SomamiR: a database for somatic mutations impacting microRNA function in cancer

Anindya Bhattacharya[1,2], Jesse D. Ziebarth[1,2] and Yan Cui[1,2,*]

[1]Department of Microbiology, Immunology and Biochemistry and [2]Center for Integrative and Translational Genomics, The University of Tennessee Health Science Center, 858 Madison Avenue, Memphis, TN 38163, USA

ABSTRACT

Whole-genome sequencing of cancers has begun to identify thousands of somatic mutations that distinguish the genomes of normal tissues from cancers. While many germline mutations within microRNAs (miRNAs) and their targets have been shown to alter miRNA function in cancers and have been associated with cancer risk, the impact of somatic mutations on miRNA function has received relatively little attention. Here, we have created the SomamiR database (http://compbio.uthsc.edu/SomamiR/) to provide a comprehensive resource that integrates several types of data for use in investigating the impact of somatic and germline mutations on miRNA function in cancer. The database contains somatic mutations that may create or disrupt miRNA target sites and integrates these somatic mutations with germline mutations within the same target sites, genome-wide and candidate gene association studies of cancer and functional annotations that link genes containing mutations with cancer. Additionally, the database contains a collection of germline and somatic mutations in miRNAs and their targets that have been experimentally shown to impact miRNA function and have been associated with cancer.

INTRODUCTION

MicroRNAs (miRNAs) are small, non-coding RNAs that function as post-transcriptional regulators of mRNA expression. A single miRNA can target hundreds of mRNAs and a single mRNA can contain functional binding sites for several miRNAs, creating miRNA networks (1) that regulate the expression of a large portion of the human transcriptome (2). These networks have been shown to control many critical cellular processes that are altered in cancers, including differentiation, proliferation and apoptosis, and therefore, dysregulation of miRNA networks has been shown to be a key feature of many human cancers (1,3,4). Genetic variants within miRNAs or their target sites are one potential cause of the changes in miRNA function observed in cancers (5–8). For example, many patients with B-cell chronic lymphocytic leukemia have three large deletions that cause the loss of miR-15a and miR-16-1, preventing repression of the oncogene *BCL2*, and miR-34b and miR-34c, which are activated by *TP53* (9,10). Additionally, rs61764370, a single-nucleotide polymorphism that has been associated with lung, ovarian, oral and breast cancer, disrupts a binding site for let-7 in the 3′-untranslated region (3′-UTR) of *KRAS*, resulting in increased *KRAS* expression (11–14).

While germline mutations can impact cancer risk and treatment outcomes, understanding tumorigenesis often requires the identification of the somatic mutations that differentiate cancer from normal cells. Advances in sequencing technologies have enabled the identification of increasing numbers of somatic mutation through whole exome and, more recently, whole-genome sequencing (15). To date, most analysis of somatic mutations from these sequencing efforts has focused on the impact of somatic mutations in coding regions (16), neglecting mutations in the 3′-UTRs that are believed to contain the majority of functional miRNA target sites. However, one recent study investigated a somatic mutation in the 3′-UTR of *TNFAIP2* identified from sequencing the genome of a patient with acute myeloid leukemia and found that the mutation creates a new miRNA target site in *TNFAIP2*, a known target of the *PML-RARα* oncogene, reducing its expression (17).

To systematically investigate the impact of somatic mutations on miRNA dysregulation in cancer, we have created the SomamiR database (Figure 1). We collected somatic mutations identified by the sequencing of paired normal and cancer genomes that are located within miRNAs or their potential target sites. We then integrated these mutations with several other types of data, including the results of association studies, miRNA expression

*To whom correspondence should be addressed. Tel: +1 901 448 3240; Fax: +1 901 448 7360; Email: ycui2@uthsc.edu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.
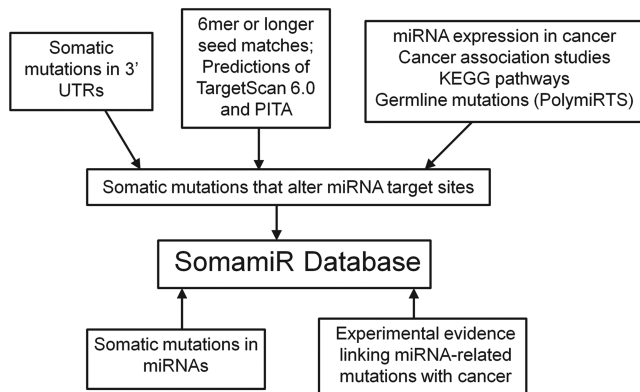
Figure 1. Overview of SomamiR. Somatic mutations that alter predicted miRNA target sites are determined and these mutations are integrated with several additional cancer data sets. The database also contains somatic mutations in miRNAs and summaries of studies that have investigated functional roles for miRNA-related germline and somatic mutations in cancer.

levels, gene functional annotations and germline mutations. Another key feature of SomamiR is a collection of mutations in miRNAs and miRNA target sites that have been experimentally shown to alter miRNA function and have also been linked with cancer through association studies or other methods.

DATABASE CONTENT

Somatic mutations within miRNA target sites

To determine somatic mutations that may alter miRNA target sites, we analyzed 15 sources of somatic mutations that have been identified from whole-genome sequencing of paired normal and cancer samples (Supplementary Table S1) and selected those somatic mutations located within the 3′-UTRs of RefSeq genes in the hg19 genome. We then assessed if the mutation alters putative miRNA target sites, focusing on mutations that alter sequence complementarity to miRNA seed regions for miRNAs included in miRBase release 17 (18–21). Specifically, we determined if a somatic mutation disrupts a seed match found in the normal genome, creates a new seed match in the cancer genome or modifies a seed match by, for example, changing an 8mer seed match to a 7mer seed match through disruption of the complementarity between of the final base of the target and the miRNA seed. As an analysis of experimentally identified miRNA target sites found that while predicted target sites containing longer (i.e. 7- and 8-mer) seed matches provide higher specificities, many experimental target sites contain only 6mer seed matches (22), we used three methods to predict how mutations may impact target sites. First, we created a comprehensive list of how somatic mutations may alter miRNA-binding sites by identifying 3′-UTR somatic mutations that create, disrupt or modify a 6mer or longer sequence complementary to a miRNA seed, using the 6 seed classes of miRNA seeds described by Ellwanger *et al.* (22). We also used two popular miRNA target prediction algorithms, TargetScan (23) and PITA (24), to determine mutations

that are more likely to alter functional binding sites. TargetScan 6.0 (23), which uses a more restrictive seed match criteria, provides a context+ score that considers how the site type, 3′-pairing, local AU sequence and position of the target site contribute to miRNA binding, while PITA (24) calculates a target score by comparing the energy required to unpair the target site and make it accessible for binding with the energy gained through binding the miRNA with the target. Both the TargetScan context+ score and the PITA score predict the binding of a miRNA to an entire 3′-UTR by summing over contributions made by individual sites within the 3′-UTR that are complementary to the miRNA seed region. The database currently contains 1886 somatic mutations in 3′-UTRs that alter miRNA target sites predicted by TargetScan 6.0 and 2255 somatic mutations that alter a 6mer or longer match to a miRNA seed.

To aid users in identifying somatic mutations likely to play a role in cancer, we have annotated miRNAs, genes and target locations with several types of information.

(1) Association studies results: we collected high-scoring markers from genome-wide association studies (GWAS) of cancer from the NHGRI GWAS Catalog (25) as well as meta-analysis of cancer candidate gene association studies (CGAS) from the Cancer GAMAdb (26). The database displays all markers associated with cancer from these sources for each gene that contains a somatic mutation altering a putative miRNA target site in its 3′-UTR.

(2) Gene pathways: functional annotation of genes containing somatic mutations that alter miRNA target sites was performed using the KEGG pathway database (27). Genes with somatic mutations in miRNA target sites are highlighted in each pathway.

(3) Sequence conservation: the conservation of a target site sequence across species has been used to improve miRNA target prediction. We used the 46-way MultiZ alignment of vertebrate genomes to determine vertebrate genomes in which the sequence of a predicted target site was conserved.

(4) Expression of miRNAs in cancer: the expression of miRNAs has been found to vary significantly across both normal and cancer tissues and therefore somatic mutations altering target sites of miRNAs that are highly expressed in a particular cancer type may be more likely to have a significant functional impact. We collected miRNA expression data from The Cancer Genome Atlas (TCGA) (http://cancergenome.nih.gov/), which contains read counts from RNA-Seq experiments measuring the expression of miRNAs in hundreds of cancer tissue samples for several cancer types (28,29). For the type of cancer in which a somatic mutation was identified, we selected a similar cancer type in TCGA, if available, and calculated the average miRNA expression read count in samples of that cancer in the TCGA database.

(5) Germline mutations: to identify germline mutations that could have a similar impact on miRNA targeting as the somatic mutations, we integrated somatic

mutations in the SomamiR database with data available in the PolymiRTS database (30,31), which contains germline mutations that alter predicted and experimental miRNA target sites. Specifically, we search PolymiRTS to collect germline mutations for each gene containing 3′-UTR somatic mutations that impact target sites of the same miRNAs whose target sites are impacted by somatic mutations. In some cases, the somatic and germline mutations are within a single predicted target site for a miRNA, but the somatic and germline mutations may alter distinct target sites within the same 3′-UTR.

Somatic mutations within miRNAs

To identify somatic mutations within miRNAs, we collected mutations from whole-exome sequences and other sources, including TCGA, that specifically reported somatic mutations within miRNAs, in addition to the whole-genome sequences used to identify somatic mutations in 3′-UTRs. We then determined if the somatic mutation was within the pre-miRNA sequence, the mature miRNA sequence or the seed region of the miRNA using miRBase 18. While mutations within pre-miRNAs have the potential to impact miRNA processing and the expression of miRNAs, mutations within mature miRNAs, and their seed regions in particular, may also rewire miRNA networks, redirecting a miRNA from its original targets to a new set of targets. In total, we identified 26 somatic mutations within miRNAs, including nine mutations within the miRNA seed region. For each miRNA containing a somatic mutation, we provide a link to the TarBase 6.0 (32), a database of experimentally identified miRNA–mRNA interactions, which displays experimentally supported targets of the miRNA.

Experimental evidence linking miRNA-related mutations with cancer

In recent years, over a hundred studies have investigated associations between cancer and germline mutations that may impact miRNA function (33). While the majority of these studies contain only association study results, there is a relatively small, but growing, number of cases where the functional impact of cancer-associated mutations on miRNA function has been extensively experimentally investigated. For example, rs2910164, a mutation in miR-146a, was shown to not only be associated with several cancers (34–37) but also increase both the expression of the miRNA and binding to *BRCA1* (34) and other target genes (37). Also, rs1434536, a mutation in strong linkage disequilibrium with high-scoring markers from breast cancer association studies, was shown to be located in an experimentally validated target site of miR-125b in *BMPR1B*, and the mutation was shown to disrupt the targeting of *BMPR1B* by miR-125b, increasing *BMPR1B* expression (38). We manually curated a collection of studies containing experimental evidence providing a functional role for miRNA-related mutations in cancer. In general, we used the following criteria to select mutations for this section of the database: (i) the mutation was

found to be associated with cancer risk or identified in cancer samples and (ii) the mutation was experimentally shown to affect miRNA expression or a specific mRNA–miRNA interaction. We have collected 16 and 27 cases in which mutations in miRNAs and their target sites, respectively, were shown to increase cancer risk and were experimentally shown to impact expression.

DATABASE FEATURES AND ACCESS

Presentation of somatic mutations in miRNA target sites

The majority of the SomamiR database is organized to present a single page for each gene that provides all somatic mutations that alter miRNA target sites in the gene as well as annotations linking the gene and mutations with cancer (Figure 2). These pages are divided into three main sections. The top section contains a description of the gene and a table showing how somatic mutations alter miRNA target sites. The middle section contains additional information that links the gene or mutation with cancer. This section includes, if applicable, germline mutations that impact target sites of the same miRNAs as the somatic mutations, association study results linking the gene with cancer risk, experimental results from investigations of the functional role of miRNA-related mutations in the gene and KEGG pathways that include the gene. The final section of the page is a Genome Browser window that includes custom tracks containing somatic and germline mutations and the putative miRNA-binding sites that these mutations alter. These tracks allow for quick visual identification of somatic and germline mutations that may impact the same miRNA target sites.

Browsing, searching and downloading SomamiR

Users can access the contents of SomamiR pages through several browsable tables that are linked from the database homepage. The browse options include separate tables showing somatic mutations in miRNA target sites, somatic mutations in miRNAs, experimental evidence linking mutations in miRNA target sites with cancer and experimental evidence linking mutations in miRNAs with cancer. Additionally, two separate tables can be used to browse database records in the context of gene pathways and association studies. One of these tables displays the number of genes in each KEGG pathway that contain somatic mutations in putative miRNA target sites. The KEGG pathways are separated into three sections: cancer pathways, signaling pathways and other pathways. Selecting any of the pathways provides a figure showing the pathway in which genes containing somatic mutations in miRNA target sites are highlighted. Second, a browsable table displays genes that contain 3′-UTR somatic mutations and also have been associated with cancer in a GWAS or CGAS. The records in this table can be filtered by the cancer type with which the gene was associated or by association study type (CGAS, GWAS or both).

SomamiR can be searched by chromosome location, miRNA, gene symbol and RefSeq ID through entering search terms directly on the website or uploading a file
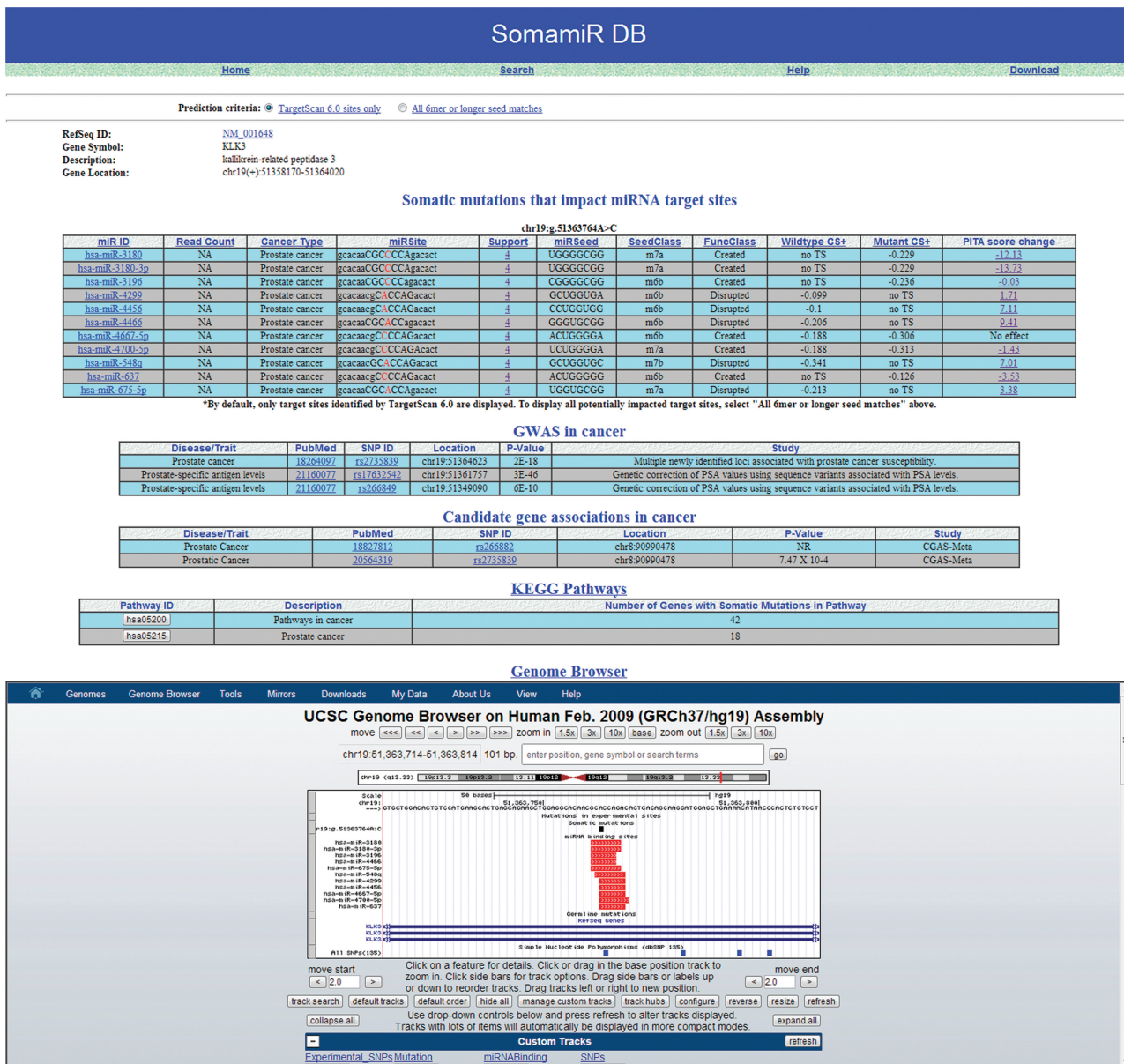
Figure 2. A screenshot of an example record in the SomamiR database. A somatic mutation in the 3′-UTR of *KLK3*, a gene in cancer pathways that has been associated with cancer in GWAS and CGAS, impacts predicted miRNA target sites.

containing search terms. The search results can also be filtered to select only somatic mutations that were identified in particular cancer types. Tab-delimited text files containing all of the data in the database are also available for download.

SUMMARY AND FUTURE DIRECTIONS

The SomamiR database provides a resource for the investigation of mutations that may affect miRNA function and play a role in cancer. The database integrates multiple types of data to facilitate the assessment of the functional significance of miRNA-related somatic mutations. For example, somatic mutations that are within sites targeted by miRNAs with high expression in cancer and

result in large changes in TargetScan context+ scores and/or PITA target scores are more likely to impact functional miRNA target sites. Additionally, cancer pathways and association study results can be used to identify mutation-harboring genes that are known to be involved in tumorigenesis. The analysis of cancer pathways in SomamiR could also potentially be used to identify particular branches of pathways that commonly contain miRNA-related somatic mutations. Integrating germline and somatic mutations may enable identification of functionally significant miRNAs and target sites that may contain germline mutations that are associated with cancer and somatic mutations that differentiate normal and cancer tissues in different populations or individuals. The Genome Browser tracks provided in SomamiR

provide a user-friendly way to identify target sites that contain both germline and somatic mutations. Finally, the collection of the results of experiments that have investigated the functional impact of miRNA-related mutations provided in the database.

In the future, we expect the utility of the SomamiR database to increase significantly as new data sets that expand knowledge of somatic mutations and miRNA function become available. Several large-scale projects, including TCGA, the International Cancer Genome Consortium (ICGC) (39) and the Pediatric Cancer Genome Project (PCGP) (40), are underway that will determine somatic mutations and miRNA expression levels in hundreds of cancer samples for dozens of cancer types. We have already included some of the initial publically available data from these projects, including somatic mutations and miRNA expression data. In addition to increasing the raw number of miRNA-related somatic mutations identified in cancer samples, these projects may provide somatic mutations that are found in genomes of multiple individuals with the same cancer type or multiple cancer types. These resources will also be greatly beneficial in future updates of the SomamiR database. While the majority of the current contents of the database were collected from supplementary information of individual papers, we will, in the future, focus on obtaining somatic mutations from the large data sets produced by these projects and from other collections of somatic mutations, such as COSMIC (41). We plan on collecting new somatic mutation to add to the database from these sources on a semi-annual basis and have developed a set of Python scripts that identify mutations in miRNAs and 3′-UTRs and predict how they impact miRNA targeting.

In addition to these large-scale projects studying somatic mutations and miRNAs in cancers, general understanding of miRNA function is likely to increase in the near future, and this understanding can be used to improve the SomamiR database. Recently, new experiment techniques, such as HITS-CLIP (42) and PAR-CLIP (43), have determined thousands of specific genomic sequences that are bound to miRNAs in mRNA–miRNA-Ago ternary complexes. The results of these and other lower-throughput experiments have greatly increased the number of experimentally identified miRNA target sites in the past few years, and this increase is likely to continue in the future (32). We used the results of PAR-CLIP experiments, as well as databases, such as TarBase (32), miRTarBase (44) and miRecords (45), that collect experimentally determined miRNA–mRNA interactions to determine if any somatic mutations are located within experimentally determined miRNA target sites, but found only a very small number of somatic mutations in these sites. In the future, as the number of experimental miRNA target sites and somatic mutations increases, we will expand the SomamiR database to highlight these somatic mutations that alter experimentally supported target sites. The growing number of experimentally supported miRNA target sites is also likely to improve the understanding of the specifics of miRNA target selection and binding, potentially improving miRNA target

prediction algorithms. Currently, the SomamiR database only examines how somatic mutations within 3′-UTRs alter sequences complementary to miRNA seeds. While we have adopted this approach because this class of mRNA–miRNA interactions has been the most highly studied and is believed to be a common factor to most mRNA–miRNA target pairs, it ignores that target sites may also interact with the 5′-end of miRNAs (46) and may be located with coding sequences (47) and the 5′-UTR of genes (48). As the understanding of miRNA target sites improves, it may soon become possible to more reliably predict how mutations impact miRNA binding, greatly improving the usefulness of the SomamiR database.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1.

REFERENCES

1. Volinia,S., Galasso,M., Costinean,S., Tagliavini,L., Gamberoni,G., Drusco,A., Marchesini,J., Mascellani,N., Sana,M.E., Abu Jarour,R. *et al.* (2010) Reprogramming of miRNA networks in cancer and leukemia. *Genome Res.*, **20**, 589–599.
2. Friedman,R.C., Farh,K.K., Burge,C.B. and Bartel,D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.*, **19**, 92–105.
3. Esquela-Kerscher,A. and Slack,F.J. (2006) Oncomirs - microRNAs with a role in cancer. *Nat. Rev. Cancer*, **6**, 259–269.
4. Medina,P.P. and Slack,F.J. (2008) microRNAs and cancer: an overview. *Cell Cycle*, **7**, 2485–2492.
5. Ryan,B.M., Robles,A.I. and Harris,C.C. (2010) Genetic variation in microRNA networks: the implications for cancer research. *Nat. Rev. Cancer*, **10**, 389–402.
6. Sethupathy,P. and Collins,F.S. (2008) MicroRNA target site polymorphisms and human disease. *Trends Genet.*, **24**, 489–497.
7. Bhattacharya,A., Ziebarth,J.D. and Cui,Y. (2012) Systematic analysis of microRNA targeting impacted by small insertions and deletions in human genome. *PLoS One*, **7**, e46176.
8. Mendell,J.T. and Olson,E.N. (2012) MicroRNAs in stress signaling and human disease. *Cell*, **148**, 1172–1187.
9. Fabbri,M., Bottoni,A., Shimizu,M., Spizzo,R., Nicoloso,M.S., Rossi,S., Barbarotto,E., Cimmino,A., Adair,B., Wojcik,S.E. *et al.* (2011) Association of a microRNA/TP53 feedback circuitry with pathogenesis and outcome of B-cell chronic lymphocytic leukemia. *JAMA*, **305**, 59–67.
10. Allendorf,D.J. and Davis,R.S. (2011) Unraveling the molecular pathogenesis of chronic lymphocytic leukemia: dissecting a microRNA regulatory network. *JAMA*, **305**, 95–97.
11. Chin,L.J., Ratner,E., Leng,S., Zhai,R., Nallur,S., Babar,I., Muller,R.U., Straka,E., Su,L., Burki,E.A. *et al.* (2008) A SNP in a let-7 microRNA complementary site in the KRAS 3′-untranslated region increases non-small cell lung cancer risk. *Cancer Res.*, **68**, 8535–8540.
12. Ratner,E., Lu,L., Boeke,M., Barnett,R., Nallur,S., Chin,L.J., Pelletier,C., Blitzblau,R., Tassi,R., Paranjape,T. *et al.* (2010) A KRAS-variant in ovarian cancer acts as a genetic marker of cancer risk. *Cancer Res.*, **70**, 6509–6515.

13. Christensen,B.C., Moyer,B.J., Avissar,M., Ouellet,L.G., Plaza,S.L., McClean,M.D., Marsit,C.J. and Kelsey,K.T. (2009) A let-7 microRNA-binding site polymorphism in the KRAS 3′-UTR is associated with reduced survival in oral cancers. *Carcinogenesis*, **30**, 1003–1007.

14. Paranjape,T., Heneghan,H., Lindner,R., Keane,F.K., Hoffman,A., Hollestelle,A., Dorairaj,J., Geyda,K., Pelletier,C., Nallur,S. *et al.* (2011) A 3′-untranslated region KRAS variant and triple-negative breast cancer: a case-control and genetic analysis. *Lancet Oncol.*, **12**, 377–386.

15. Meyerson,M., Gabriel,S. and Getz,G. (2010) Advances in understanding cancer genomes through second-generation sequencing. *Nat. Rev. Genet.*, **11**, 685–696.

16. Berger,M.F., Lawrence,M.S., Demichelis,F., Drier,Y., Cibulskis,K., Sivachenko,A.Y., Sboner,A., Esguerra,R., Pflueger,D., Sougnez,C. *et al.* (2011) The genomic complexity of primary human prostate cancer. *Nature*, **470**, 214–220.

17. Ramsingh,G., Koboldt,D.C., Trissal,M., Chiappinelli,K.B., Wylie,T., Koul,S., Chang,L.W., Nagarajan,R., Fehniger,T.A., Goodfellow,P. *et al.* (2010) Complete characterization of the microRNAome in a patient with acute myeloid leukemia. *Blood*, **116**, 5316–5326.

18. Griffiths-Jones,S. (2004) The microRNA registry. *Nucleic Acids Res.*, **32**, D109–D111.

19. Griffiths-Jones,S., Grocock,R.J., van Dongen,S., Bateman,A. and Enright,A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.*, **34**, D140–D144.

20. Griffiths-Jones,S., Saini,H.K., van Dongen,S. and Enright,A.J. (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res.*, **36**, D154–D158.

21. Kozomara,A. and Griffiths-Jones,S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.*, **39**, D152–D157.

22. Ellwanger,D.C., Buttner,F.A., Mewes,H.W. and Stumpflen,V. (2011) The sufficient minimal set of miRNA seed types. *Bioinformatics*, **27**, 1346–1350.

23. Garcia,D.M., Baek,D., Shin,C., Bell,G.W., Grimson,A. and Bartel,D.P. (2011) Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat. Struct. Mol. Biol.*, **18**, 1139–1146.

24. Kertesz,M., Iovino,N., Unnerstall,U., Gaul,U. and Segal,E. (2007) The role of site accessibility in microRNA target recognition. *Nat. Genet.*, **39**, 1278–1284.

25. Hindorff,L.A., Sethupathy,P., Junkins,H.A., Ramos,E.M., Mehta,J.P., Collins,F.S. and Manolio,T.A. (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl Acad. Sci. USA*, **106**, 9362–9367.

26. Schully,S.D., Yu,W., McCallum,V., Benedicto,C.B., Dong,L.M., Wulf,A., Clyne,M. and Khoury,M.J. (2011) Cancer GAMAdb: database of cancer genetic associations from meta-analyses and genome-wide association studies. *Eur. J. Hum. Genet.*, **19**, 928–930.

27. Kanehisa,M., Goto,S., Sato,Y., Furumichi,M. and Tanabe,M. (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.*, **40**, D109–D114.

28. Cancer Genome Atlas Research Network. (2011) Integrated genomic analyses of ovarian carcinoma. *Nature*, **474**, 609–615.

29. Cancer Genome Atlas Research Network. (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, **487**, 330–337.

30. Bao,L., Zhou,M., Wu,L., Lu,L., Goldowitz,D., Williams,R.W. and Cui,Y. (2007) PolymiRTS database: linking polymorphisms in microRNA target sites with complex traits. *Nucleic Acids Res.*, **35**, D51–D54.

31. Ziebarth,J.D., Bhattacharya,A., Chen,A. and Cui,Y. (2012) PolymiRTS database 2.0: linking polymorphisms in microRNA target sites with human diseases and complex traits. *Nucleic Acids Res.*, **40**, D216–D221.

32. Vergoulis,T., Vlachos,I.S., Alexiou,P., Georgakilas,G., Maragkakis,M., Reczko,M., Gerangelos,S., Koziris,N., Dalamagas,T. and Hatzigeorgiou,A.G. (2012) TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. *Nucleic Acids Res.*, **40**, D222–D229.

33. Bruno,A.E., Li,L., Kalabus,J.L., Pan,Y., Yu,A. and Hu,Z. (2012) miRdSNP: a database of disease-associated SNPs and microRNA target sites on 3′-UTRs of human genes. *BMC Genomics*, **13**, 44.

34. Shen,J., Ambrosone,C.B., DiCioccio,R.A., Odunsi,K., Lele,S.B. and Zhao,H. (2008) A functional polymorphism in the miR-146a gene and age of familial breast/ovarian cancer diagnosis. *Carcinogenesis*, **29**, 1963–1966.

35. Xu,B., Feng,N.H., Li,P.C., Tao,J., Wu,D., Zhang,Z.D., Tong,N., Wang,J.F., Song,N.H., Zhang,W. *et al.* (2010) A functional polymorphism in pre-miR-146a gene is associated with prostate cancer risk and mature miR-146a expression in vivo. *Prostate*, **70**, 467–472.

36. Yue,C., Wang,M., Ding,B., Wang,W., Fu,S., Zhou,D., Zhang,Z. and Han,S. (2011) Polymorphism of the pre-miR-146a is associated with risk of cervical cancer in a Chinese population. *Gynecol. Oncol.*, **122**, 33–37.

37. Jazdzewski,K., Murray,E.L., Franssila,K., Jarzab,B., Schoenberg,D.R. and de la Chapelle,A. (2008) Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. *Proc. Natl Acad. Sci. USA*, **105**, 7269–7274.

38. Saetrom,P., Biesinger,J., Li,S.M., Smith,D., Thomas,L.F., Majzoub,K., Rivas,G.E., Alluin,J., Rossi,J.J., Krontiris,T.G. *et al.* (2009) A risk variant in an miR-125b binding site in BMPR1B is associated with breast cancer pathogenesis. *Cancer Res.*, **69**, 7459–7465.

39. Hudson,T.J., Anderson,W., Artez,A., Barker,A.D., Bell,C., Bernabe,R.R., Bhan,M.K., Calvo,F., Eerola,I., Gerhard,D.S. *et al.* (2010) International network of cancer genome projects. *Nature*, **464**, 993–998.

40. Downing,J.R., Wilson,R.K., Zhang,J., Mardis,E.R., Pui,C.H., Ding,L., Ley,T.J. and Evans,W.E. (2012) The pediatric cancer genome project. *Nat. Genet.*, **44**, 619–622.

41. Forbes,S.A., Bindal,N., Bamford,S., Cole,C., Kok,C.Y., Beare,D., Jia,M., Shepherd,R., Leung,K., Menzies,A. *et al.* (2011) COSMIC: mining complete cancer genomes in the catalogue of somatic mutations in cancer. *Nucleic Acids Res.*, **39**, D945–D950.

42. Chi,S.W., Zang,J.B., Mele,A. and Darnell,R.B. (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*, **460**, 479–486.

43. Hafner,M., Landthaler,M., Burger,L., Khorshid,M., Hausser,J., Berninger,P., Rothballer,A., Ascano,M. Jr, Jungkamp,A.C., Munschauer,M. *et al.* (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*, **141**, 129–141.

44. Hsu,S.D., Lin,F.M., Wu,W.Y., Liang,C., Huang,W.C., Chan,W.L., Tsai,W.T., Chen,G.Z., Lee,C.J., Chiu,C.M. *et al.* (2011) miRTarBase: a database curates experimentally validated microRNA-target interactions. *Nucleic Acids Res.*, **39**, D163–D169.

45. Xiao,F., Zuo,Z., Cai,G., Kang,S., Gao,X. and Li,T. (2009) miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res.*, **37**, D105–D110.

46. Brennecke,J., Stark,A., Russell,R.B. and Cohen,S.M. (2005) Principles of microRNA-target recognition. *PLoS Biol.*, **3**, e85.

47. Fang,Z. and Rajewsky,N. (2011) The impact of miRNA target sites in coding sequences and in 3′-UTRs. *PLoS One*, **6**, e18067.

48. Lee,I., Ajay,S.S., Yook,J.I., Kim,H.S., Hong,S.H., Kim,N.H., Dhanasekaran,S.M., Chinnaiyan,A.M. and Athey,B.D. (2009) New class of microRNA targets containing simultaneous 5′-UTR and 3′-UTR interaction sites. *Genome Res.*, **19**, 1175–1183.