found to be edited. Shown are the precursors in their predicted folding structure. Red sequences represent miRNAs that are covered in our DGE profiles, bold "a"s along numbers indicate known and newly detected editing positions. **B)** Fractions of edited vs consensus miRNAs. Columns represent duplicates from each tissue, rows indicate miRNAs and edited position on the precursor. The red intensity reflects the fraction of edited miRNAs (edited miRNAs on this position/(edited + consensus miRNAs)). The triplets on the right side reflect the wild-type allele of which the middle adenosine is edited. **C)** qPCR data for *mir-377*. The first two bars describe results obtained from brain, the only tissue from which *mir-377* was cloned. The ΔCt in brain corresponds to a ~2.7 fold higher expression of the consensus compared to the edited miRNA. Testis served as a negative input control. The specificity of the primers is indicated by the synthetic consensus and edited miRNA, mixed in a 1:4 ratio (cs: consensus, ed: edited). nc: primers only.

be expressed as a single transcript [54], is strain-specific. This suggests that miRNA-specific post-transcriptional regulation could strongly affect miRNA expression. In order to investigate the conservation of this cluster, we aligned its pre-miRNAs to the genome of the mouse. The mouse is the best annotated close relative of the rat and shares common ancestry with the rat 12-24 Mya [55]. We find that the *mir-463* cluster is located in a syntentic region on chromosome X and is most prominently, but not uniquely, expressed in testis in both mouse and rat,

indicating conservation of location and expression (Additional file 12, **Figure S4A, B)** [34]. *rno-mir-741* is partially conserved; the syntenic mouse miRNA is *mmu-mir-741*. Notably the strongest diverged arm of *rno-mir-741* is the most abundant arm in our cloning data (Additional file 12, **Figure S4B)** and is the only miRNA arm that is described in *mmu-mir-741*, suggesting that expression is conserved while the sequence has diverged.

While most miRNA seeds on this cluster are strongly conserved between rat and mouse, divergence is an ongo-

Figure 3 Strain-specific miRNA expression. **A)** Strain-differentially expressed miRNAs in the liver. The error bars indicate the estimated standard deviation (see methods for details). Only miRNAs of which the expression was significantly different ($p < 0.05$) are shown. The read frequency is normalized to the ocm (see methods for details). **B)** Differential expression of *mir-34a* confirmed by qPCR. The barplots show ΔCt values. This data was obtained from 3×3 replicates of each strain. The error bars show the combined standard deviation of all replicates for each miRNA. The student's T-test (2-sided) only assigned significance ($p < 0.01$) to mir-34a. NS, not significant.

ing process that occasionally introduces novel functional miRNAs. Strain-specific processing of such miRNAs potentially precedes divergence: if by such regulation the miRNA expression is low, the miRNA itself can diverge without serious consequences. This may be one mechanism by which miRNAs evolve within this *mir-463* cluster, in which the *rno-mir-741/mmu-mir-741* ancestral miRNA has diverged between mouse and rat.

piRNAs

The majority of small RNAs in the testis that we identified in our experiments were either repeat-derived or lacked annotation. Most likely these small RNA reads represent germline-specific piwi-assiociated RNAs (piRNAs). Because piRNAs function in genome maintenance by silencing of transposon sequences [16,17,56], we first investigated the piRNAs that mapped to annotated repeats like long tandem repeats an other transposon-derived repeats. Repeat-derived piRNAs are thought to be generated via a "ping-pong" mechanism [17]: a piRNA hybridizes to a target, which is subsequently cleaved along position 10 of the piRNA. By further processing, a novel piRNA is generated from the original target. In vertebrates, the maturation of piRNAs does not depend on *Dicer*, but on *piwi* homologues [18,56,57]. They archetypically start with a 5' uracil and posess an adenine on position 10, which suggest that they are *Argonaute*-dependent cleavage products [17]. As a result, one of the characteristics of the model is a 10 nt 5' region that is complementary between two piRNAs. We therefore tested for the presence of this characteristic signature, by selecting all piRNAs that mapped to a specific type of transposon (*i.e.* L1_Rn and RMER3) and looked for presence and length of overlap which, as expected, peaks at 10 nt (Figure 4A, C). Consistent with this overlap, we find the characteristic nucleotide distribution, with piRNAs mapping to the sense strand having a 5' terminal uracil, and piRNAs

mapping to antisense strand having an adenine at 10th position (Figure 4B, D). This signature is characteristic for piRNAs that have been generated via the ping-pong mechanism mentioned above.

An additional class of piRNAs is the pachytene piRNAs, which are found in large strand-specific clusters throughout the genome [20]. By employing a cluster identification algorithm based on the density of small RNA sequences within a specific region (see Methods), we found 365 strand specific clusters (Additional file 13, **Figure S5**, Additional file 14, **Table S4**). piRNA clusters are found on all the autosomes but are excluded from the X chromosome. The genomic distribution shows that clusters on the negative and the positive strand are often found in close proximity. Close inspection reveals that these clusters are often adjacent and hardly overlapping. After merging adjacent clusters we obtained 321 unique locations in the genome with piRNA clusters. Whenever we find two large non-overlapping piRNA clusters on opposite strands in close proximity these are found almost exclusively upstream of each other, with transcription pointing outwards (Figure 4E), suggesting a nucleation mechanism on the border between these two clusters [18].

The identified piRNA clusters allow for probing the characteristics of these sequences compared to the rest of the genome (Figure 4F). Only 35 RefSeq curated genes (n = 12,692) overlap with piRNA clusters, and only 4 reside completely within a cluster. The gene products cover a broad range of functions, but no common function could be pinpointed (data not shown). Interestingly, the gene density in these clusters is more than three-fold increased compared to regions outside of piRNA cluster. As previously reported, we confirm that rat piRNA clusters are strongly depleted in repeat sequences compared to the rest of the genome [18,19,21]. Although a high level of synteny among piRNA clusters has been reported [19], piRNA clusters diverge at an elevated rate [58]. Indeed, we find that the amount of conserved sequences, defined as having a phastCons score > 0.5 in a multiple alignment of 8 vertebrate species [59], showed a depletion of conserved sequences in these regions. In conclusion, piRNA clusters are genomic regions with distinct sequence characteristics. Their location and expression is maintained while their sequence diverges quickly. This would favor the idea that they target the region on which they are encoded [18]. Alternatively, it may well be that they generate small RNAs which act sequence independently.

Conclusions

We performed small RNA DGE profiling for six rat tissues to expand the set of experimentally verified miRNAs and to gain biological insight into role of specific miRNAs. In total, we added 61 miRNA loci to the compen-
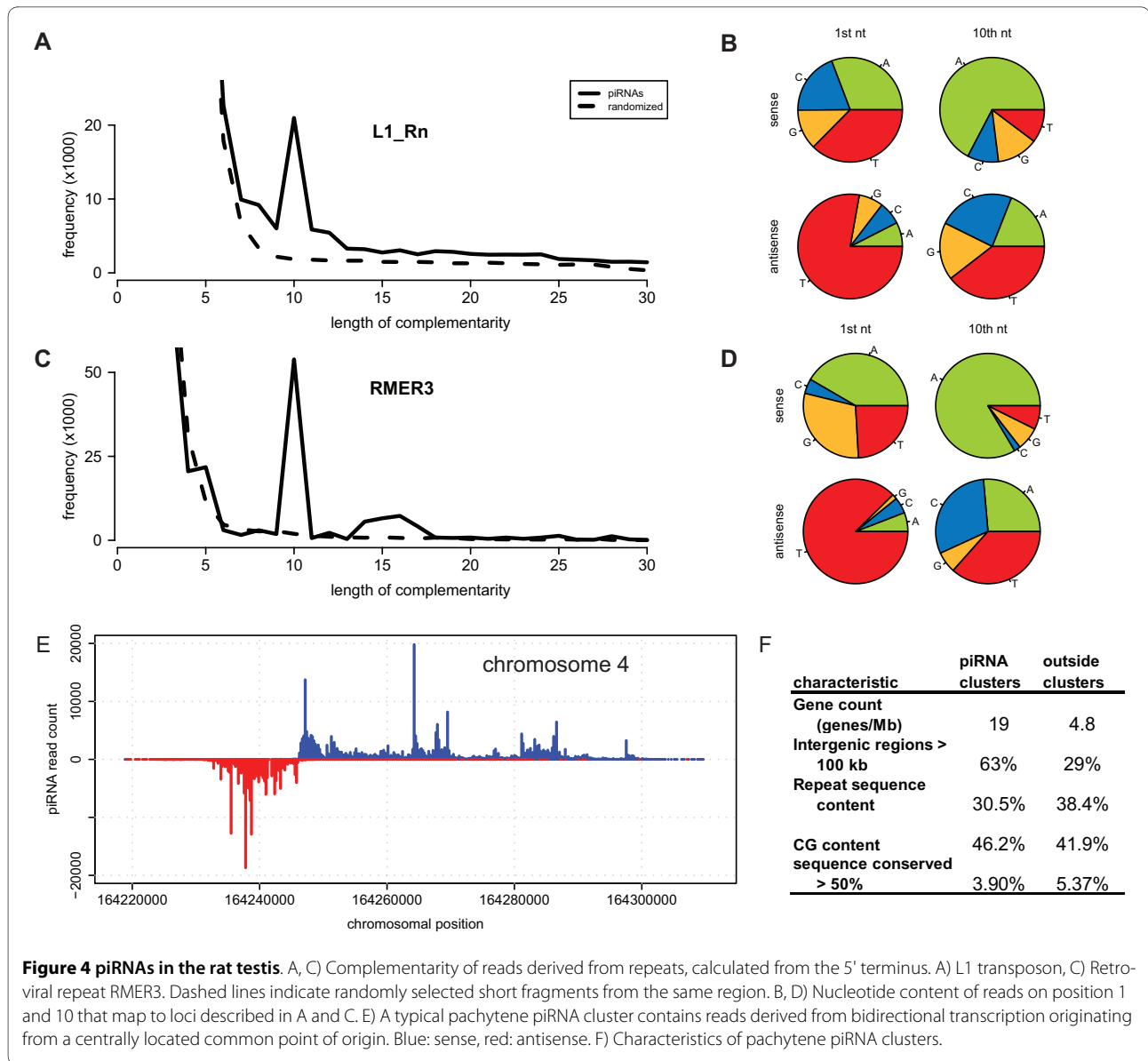
Figure 4 piRNAs in the rat testis. A, C) Complementarity of reads derived from repeats, calculated from the 5' terminus. A) L1 transposon, C) Retroviral repeat RMER3. Dashed lines indicate randomly selected short fragments from the same region. B, D) Nucleotide content of reads on position 1 and 10 that map to loci described in A and C. E) A typical pachytene piRNA cluster contains reads derived from bidirectional transcription originating from a centrally located common point of origin. Blue: sense, red: antisense. F) Characteristics of pachytene piRNA clusters.

dium of known rat miRNAs [40], providing experimental evidence for homologues of known mammalian miRNA as well as novel highly confident candidate miRNAs specific to the rat lineage. In addition to discovering novel miRNAs we quantified expression levels of miRNAs, creating a deep sequencing expression atlas of 6 tissues in two different rat inbred strains. This expression atlas can be an important instrument for understanding the evolution of miRNA regulation, *e.g.* conserved expression patterns among mammalian species, but also species-specific expression patterns. Our own experiments show strain-specific differences in miRNA expression and could form the basis for understanding this process at the molecular level.

Strain-specific miRNA expression patterns as determined by DGE can make important contributions to

understand regulatory expression networks. The role of miRNAs in complex traits has not been fully addressed yet. Such traits are not inherited in a monogenic mendelian fashion, but are expected to be caused or modulated by subtle differences in multiple different types of regulation, including miRNA-mediated regulation. Our evidence for strain-specific miRNA expression could imply that eQTLs, which have been observed at the level of expression of protein-coding genes, could be partially explained by miRNAs, especially when linked regions span large genomic distances. In such cases, these so-called *trans* eQTLs could potentially be explained by differential miRNA expression (caused by a *cis*-effect of a causal local variant) and the linked region(s) could thus be expected to be enriched for miRNA target genes for this specific miRNA. In conclusion, strain-specific

expression of miRNAs may help to identify and study these regulatory networks and facilitate target discovery for specific diseases.

In addition to transcriptional regulation, we confirmed the editing of known ADAR targets in rat, and discovered that, in addition to the known edited miRNAs, *mir-377* is edited in rat brains. Since the physiology and neurobiology of the rat has been intensively studied, rat models could be an important experimental system for addressing the biological relevance of miRNA editing. Finally we provide an exhaustive list of piRNA candidates.

The small RNA inventories and profiles described here further improve the annotation of the rat genome, thereby facilitating functional genomics research and the investigation of the involvement of miRNAs in complex genetic traits.

Methods
Animal care and treatment
All animal experiments have been carried out according to the Dutch legal ethical guidelines and were approved by the Animal Care Committee of the Royal Dutch Academy of Sciences.

MiRNA isolation and library preparation
Fresh tissues (i.e. brain, liver, spleen, heart, testis, kidney) from two rats (BN-Lx and SHR) and four additional liver samples (from BN-Lx and SHR) were directly frozen in liquid nitrogen and stored at -80°C until RNA isolation. Whole tissue was grained in liquid nitrogen and RNA isolation was performed by the Ambion Mirvana kit, according to the manufacturer's guidelines for total RNA isolation. After quality assessment on a 1% Agarose-TAE gel (data not shown), we prepared miRNA libraries by Ambion SREK kit according to the manufacturer's guidelines with the following specifications or deviations. We started with 2.5 µg total RNA from each sample and applied half of the suggested reagents for each step until PCR. We size fractionated total RNA in a 15-40 nt fraction on a 15% PAA denaturing (7 M Urea, TBE buffered) gel to collect small RNAs only. RNAs were eluted in 0.3 M NaCl fro 3 hrs at RT, precipitated at -80°C for 30 min in 3V EtOH, 10% 5 M ammonium acetate and 0.5 ul glycoblue (Ambion), spun down at 20 K G for 25 min at 4°C, washed in ice-cold 70% EtOH, spun down for 10 at 4°C and dissolved in nuclease-free H_2O. We ligated adapters from adapter mix A for 2 hrs at 16°C. For all but testis (from both rats) and liver (from the SHR rat) we applied 14 cycli of PCR. For testis samples we applied 16 cycles of PCR and for SHR liver we applied 17 cycli. Directly after PCR we excised the amplified libraries from a 3% agarose-TAE gel, without exposing the libraries to UV light. The libraries were purified by the GFX PCR DNA and Gel Band purification kit (GE life sciences), according to

the manufacturer's guidelines. The DNA size and quantity was measured on a Lonza 2.2% flash gel (Lonza) and Quant-iT dsDNA high sensitivity kit (Invitrogen), respectively (data not shown). Samples were mixed in to a 100 pg/µl solution. The emulsion PCR and Solid (AB) sequencing was performed as previously described [41]; the read length was set to 30 nt.

Source specifications
The rat genome version RGSC3.4 and the mousse genome NCBI build 37 were used for mapping. miRNA identities and sequences were obtained from miRBase V12 [40,60]. All statistics and visualization was performed with the R package [61] or Microsoft excel. Genome sequences were obtained from Ensemble release 55 [62] and alignments performed by blast [63]. We used RNAshapes [64] to predict secondary structures of RNA. Scripts were written in perl [65] and R.

Classifying small RNAs
All reads were submitted to the miRIntess small RNA analysis pipeline [38,49,66] (InteRNA Genomics B.V., Bilthoven, The Netherlands). The output of this analysis has been deposited on http://www.internagenomics.com/public/rat0812. This pipeline provided us with known miRNA reads and predicted miRNA reads with different levels of confidence. For instance, reads mapping to known miRNAs, but derived from the other arm, or from the opposite strand, were annotated with high confidence if they represented expected *Dicer* products and the pre-miRNA would form a stable hairpin. Homologous miRNAs were accordingly annotated. Similar rules applied to reads derived from yet unknown miRNAs, but the number of total reads should match at least 50 before these were assigned to be highly confident. Only low-confident novel miRNAs were excluded from the analysis, since predictions that fall into this category were indistinguishable from predictions performed on hairpins that had been randomly selected from genomic sequence (data not shown). Simultaneously the reads with a perfectly matching linker sequence were collected. The pre-miRNAs annotated by the mirIntess pipeline, except the category of low confident novel ones, were used to re-map the linker-trimmed reads, to obtain only perfectly mapping miRNA reads. Perl scripts are available upon request.

For the fine mapping of the small RNA reads we used the pre-miRNA hairpins that were identified by the miRIntess pipeline. Trimming of the sequencing linker was performed on the sequencing reads. Reads were subsequently remapped using SHRiMP [67], with the following settings (-h 50% -v 50% -s 1110111). Because this software package can natively map ABI colorspace reads, it has a higher sensitivity than the BLAST algorithm used in the miRIntess pipline. For all subsequent miRNA analy-

ses, except the editing analysis (see below), only perfectly matching reads were used.

Submission Gene expression omnibus, miRBase
Sequence reads have been submitted to GEO (accession number: GSE19054); microRNA sequences have been submitted to miRBase.

Normalizing miRNA coverage
The coverage of each miRNA arm derived from a known miRNA locus (sense or antisense) was counted for each tissue. In almost all tissues we observed an extremely high coverage of a selected number of miRNAs. For instance, reads from the *rno-mir-122-5p* contribute to more than 20% of all reads from known miRNAs from liver. Such outliers would have a substantial impact on the outcome of differential expression when the reads would be normalized to the mean. To reduce the impact of outliers to the correction parameter, we removed the three highest covered miRNAs from the calculation of the outlier-corrected mean (ocm). The read coverage of each miRNA (known and novel) was divided by this ocm. We chose not to take the median as the correction parameter because of the high number of lowly covered miRNAs in our dataset.

Editing analysis
To pick up edited miRNAs we scanned the mapped reads for A to G mutations. Because in ABI colorspace this requires two coherent colorbase changes, providing additional confidence that editing has indeed occurred in this miRNA.

miRNA profiles
For profiles, reads were normalized according to the ocm (see earlier in text). We calculated the Euclidean distance between each miRNA and built a tree by hierarchical clustering. The $\log_{10}$ of each read frequency was visualized in heatmaps, for known, novel but homologous to known, and novel miRNAs.

For the determination of tissue-specific miRNAs, reads were normalized as described and the normalized frequency in one tissue should be at least 5-fold higher than in any other tissue, for each strain individually, to be included in our set of tissue-specific miRNAs. Similarly, this frequency should be 5-fold lower in one tissue than in any other tissue, for each strain individually, for the miRNA to be included as a tissue-avoidant miRNA.

Differential miRNA expression in the liver
Read counts were normalized according to the ocm as described. We assigned significance $\alpha = 0.05$, because of the low number of replicates and the high expected variation. We took means of all replicates from each strain: mean(BN-Lx) = mB, mean (SHR) = mS. We calculated

average read frequency A = (mB + mS)/2 for both strains and relative expression values $D = \log_2(mB/mS)$. Subsequently, the data was ordered based on A and using a sliding window approach, the mean and standard deviation were calculated for windows of 41 miRNAs surrounding every miRNA (with 20 identical values at both extremes). Those means and standard deviations were estimations, which we applied to calculate the Z-score of each D. We obtained the P-value defined by the one-sided normal distribution ($P(X>Z)$) and corrected P-values by the false discovery rate [68].

qPCR analysis
The same RNA samples as those described in Methods section *MiRNA isolation and library preparation* were used. For the strain-specific miRNA expression, i.e. *mir-34a* and *let-7a*, we applied the miRNA taqman assay according to the manufacturer's guidelines, with 5 ng total RNA as input material. For *mir-377* we designed circular probes of which the termini were complementary to the miRNA, leaving a gap on the edited site (RT-mir377: TTGTGTGATTCAaaacaagagaagaagtagatgtactagtgctgcca caacattaatcaagaAAAAGTTGCC) We applied 11 μg total RNA (biological sample) or 1 pmol (synthetic miRNAs). RT was performed (total volume: 20 μl; 2 μl buffer, 6 μmol dNTPs each, 1 pmol RT primer, incubation at 72°C for 5 min, add 100 U RT enzyme (promega, MMLV-), 37°C 40 min). Samples were diluted 10 × in H_2O). PCR primers were diluted to 250 nM each (FWD-mir377-WT: CCCA-TGTTGAATCACACAAA, FWD-mir377-ED: CCCCAT-GTTGAATCACACAAG, REV-mir377: TTGTGTGATTCAaaacaagag,). Quantitative PCR was performed in 12 μl sybrgreen (Biorad IQ), 10 μl primer mix and 3 μl template. On a thermocylcer (Biorad MyIQ), we performed qPCR (1× 96°C for 02:00; 2× 96.0°C for 00:15, 55.0°C for 00:15, 72.0°C for 00:30; 60× 96.0°C for 00:15, 60.0°C for 00:15, 72.0°C for 00:30, 80.0°C for 00:05; 1× 80.0°C for 10:00).

mir-463 cluster comparison with mouse
Pre-miRNAs were aligned with the mouse genome by blastn (parameter deviating from default:-q -2). Loci with a minimal match length of 40 nt were investigated for a hairpin with approximately the same length of the original rat pre-miRNA. Only putative hairpins with at least 8 nt overlap with the cloned miRNA were included. Cloned miRNAs that aligned to the detected homologous locus had to match to at least one stem arm of the predicted pre-miRNA.

Definition of piRNAs and piRNA clusters
We mapped reads to annotated repeats (LTRs and LINE elements) and counted the amount of possible reverse complementary interactions, for lengths from 1 to 30 nt.

We also performed this analysis for a randomized set of 30-mers derived from the same repeat sequence.

After filtering out reads that are aligned to miRNAs, sn(o)RNAs, small cytoplasmic RNAs, rRNA genes or known coding genes, we found 271,382 unique start positions on the positive strand and 309,051 unique start positions on the negative strand. We have observed that distribution of these reads is non-random in the genome, *i.e.* clustered in certain locations. To rule out effects of clonality or cloning bias we only considered unique start positions. Regions that contain within a 2 kb window more than 20 unique start positions on the same strand are selected as the core of a piRNA cluster. From these piRNA core regions the start and end of the cluster is extended up- and downstream when a read is found on the same strand within 1 kb. This is continued until a read is found more than 1 kb from either start or end of the cluster.

Additional material

Additional file 1 Table S1. Read statistics.
Additional file 2 Figure S1. Small RNA read description.
Additional file 3 File S1. fastafile of all identified miRNAs in 6 tissues from the SHR and BN-Lx rat.
Additional file 4 File S2. GFF of all known pre-miRNAs.
Additional file 5 File S3. GFF of all homologous pre-miRNAs.
Additional file 6 File S4. GFF of all novel pre-miRNAs.
Additional file 7 File S5. fastafile of miRNAs without known seeds in vertebrates.
Additional file 8 Figure S2. correlation between 6 datasets.
Additional file 9 Figure S3. miRNAs derived from all known loci in the rat.
Additional file 10 Table S2. Optimization qPCR for *mir-377*.
Additional file 11 Table S3. Strain specific and avoidant miRNAs in six rat tissues.
Additional file 12 Figure S4. evolving miRNAs in the rat *mir-463* cluster.
Additional file 13 Figure S5. pachytene piRNA clusters on the rat genome.
Additional file 14 Table S4. piRNA clusters.

Authors' contributions

SL prepared the small RNA libraries, performed data analysis and wrote the manuscript. EB coordinated sequencing and handled the sequencer. EW performed data analysis and assisted in writing the manuscript. EC designed and coordinated the study. All authors read and approved the final manuscript.

Author Details

Hubrecht Institute-KNAW & University Medical Center Utrecht, Cancer Genomics Center, Utrecht, The Netherlands

References

1. Filipowicz W: RNAi: the nuts and bolts of the RISC machine. *Cell* 2005, **122**:17-20.
2. Lai EC: Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet* 2002, **30**:363-364.
3. Filipowicz W, Bhattacharyya SN, Sonenberg N: Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* 2008, **9**:102-114.
4. Gu S, Jin L, Zhang F, Sarnow P, Kay MA: Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. *Nat Struct Mol Biol* 2009, **16**:144-150.
5. Bartel DP: MicroRNAs: target recognition and regulatory functions. *Cell* 2009, **136**:215-233.
6. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB: Prediction of mammalian microRNA targets. *Cell* 2003, **115**:787-798.
7. Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS: MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell* 2009, **137**:647-658.
8. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G: The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* 2000, **403**:901-906.
9. Johnston RJ, Hobert O: A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. *Nature* 2003, **426**:845-849.
10. Karres JS, Hilgers V, Carrera I, Treisman J, Cohen SM: The conserved microRNA miR-8 tunes atrophin levels to prevent neurodegeneration in Drosophila. *Cell* 2007, **131**:136-145.
11. Xiao C, Calado DP, Galler G, Thai TH, Patterson HC, Wang J, Rajewsky N, Bender TP, Rajewsky K: MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb. *Cell* 2007, **131**:146-159.
12. 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.
13. 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.
14. Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP: The impact of microRNAs on protein output. *Nature* 2008, **455**:64-71.
15. Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N: Widespread changes in protein synthesis induced by microRNAs. *Nature* 2008, **455**:58-63.
16. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ: A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell* 2008, **31**:785-799.
17. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ: Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* 2007, **128**:1089-1103.
18. Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP, Kingston RE: Characterization of the piRNA complex from rat testes. *Science* 2006, **313**:363-367.
19. Girard A, Sachidanandam R, Hannon GJ, Carmell MA: A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* 2006, **442**:199-202.
20. Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ: Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* 2007, **316**:744-747.
21. Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, *et al.*: A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* 2006, **442**:203-207.
22. Stefani G, Slack FJ: Small non-coding RNAs in animal development. *Nat Rev Mol Cell Biol* 2008, **9**:219-230.
23. Yekta S, Shih IH, Bartel DP: MicroRNA-directed cleavage of HOXB8 mRNA. *Science* 2004, **304**:594-596.
24. Aravin AA, Hannon GJ, Brennecke J: The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 2007, **318**:761-764.
25. Jacob HJ: Functional genomics and rat models. *Genome Res* 1999, **9**:1013-1016.

26. Zan Y, Haag JD, Chen KS, Shepel LA, Wigington D, Wang YR, Hu R, Lopez-Guajardo CC, Brose HL, Porter KI, *et al.*: **Production of knockout rats using ENU mutagenesis and a yeast-based screening assay.** *Nat Biotechnol* 2003, **21**:645-651.

27. Smits BM, Mudde JB, van de BJ, Verheul M, Olivier J, Homberg J, Guryev V, Cools AR, Ellenbroek BA, Plasterk RH, *et al.*: **Generation of gene knockouts and mutant models in the laboratory rat by ENU-driven target-selected mutagenesis.** *Pharmacogenet Genomics* 2006, **16**:159-169.

28. van Boxtel R, Toonen PW, Verheul M, van Roekel HS, Nijman IJ, Guryev V, Cuppen E: **Improved generation of rat gene knockouts by target-selected mutagenesis in mismatch repair-deficient animals.** *BMC Genomics* 2008, **9**:460.

29. Ryu BY, Orwig KE, Oatley JM, Lin CC, Chang LJ, Avarbock MR, Brinster RL: **Efficient generation of transgenic rats through the male germline using lentiviral transduction and transplantation of spermatogonial stem cells.** *J Androl* 2007, **28**:353-360.

30. Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, Jenkins SS, Wood A, Cui X, Meng X, *et al.*: **Knockout rats via embryo microinjection of zinc-finger nucleases.** *Science* 2009, **325**:433.

31. Pravenec M, Simonet L, Kren V, Kunes J, Levan G, Szpirer J, Szpirer C, Kurtz T: **The rat renin gene: assignment to chromosome 13 and linkage to the regulation of blood pressure.** *Genomics* 1991, **9**:466-472.

32. Grieve IC, Dickens NJ, Pravenec M, Kren V, Hubner N, Cook SA, Aitman TJ, Petretto E, Mangion J: **Genome-wide co-expression analysis in multiple tissues.** *PLoS One* 2008, **3**:e4033.

33. Lazar J, Moreno C, Jacob HJ, Kwitek AE: **Impact of genomics on research in the rat.** *Genome Res* 2005, **15**:1717-1728.

34. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, *et al.*: **A mammalian microRNA expression atlas based on small RNA library sequencing.** *Cell* 2007, **129**:1401-1414.

35. Reddy AM, Zheng Y, Jagadeeswaran G, Macmil SL, Graham WB, Roe BA, Desilva U, Zhang W, Sunkar R: **Cloning, characterization and expression analysis of porcine microRNAs.** *BMC Genomics* 2009, **10**:65.

36. Murchison EP, Kheradpour P, Sachidanandam R, Smith C, Hodges E, Xuan Z, Kellis M, Grutzner F, Stark A, Hannon GJ: **Conservation of small RNA pathways in platypus.** *Genome Res* 2008, **18**:995-1004.

37. Wheeler BM, Heimberg AM, Moy VN, Sperling EA, Holstein TW, Heber S, Peterson KJ: **The deep evolution of metazoan microRNAs.** *Evol Dev* 2009, **11**:50-68.

38. de Wit E, Linsen SE, Cuppen E, Berezikov E: **Repertoire and evolution of miRNA genes in four divergent nematode species.** *Genome Res* 2009, **19**:2064-2074.

39. Pravenec M, Gauguier D, Schott JJ, Buard J, Kren V, Bila V, Szpirer C, Szpirer J, Wang JM, Huang H, *et al.*: **Mapping of quantitative trait loci for blood pressure and cardiac mass in the rat by genome scanning of recombinant inbred strains.** *J Clin Invest* 1995, **96**:1973-1978.

40. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ: **miRBase: tools for microRNA genomics.** *Nucleic Acids Res* 2008, **36**:D154-D158.

41. Linsen SE, de Wit E, Janssens G, Heater S, Chapman L, Parkin RK, Fritz B, Wyman SK, de Bruijn E, Voest EE, *et al.*: **Limitations and possibilities of small RNA digital gene expression profiling.** *Nat Methods* 2009, **6**:474-476.

42. Miska EA, Alvarez-Saavedra E, Abbott AL, Lau NC, Hellman AB, McGonagle SM, Bartel DP, Ambros VR, Horvitz HR: **Most Caenorhabditis elegans microRNAs are individually not essential for development or viability.** *PLoS Genet* 2007, **3**:e215.

43. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP: **MicroRNA targeting specificity in mammals: determinants beyond seed pairing.** *Mol Cell* 2007, **27**:91-105.

44. Saetrom P, Heale BS, Snove O Jr, Aagaard L, Alluin J, Rossi JJ: **Distance constraints between microRNA target sites dictate efficacy and cooperativity.** *Nucleic Acids Res* 2007, **35**:2333-2342.

45. 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.

46. Bannister SC, Tizard ML, Doran TJ, Sinclair AH, Smith CA: **Sexually dimorphic microRNA expression during chicken embryonic gonadal development.** *Biol Reprod* 2009, **81**:165-176.

47. van Rooij E, Sutherland LB, Qi X, Richardson JA, Hill J, Olson EN: **Control of stress-dependent cardiac growth and gene expression by a microRNA.** *Science* 2007, **316**:575-579.

48. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T: **Identification of tissue-specific microRNAs from mouse.** *Curr Biol* 2002, **12**:735-739.

49. 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.

50. 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.

51. Kawahara Y, Megraw M, Kreider E, Iizasa H, Valente L, Hatzigeorgiou AG, Nishikura K: **Frequency and fate of microRNA editing in human brain.** *Nucleic Acids Res* 2008, **36**:5270-5280.

52. Kawahara 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.

53. Tarasov V, Jung P, Verdoodt B, Lodygin D, Epanchintsev A, Menssen A, Meister G, Hermeking H: **Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest.** *Cell Cycle* 2007, **6**:1586-1593.

54. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**:281-297.

55. Gibbs RA, Weinstock GM, Metzker ML, Muzny DM, Sodergren EJ, Scherer S, Scott G, Steffen D, Worley KC, Burch PE, *et al.*: **Genome sequence of the Brown Norway rat yields insights into mammalian evolution.** *Nature* 2004, **428**:493-521.

56. Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den EH, Filippov DV, Blaser H, Raz E, Moens CB, *et al.*: **A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish.** *Cell* 2007, **129**:69-82.

57. Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD: **A distinct small RNA pathway silences selfish genetic elements in the germline.** *Science* 2006, **313**:320-324.

58. Assis R, Kondrashov AS: **Rapid repetitive element-mediated expansion of piRNA clusters in mammalian evolution.** *Proc Natl Acad Sci USA* 2009, **106**:7079-7082.

59. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, *et al.*: **Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes.** *Genome Res* 2005, **15**:1034-1050.

60. miRBAse: the microRNA database [http://www.mirbase.org/]

61. The R Project for Statistical Computing [http://www.r-project.org/]

62. Ensembl Genome Browser [http://www.ensembl.org/]

63. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403-410.

64. Steffen P, Voss B, Rehmsmeier M, Reeder J, Giegerich R: **RNAshapes: an integrated RNA analysis package based on abstract shapes.** *Bioinformatics* 2006, **22**:500-503.

65. Comprehensive Perl Archive Network [http://www.cpan.org/]

66. Interna Genomics [http://www.interna-genomics.com/]

67. Rumble SM, Lacroute P, Dalca AV, Fiume M, Sidow A, Brudno M: **SHRiMP: accurate mapping of short color-space reads.** *PLoS Comput Biol* 2009, **5**:e1000386.

68. Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing.** *Journal of the Royal Statistical Society B* 2005, **57**:289-300.