

Article

oncoNcRNA: A Web Portal for Exploring the Non-Coding RNAs with Oncogenic Potentials in Human Cancers

Ze-Lin Wang ^{1,2}, Xiao-Qin Zhang ³, Hui Zhou ^{1,2}, Jian-Hua Yang ^{1,2,*} and Liang-Hu Qu ^{1,2,*}

¹ Key Laboratory of Gene Engineering of the Ministry of Education, Guangzhou 510275, China; wangzl6@mail2.sysu.edu.cn

² State Key Laboratory for Biocontrol, Sun Yat-sen University, Guangzhou 510275, China; wangzl6@mail2.sysu.edu.cn

³ School of medicine, South China University of Technology, Guangzhou 510640, China; zhxq@connect.hku.hk

* Correspondence: lssqlh@mail.sysu.edu.cn (L.-H.Q.); yangjh7@mail.sysu.edu.cn (J.-H.Y.);
Tel.: +86-20-84112399 (J.-H.Y.)

Academic Editor: George A. Calin

Received: 30 November 2016; Accepted: 20 January 2017; Published: 7 February 2017

Abstract: Non-coding RNAs (ncRNAs) have been shown to contribute to tumorigenesis and progression. However, the functions of the majority of ncRNAs remain unclear. Through integrating published large-scale somatic copy number alterations (SCNAs) data from various human cancer types, we have developed oncoNcRNA, a user-friendly web portal to explore ncRNAs with oncogenic potential in human cancers. The portal characterizes the SCNAs of over 58,000 long non-coding RNAs (lncRNAs), 34,000 piwi-interacting RNAs (piRNAs), 2700 microRNAs (miRNAs), 600 transfer RNAs (tRNAs) and 400 small nucleolar RNAs (snoRNAs) in 64 human cancer types. It enables researchers to rapidly and intuitively analyze the oncogenic potential of ncRNAs of interest. Indeed, we have discovered a large number of ncRNAs which are frequently amplified or deleted within and across tumor types. Moreover, we built a web-based tool, Correlations, to explore the relationships between gene expression and copy number from ~10,000 tumor samples in 36 cancer types identified by The Cancer Genome Atlas (TCGA). oncoNcRNA is a valuable tool for investigating the function and clinical relevance of ncRNAs in human cancers. oncoNcRNA is freely available at <http://rna.sysu.edu.cn/onconcrna/>.

Keywords: ncRNAs; cancer; somatic copy number alteration; biomarker

1. Introduction

It has been clear that the majority of the human (or mammalian) genome is transcribed as non-coding RNAs (ncRNAs) such as well-studied miRNAs and long non-coding RNAs (lncRNAs) [1–9]. The most recent study has identified more than 58,000 lncRNA genes in the human genome [10]. However, only a small portion of lncRNAs have been characterized. Similarly, for other ncRNA classes like small nucleolar RNA (snoRNA), piwi-interacting RNA (piRNA) and transfer RNA (tRNA), whether they can contribute to carcinogenesis also remains unclear.

Somatic copy number alterations (SCNAs) are extremely common in cancer [11,12], which can promote tumorigenesis by altering the expression of oncogenic genes [13]. Recent advances in genome characterization technologies have increasingly enabled the systematic characterization of these alterations in human cancers [14,15]. However, a large number of these identified recurrent somatic copy number alterations (SCNAs) regions usually have no known oncogene or tumor suppressor gene targets (protein-coding genes). This leads researchers to focus their attentions on ncRNAs. Indeed, several studies have identified oncogenic ncRNAs from these recurrent SCNA regions. For example,

lncRNA FAL1 is amplified in ovarian cancer, and its small interfering RNAs (siRNAs) can significantly inhibit tumor growth in vivo [16]; the SNORD50A/B snoRNA locus is deleted at a frequency of >10% in each of 12 common cancers, where its loss is associated with reduced survival [17]. This suggests that SCNA is an important feature for estimating the function of ncRNAs.

With the initiation of large-scale cancer genomic projects like The Cancer Genome Atlas (TCGA) [18], the International Cancer Genome Consortium (ICGC) cancer genome projects [19] and the Cancer Cell Line Encyclopedia (CCLE) projects [20], a mass of oncogenomic data has been produced. A number of tools have been created to facilitate the access to multidimensional oncogenomic data and assist with the interpretation of the data, such as Catalogue Of Somatic Mutations In Cancer (COSMIC) [21] and cBioPortal [22,23]. However, these tools mainly focus their attention on protein-coding genes instead of ncRNAs. Although some databases/platforms have also been developed to study the function of ncRNAs [24] such as starBase2.0 [25], lncRNADB [26], LNCipedia [27] and deepBase2.0 [28] (Supplementary Table 1), they usually ignore the SCNA information of ncRNAs. Thus, a comprehensive catalog for SCNAs of ncRNAs in human cancers is urgently needed.

In this study, we developed oncoNcRNA to explore the oncogenic potential of ncRNAs from large-scale SCNA data in human cancers. In oncoNcRNA, we performed a large-scale integration of public SCNA data and provided the SCNA information of all well-annotated ncRNAs in up to 64 human cancer types (Figure 1).

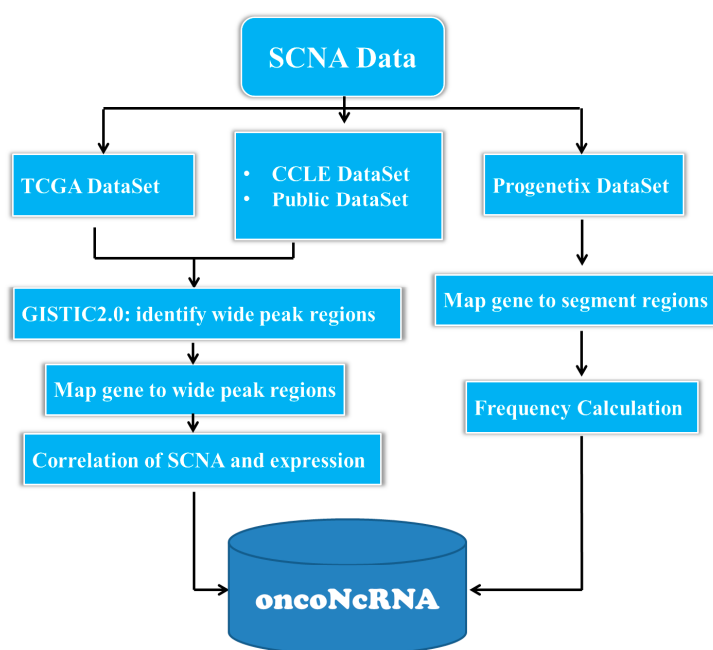


Figure 1. The workflow of oncoNcRNA web portal. We integrated large-scale somatic copy number alteration (SCNA) data from The Cancer Genome Atlas (TCGA), Cancer Cell Line Encyclopedia (CCLE) and other published papers, and provided the correlation of expression level and copy number for over 50,000 genes in 36 TCGA cancer types. Recurrent SCNA regions were identified by GISTIC 2.0 software [29] (Genomic Identification of Significant Targets in Cancer). The oncoNcRNA provided user-friendly web interfaces to query and browse the SCNA information of all well-annotated ncRNAs in up to 64 human cancer types.

Moreover, we also estimated the correlation of expression level and copy number for over 50,000 genes in 36 TCGA cancer types. oncoNcRNA provides user-friendly web interfaces to query and browse the SCNA information of the gene of interest in human cancers. It is expected to help researchers to investigate the potential functions and mechanisms of ncRNAs.

2. Results

2.1. Identification of ncRNAs with Recurrent SCNAs in Human Cancers

To help researchers rapidly and conveniently estimate whether a ncRNA has oncogenic roles, we systematically estimated the SCNAs of all well-annotated ncRNAs in 64 human cancers from TCGA [18], Progenetix database [30] and several independent studies [15,20,31]. As a result, we discovered that a larger number of ncRNAs show extensive copy number gain/loss within and across cancer types. For example, for 13,870 lncRNAs annotated in GENCODE version 19, 72.1% (10,001/13,870) of lncRNAs showed significant SCNAs in at least one cancer type for all 37 TCGA cancer types; for 2794 miRNA annotated in miRBase version 20, up to 99.1% (2768/2794) of miRNAs showed significant SCNAs in TCGA tumor cohorts. This finding reconfirms that miRNAs frequently show genomic alterations in human cancers [32,33]. Notably, the majority of ncRNAs possessed SCNAs in at least two cancer types, whereas only a small portion of ncRNAs presented cancer-unique alterations. Figure 2 characterizes the signatures consisting of the top 50 lncRNAs and 50 miRNAs with the most cancer types for amplification. For 50 lncRNAs, cancer susceptibility candidate 11 (CASC11) is ranked the first, while two canonical oncogenic lncRNA, CCAT1 [34–38] and PVT1 [39–41], are ranked 9th and 12th respectively. These molecules may be worthy of further experimental investigations.

2.2. Estimation of Correlation of ncRNA Gene Copy Number and Expression

Expression alteration is one of the important features to estimate the oncogenic potential of a ncRNA. We speculated that the ncRNA whose expression has a significantly positive correlation with SCNA is more likely to be oncogenic. Therefore, we integrated the available expression values from 36 TCGA cancer types to estimate the 54,312 genes including 1227 miRNAs. This analysis revealed that 54.4% of genes showed significant correlations between expression and copy numbers (Pearson $r > 0$ and $p < 0.05$). The top five genes were protein-coding genes with copy number loss in more than 23 cancer types, including canonical tumor suppressor genes *RB1* [42] and *PTEN* [43]. The top non-coding RNA was found to be a pseudogene named as RPL23AP82, which was deleted in 23 cancer types. Some well-known oncogenic ncRNAs were also included in our results. For example, lncRNA PVT1 [39–41] showed significant amplification and overexpression in 12 cancer types. Two of the most significant cancers for lncRNA PVT1 are presented in Figure 3. These data provide useful information for ncRNA studies and clinical biomarker discovery.

2.3. The Web-Based Exploration for the Oncogenic Potential of ncRNAs

To help researchers conveniently obtain our analysis results and inquire after a gene of interest, we have developed a user-friendly web portal, oncoNcRNA, for querying and browsing SCNA information for a specific gene. The web portal includes six main modules: lncRNAs, miRNAs, OtherNcRNAs, Proteins, Correlations and Download. The “Download” module allows users to obtain all our analysis results for further investigation.

According to different gene types, we provide four modules, including lncRNAs, miRNAs, OtherNcRNAs and Proteins, for querying SCNA information about genes in human cancers. With these modules, users can rapidly learn about whether the gene of interest shows recurrent SCNA in a specific cancer type and whether the alteration is extensive across tumor types. Given lncRNA PVT1 for an example, users can rapidly find that PVT1 is frequently amplified in 15 cancer types when selecting the TCGA data cohort through the “lncRNA” module (Figure 4A). When users continue to select other data cohorts, such as the progenetix dataset, PVT1 shows amplification in more than 10% of samples within individual tumors across 33 cancer types (Figure 4B). Users can further obtain the detailed information on genes about cancer types, amplification or deletion, GISTIC 2.0 [29] q value or frequency, residual q value or sample numbers that matched the frequency in result tables when clicking on the gene. The information is also visualized when clicking on the “Frequency” sub-module, and the figure can be exported in multiple formats including PDF file.

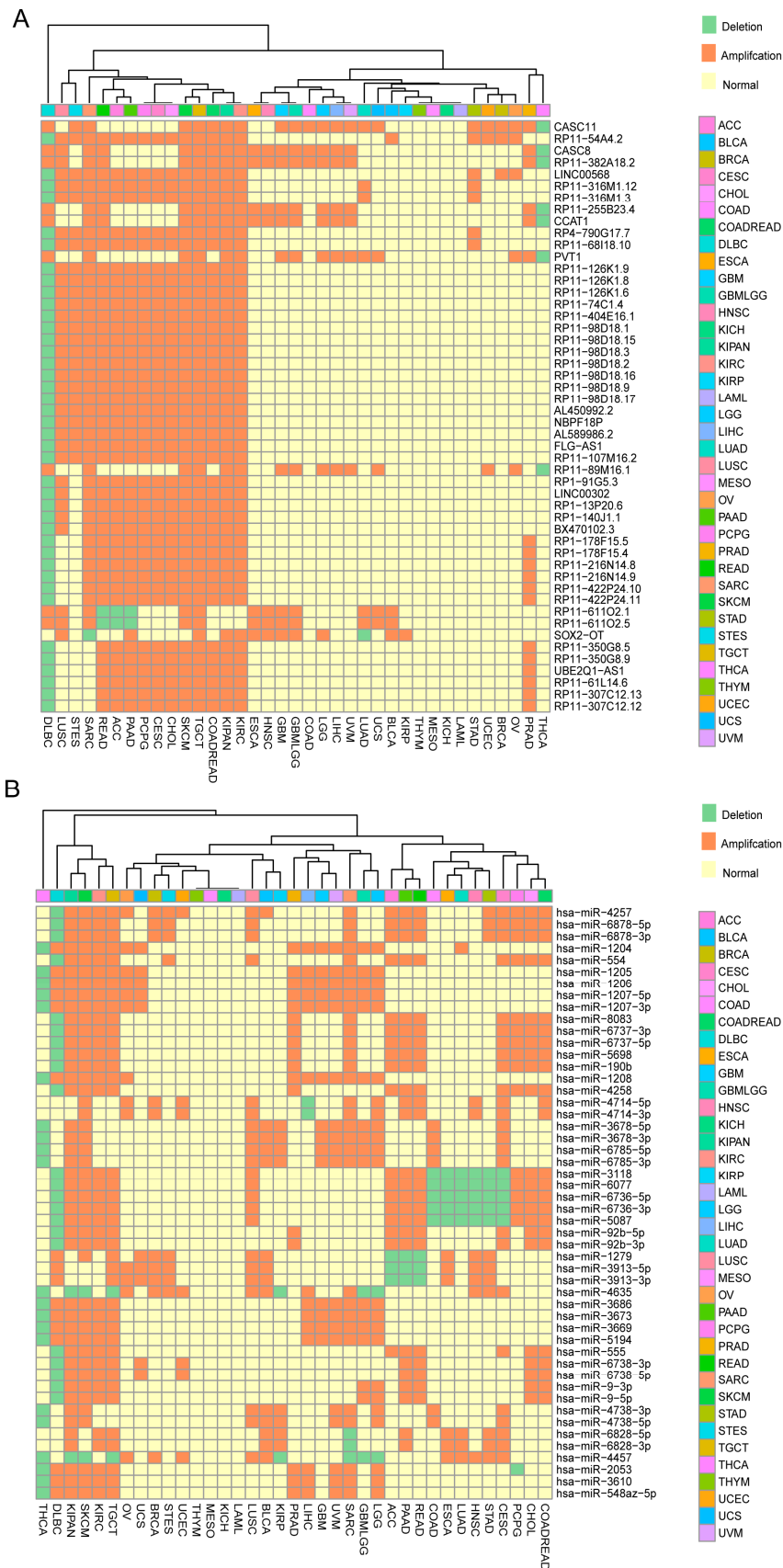


Figure 2. The amplification signatures for the top long non-coding RNAs (lncRNAs) and microRNAs (miRNAs) in the most cancer types. (A) The top 50 lncRNAs annotated by GENCODE V19 with the most cancer types for amplification; (B) The top 50 miRNAs with the most cancer types for amplification in 37 TCGA tumors.

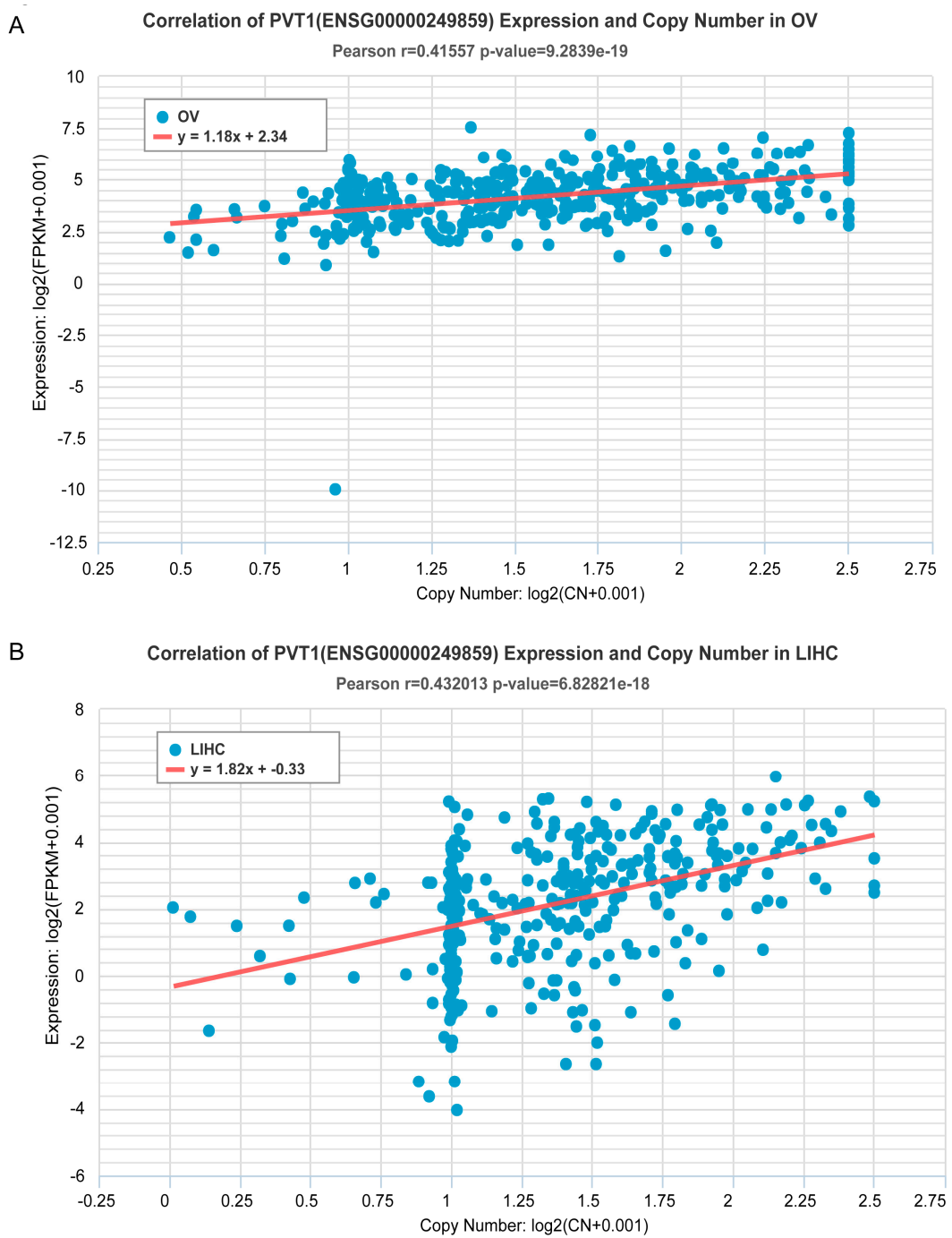


Figure 3. lncRNA PVT1 with significant amplification and overexpression. (A) PVT1 shows the significant and positive correlations between expressions and SCNAs in Ovarian Cancer (OV); (B) PVT1 shows the significant and positive correlations between expressions and SCNAs in Liver Hepatocellular Carcinoma (LIHC).

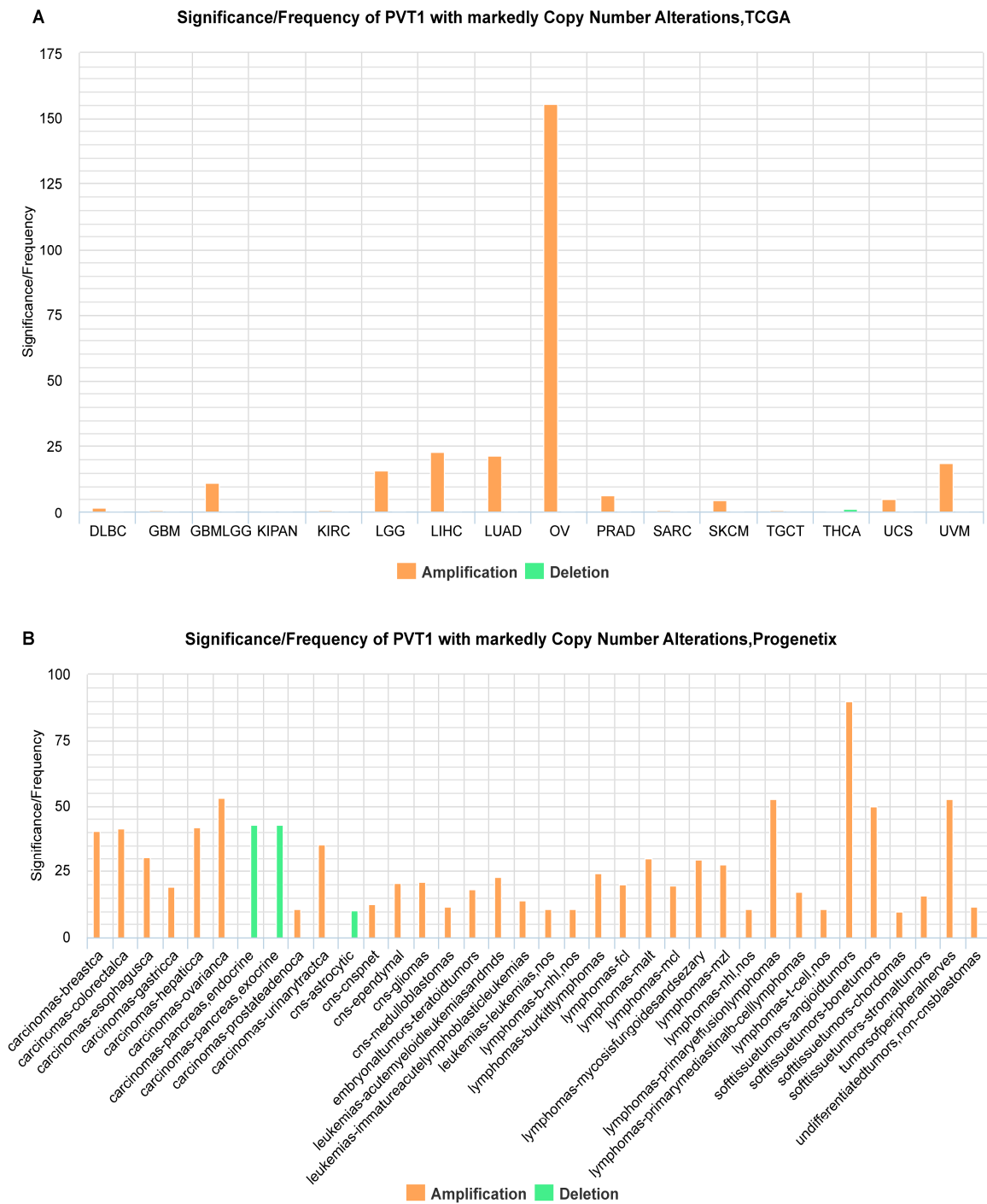


Figure 4. An example showing how users discover the frequently amplified lncRNAs using oncoNcRNA web portal. (A) PVT1 shows recurrent SCNAs in 16 TCGA cancer types. The y axis shows the $-\log_{10}$ (q value); (B) PVT1 shows recurrent SCNAs in 36 Progenetix cancer types. The y axis shows the frequency.

The SCNA usually disrupts the expression of genes within it, therefore we estimated the correlation between gene copy number and expression level through the ‘‘Correlations’’ module. With this module, users can easily find whether the expression of the gene of interest is positively correlated with copy numbers in specific cancer type. Meanwhile, users can identify the cancer with the strongest significance for the gene of interest by ranking the correlations and p values. When clicking on each cancer type, users can obtain a visual scatter plot which presents the correlation. Through

the oncoNcRNA portal, users can rapidly estimate whether the gene of interest has a significant copy number gain/loss within and across tumor types and simultaneously obtain publication-quality figures in a convenient way. A general tutorial is displayed in the “Help and Tutorials” module of website.

3. Discussion

With the application of tiling microarrays and high-throughput sequencing technologies into the investigation of whole genomes and transcriptomes, it is now evident that less than 2% of the genome encodes proteins, whereas at least 75% is actively transcribed into ncRNAs [1]. Accumulating evidence has suggested the extensive links between ncRNAs and cancer, such as miRNAs [44–46] and lncRNAs [47–49]. The most recent study has revealed that the human genome contains over 58,000 lncRNA genes [10], which is far beyond the number of protein-coding genes (~20,000). However, the function of the majority of lncRNAs remains unknown. Despite the fact that some studies have also suggested cancer-related roles of other ncRNAs such as pseudogene [50] and snoRNA [17], our knowledge of the roles of ncRNAs in cancer development remains limited. By integrating large-scale SCNA data from TCGA [18] and other public resources [15,20,30,31], oncoNcRNA reveals a large number of ncRNAs with recurrent SCNAs in human cancers. This may provide a novel insight for the exploration of functional ncRNAs.

oncoNcRNA has unique advantages over other oncogenomic portals, such as Tumorscape [15], COSMIC [21], cBioPortal [22,23] and TANRIC [51]. First of all, oncoNcRNA is mainly designed for exploring the SCNAs of ncRNAs, which is different from other portals with a major focus on protein-coding genes. Second, oncoNcRNA illustrates the ncRNAs with recurrent SCNAs in tumors. Notably, correlations between gene expressions and copy numbers are also estimated based on available data. Third, oncoNcRNA contains all well-annotated ncRNAs, including miRNAs, tRNAs, snoRNAs and approximately 28,000 novel lncRNAs annotated in mitranscriptome [10]. Finally, oncoNcRNA presents the largest-scale integration of SCNA data for cancer tissues and cell lines from TCGA [18], CCLE [20] and other public resources [15,30,31], and covers up to 64 human cancer types. Taken together, to our knowledge, oncoNcRNA is the first oncogenomic portal for exploring oncogenic potential of all kinds of ncRNAs. It may be a valuable tool for the discovery of functional ncRNAs and clinical biomarker development.

4. Materials and Methods

4.1. Collecting and Preprocessing the SCNA Datasets in Cancer Genomes

We downloaded the level 4 data of copy numbers of 37 tumor datasets from The Broad Institute TCGA GDAC (release 20150402) [18]. The data from another two large-scale copy number analyses including over 19 tumor types [15,31] and the Cancer Cell Line Encyclopedia (CCLE) [20] were also included. Among these datasets, significantly recurrent SCNA regions were identified by GISTIC 2.0 software [29]. We also downloaded 84 tumor subsets from the type “CLINICALGROUP” on Progenetix [30] and kept 43 of them which contained at least 10 samples for downstream analysis. Progenetix provided the SCNA information of each tumor sample, and used number “1” to represent amplification while “-1” represented deletion. All genome coordinates were converted to hg19 assemblies by using the UCSC LiftOver Tool [52].

4.2. Integrating a Comprehensive Gene Set from Multiple Sources

We integrated the genes annotated by GENCODE (version 19), mitranscriptome [10], mirBase (version 20) [53], piRNABank [54], snoRNABase [55] and GtRNAdb [56]. Mitranscriptome identified over 58,648 lncRNA genes, of which 79% were previously unannotated [10]. We provided detailed information on the isoforms and sequences of each novel lncRNA in oncoNcRNA.

4.3. Identification of ncRNAs with Oncogenic Potentials in Human Cancers

We identified the ncRNAs with oncogenic potential by adopting the following way. First, the recurrent SCNA regions, the genomic regions which may contain essential oncogenic drivers, were collected from public resources or identified by ourselves. For the TCGA and several other public datasets (excluding the CCLE datasets and Progenetix dataset), their copy number analyses have been performed and recurrent SCNA regions have been identified by GISTIC 2.0 software (Genomic Identification of Significant Targets in Cancer) [29], therefore we adopted these data for our analysis directly. For the CCLE datasets, only segment copy number data were provided, and thus we performed recurrent SCNA region analysis among all cancer cell lines by using the GISTIC 2.0 software. Then, we detected the ncRNAs within the “wide peak” regions using BEDTools (v2.17.0) [57], since the “wide peak” boundaries have been reported to be most likely to contain the targeted oncogenic drivers [29]. An exception was the Progenetix dataset, which preprocessed segment copy number data with number “1” for amplification and “−1” for deletion. This made us unable to identify the recurrent SCNA regions through GISTIC 2.0 software. Therefore, the ncRNAs within these segment DNA copy number alteration regions were directly detected using BEDTools (v2.17.0). Since the ncRNAs with more frequent SCNAs are more likely to act as potential oncogenic drivers, we only selected the genes with higher frequency (>10%) as potential candidates.

4.4. Bioinformatics Analysis

The RNA-seq expression values of 36 cancer types from TCGA which were calculated by Toil pipeline were downloaded from the website [58]. We extracted samples with matched copy number and expression values for each tumor type using custom Perl script. Then both expression and copy number values were log₂ transformed. For ncRNAs with frequent SCNAs identified in the above analyses, we calculated the correlation between gene copy numbers and expression values using the Pearson correlation analysis.

4.5. Database and Web Interface Implementation

All data sets were processed and stored in a MySQL Database Management System (v5.7.11). The database query and user interface were developed using PHP and JavaScript. The query result table was based on jQueryUI (v1.11.4) and DataTables (v1.10.7), which is a highly flexible tool for sorting and filtering the search result. The figure was made with JavaScript Highcharts (v4.2.1) library [59], and allows users to export multiple picture formats. The web development was based on Bootstrap (v3.3.5).

Acknowledgments: We thank The Cancer Genome Atlas (TCGA) project, Cancer Cell Line Encyclopedia (CCLE) and their contributors for providing the valuable data set. Grant support: the National Natural Science Foundation of China (No. 31370791, 30900820, 31230042, 31471223, 91440110, 31671349); funds from Guangdong Province (No. S2012010010510, S2013010012457); The project of Science and Technology New Star in Zhujiang Guangzhou city (No. 2012J2200025); Fundamental Research Funds for the Central Universities (No. 2011330003161070, 14lgjc18); China Postdoctoral Science Foundation (No. 200902348); the project of Science and Technology of Guangzhou (201504010022). This research is supported in part by the Guangdong Province Key Laboratory of Computational Science and the Guangdong Province Computational Science Innovative Research Team. Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase).

Author Contributions: Z.L.W., J.H.Y. and L.H.Q. conceived the project and conception. Z.L.W. performed the computational and statistical analysis. Z.L.W., J.H.Y., X.Q.Z., L.H.Q., and H.Z. wrote the manuscript. L.H.Q. and J.H.Y. supervised the project. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; et al. Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108. [PubMed]

2. Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M.C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; et al. The transcriptional landscape of the mammalian genome. *Science* **2005**, *309*, 1559–1563. [[PubMed](#)]
3. Bertone, P.; Stolc, V.; Royce, T.E.; Rozowsky, J.S.; Urban, A.E.; Zhu, X.; Rinn, J.L.; Tongprasit, W.; Samanta, M.; Weissman, S.; et al. Global identification of human transcribed sequences with genome tiling arrays. *Science* **2004**, *306*, 2242–2246. [[CrossRef](#)] [[PubMed](#)]
4. Jonas, S.; Izaurralde, E. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat. Rev. Genet.* **2015**, *16*, 421–433. [[CrossRef](#)] [[PubMed](#)]
5. Croce, C.M. Causes and consequences of microRNA dysregulation in cancer. *Nat. Rev. Genet.* **2009**, *10*, 704–714. [[CrossRef](#)] [[PubMed](#)]
6. Quinn, J.J.; Chang, H.Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **2015**, *17*, 47–62. [[CrossRef](#)] [[PubMed](#)]
7. Fatica, A.; Bozzoni, I. Long non-coding RNAs: New players in cell differentiation and development. *Nat. Rev. Genet.* **2014**, *15*, 7–21. [[CrossRef](#)]
8. Spizzo, R.; Almeida, M.I.; Colombatti, A.; Calin, G.A. Long non-coding RNAs and cancer: A new frontier of translational research? *Oncogene* **2012**, *31*, 4577–4587. [[CrossRef](#)] [[PubMed](#)]
9. Qiu, M.T.; Hu, J.W.; Yin, R.; Xu, L. Long noncoding RNA: An emerging paradigm of cancer research. *Tumour biology. J. Int. Soc. Oncodev. Biol. Med.* **2013**, *34*, 613–620. [[CrossRef](#)] [[PubMed](#)]
10. Iyer, M.K.; Niknafs, Y.S.; Malik, R.; Singhal, U.; Sahu, A.; Hosono, Y.; Barrette, T.R.; Prensner, J.R.; Evans, J.R.; Zhao, S.; et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* **2015**, *47*, 199–208. [[CrossRef](#)] [[PubMed](#)]
11. Stratton, M.R.; Campbell, P.J.; Futreal, P.A. The cancer genome. *Nature* **2009**, *458*, 719–724. [[CrossRef](#)] [[PubMed](#)]
12. Garraway, L.A.; Lander, E.S. Lessons from the cancer genome. *Cell* **2013**, *153*, 17–37. [[CrossRef](#)] [[PubMed](#)]
13. Santarius, T.; Shipley, J.; Brewer, D.; Stratton, M.R.; Cooper, C.S. A census of amplified and overexpressed human cancer genes. *Nat. Rev. Cancer* **2010**, *10*, 59–64. [[CrossRef](#)] [[PubMed](#)]
14. Zack, T.I.; Schumacher, S.E.; Carter, S.L.; Cherniack, A.D.; Saksena, G.; Tabak, B.; Lawrence, M.S.; Zhang, C.Z.; Wala, J.; Mermel, C.H.; et al. Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **2013**, *45*, 1134–1140. [[CrossRef](#)] [[PubMed](#)]
15. Beroukhi, R.; Mermel, C.H.; Porter, D.; Wei, G.; Raychaudhuri, S.; Donovan, J.; Barretina, J.; Boehm, J.S.; Dobson, J.; Urashima, M.; et al. The landscape of somatic copy-number alteration across human cancers. *Nature* **2010**, *463*, 899–905. [[CrossRef](#)] [[PubMed](#)]
16. Hu, X.; Feng, Y.; Zhang, D.; Zhao, S.D.; Hu, Z.; Greshock, J.; Zhang, Y.; Yang, L.; Zhong, X.; Wang, L.P.; et al. A functional genomic approach identifies FAL1 as an oncogenic long noncoding RNA that associates with BMI1 and represses p21 expression in cancer. *Cancer Cell* **2014**, *26*, 344–357. [[CrossRef](#)] [[PubMed](#)]
17. Siprashvili, Z.; Webster, D.E.; Johnston, D.; Shenoy, R.M.; Ungewickell, A.J.; Bhaduri, A.; Flockhart, R.; Zarnegar, B.J.; Che, Y.; Meschi, F.; et al. The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer. *Nat. Genet.* **2016**, *48*, 53–58. [[CrossRef](#)] [[PubMed](#)]
18. Cancer Genome Atlas Research Network; Weinstein, J.N.; Collisson, E.A.; Mills, G.B.; Shaw, K.R.; Ozenberger, B.A.; Ellrott, K.; Shmulevich, I.; Sander, C.; Stuart, J.M. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **2013**, *45*, 1113–1120. [[PubMed](#)]
19. International Cancer Genome Consortium; Hudson, T.J.; Anderson, W.; Artez, A.; Barker, A.D.; Bell, C.; Bernabe, R.R.; Bhan, M.K.; Calvo, F.; Eerola, I.; et al. International network of cancer genome projects. *Nature* **2010**, *464*, 993–998. [[CrossRef](#)] [[PubMed](#)]
20. Barretina, J.; Caponigro, G.; Stransky, N.; Venkatesan, K.; Margolin, A.A.; Kim, S.; Wilson, C.J.; Lehar, J.; Kryukov, G.V.; Sonkin, D.; et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **2012**, *483*, 603–607. [[CrossRef](#)] [[PubMed](#)]
21. Forbes, S.A.; Beare, D.; Gunasekaran, P.; Leung, K.; Bindal, N.; Boutselakis, H.; Ding, M.; Bamford, S.; Cole, C.; Ward, S.; et al. COSMIC: Exploring the world’s knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **2015**, *43*, D805–D811. [[CrossRef](#)] [[PubMed](#)]
22. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* **2013**, *6*, pl1. [[CrossRef](#)] [[PubMed](#)]

23. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B.E.; Sumer, S.O.; Aksoy, B.A.; Jacobsen, A.; Byrne, C.J.; Heuer, M.L.; Larsson, E.; et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2012**, *2*, 401–404. [[CrossRef](#)] [[PubMed](#)]
24. Fritah, S.; Niclou, S.P.; Azuaje, F. Databases for lncRNAs: A comparative evaluation of emerging tools. *RNA* **2014**, *20*, 1655–1665. [[CrossRef](#)] [[PubMed](#)]
25. Li, J.H.; Liu, S.; Zhou, H.; Qu, L.H.; Yang, J.H. starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* **2014**, *42*, D92–D97. [[CrossRef](#)] [[PubMed](#)]
26. Quek, X.C.; Thomson, D.W.; Maag, J.L.; Bartonicek, N.; Signal, B.; Clark, M.B.; Gloss, B.S.; Dinger, M.E. lncRNADB v2.0: Expanding the reference database for functional long noncoding RNAs. *Nucleic Acids Res.* **2015**, *43*, D168–D173. [[CrossRef](#)] [[PubMed](#)]
27. Volders, P.J.; Helsens, K.; Wang, X.; Menten, B.; Martens, L.; Gevaert, K.; Vandesompele, J.; Mestdagh, P. LNCipedia: A database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res.* **2013**, *41*, D246–D251. [[CrossRef](#)] [[PubMed](#)]
28. Zheng, L.L.; Li, J.H.; Wu, J.; Sun, W.J.; Liu, S.; Wang, Z.L.; Zhou, H.; Yang, J.H.; Qu, L.H. deepBase v2.0: Identification, expression, evolution and function of small RNAs, lncRNAs and circular RNAs from deep-sequencing data. *Nucleic Acids Res.* **2016**, *44*, D196–D202. [[CrossRef](#)] [[PubMed](#)]
29. Mermel, C.H.; Schumacher, S.E.; Hill, B.; Meyerson, M.L.; Beroukhi, R.; Getz, G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **2011**, *12*, R41. [[CrossRef](#)] [[PubMed](#)]
30. Cai, H.; Kumar, N.; Ai, N.; Gupta, S.; Rath, P.; Baudis, M. Progenetix: 12 years of oncogenomic data curation. *Nucleic Acids Res.* **2014**, *42*, D1055–D1062. [[CrossRef](#)] [[PubMed](#)]
31. Kim, T.M.; Xi, R.; Luquette, L.J.; Park, R.W.; Johnson, M.D.; Park, P.J. Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes. *Genome Res.* **2013**, *23*, 217–227. [[CrossRef](#)] [[PubMed](#)]
32. Zhang, L.; Huang, J.; Yang, N.; Greshock, J.; Megraw, M.S.; Giannakakis, A.; Liang, S.; Naylor, T.L.; Barchetti, A.; Ward, M.R.; et al. microRNAs exhibit high frequency genomic alterations in human cancer. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9136–9141. [[CrossRef](#)] [[PubMed](#)]
33. Calin, G.A.; Sevignani, C.; Dumitru, C.D.; Hyslop, T.; Noch, E.; Yendamuri, S.; Shimizu, M.; Rattan, S.; Bullrich, F.; Negrini, M.; et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 2999–3004. [[CrossRef](#)] [[PubMed](#)]
34. Guo, X.; Hua, Y. CCAT1: An oncogenic long noncoding RNA in human cancers. *J. Cancer Res. Clin. Oncol.* **2016**. [[CrossRef](#)] [[PubMed](#)]
35. Ma, M.Z.; Chu, B.F.; Zhang, Y.; Weng, M.Z.; Qin, Y.Y.; Gong, W.; Quan, Z.W. Long non-coding RNA CCAT1 promotes gallbladder cancer development via negative modulation of miRNA-218–5p. *Cell Death Dis.* **2015**, *6*, e1583. [[CrossRef](#)] [[PubMed](#)]
36. Nissan, A.; Stojadinovic, A.; Mitrani-Rosenbaum, S.; Halle, D.; Grinbaum, R.; Roistacher, M.; Bochem, A.; Dayanc, B.E.; Ritter, G.; Gomceli, I.; et al. Colon cancer associated transcript-1: A novel RNA expressed in malignant and pre-malignant human tissues. *Int. J. Cancer J. Int. Cancer* **2012**, *130*, 1598–1606. [[CrossRef](#)] [[PubMed](#)]
37. Xiang, J.F.; Yin, Q.F.; Chen, T.; Zhang, Y.; Zhang, X.O.; Wu, Z.; Zhang, S.; Wang, H.B.; Ge, J.; Lu, X.; et al. Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res.* **2014**, *24*, 513–531. [[CrossRef](#)] [[PubMed](#)]
38. Zhuang, K.; Wu, Q.; Jiang, S.; Yuan, H.; Huang, S.; Li, H. CCAT1 promotes laryngeal squamous cell carcinoma cell proliferation and invasion. *Am. J. Transl. Res.* **2016**, *8*, 4338–4345.
39. Guan, Y.; Kuo, W.L.; Stilwell, J.L.; Takano, H.; Lapuk, A.V.; Fridlyand, J.; Mao, J.H.; Yu, M.; Miller, M.A.; Santos, J.L.; et al. Amplification of PVT1 contributes to the pathophysiology of ovarian and breast cancer. *Clin. Cancer Res.* **2007**, *13*, 5745–5755. [[CrossRef](#)] [[PubMed](#)]
40. Tseng, Y.Y.; Moriarity, B.S.; Gong, W.; Akiyama, R.; Tiwari, A.; Kawakami, H.; Ronning, P.; Reuland, B.; Guenther, K.; Beadnell, T.C.; et al. PVT1 dependence in cancer with MYC copy-number increase. *Nature* **2014**, *512*, 82–86. [[CrossRef](#)] [[PubMed](#)]

41. Wang, F.; Yuan, J.H.; Wang, S.B.; Yang, F.; Yuan, S.X.; Ye, C.; Yang, N.; Zhou, W.P.; Li, W.L.; Li, W.; et al. Oncofetal long noncoding RNA PVT1 promotes proliferation and stem cell-like property of hepatocellular carcinoma cells by stabilizing NOP2. *Hepatology* **2014**, *60*, 1278–1290. [[CrossRef](#)] [[PubMed](#)]
42. Chinnam, M.; Goodrich, D.W. RB1, development, and cancer. *Curr. Top. Dev. Biol.* **2011**, *94*, 129–169. [[PubMed](#)]
43. Song, M.S.; Salmena, L.; Pandolfi, P.P. The functions and regulation of the PTEN tumour suppressor. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 283–296. [[CrossRef](#)] [[PubMed](#)]
44. Lu, J.; Getz, G.; Miska, E.A.; Alvarez-Saavedra, E.; Lamb, J.; Peck, D.; Sweet-Cordero, A.; Ebert, B.L.; Mak, R.H.; Ferrando, A.A.; et al. MicroRNA expression profiles classify human cancers. *Nature* **2005**, *435*, 834–838. [[CrossRef](#)] [[PubMed](#)]
45. Esquela-Kerscher, A.; Slack, F.J. Oncomirs—MicroRNAs with a role in cancer. *Nat. Rev. Cancer* **2006**, *6*, 259–269. [[CrossRef](#)] [[PubMed](#)]
46. Yu, S.L.; Chen, H.Y.; Chang, G.C.; Chen, C.Y.; Chen, H.W.; Singh, S.; Cheng, C.L.; Yu, C.J.; Lee, Y.C.; Chen, H.S.; et al. MicroRNA signature predicts survival and relapse in lung cancer. *Cancer Cell* **2008**, *13*, 48–57. [[CrossRef](#)] [[PubMed](#)]
47. Yan, X.; Hu, Z.; Feng, Y.; Hu, X.; Yuan, J.; Zhao, S.D.; Zhang, Y.; Yang, L.; Shan, W.; He, Q.; et al. Comprehensive Genomic Characterization of Long Non-coding RNAs across Human Cancers. *Cancer Cell* **2015**, *28*, 529–540. [[CrossRef](#)] [[PubMed](#)]
48. Wang, Z.L.; Li, B.; Piccolo, S.R.; Zhang, X.Q.; Li, J.H.; Zhou, H.; Yang, J.H.; Qu, L.H. Integrative analysis reveals clinical phenotypes and oncogenic potentials of long non-coding RNAs across 15 cancer types. *Oncotarget* **2016**, *7*, 35044–35055. [[CrossRef](#)] [[PubMed](#)]
49. Huarte, M. The emerging role of lncRNAs in cancer. *Nat. Med.* **2015**, *21*, 1253–1261. [[CrossRef](#)] [[PubMed](#)]
50. Han, L.; Yuan, Y.; Zheng, S.; Yang, Y.; Li, J.; Edgerton, M.E.; Diao, L.; Xu, Y.; Verhaak, R.G.; Liang, H. The Pan-Cancer analysis of pseudogene expression reveals biologically and clinically relevant tumour subtypes. *Nat. Commun.* **2014**, *5*, 3963. [[CrossRef](#)] [[PubMed](#)]
51. Li, J.; Han, L.; Roebuck, P.; Diao, L.; Liu, L.; Yuan, Y.; Weinstein, J.N.; Liang, H. TANRIC: An interactive open platform to explore the function of lncRNAs in cancer. *Cancer Res.* **2015**, *75*, 3728–3737. [[CrossRef](#)] [[PubMed](#)]
52. Meyer, L.R.; Zweig, A.S.; Hinrichs, A.S.; Karolchik, D.; Kuhn, R.M.; Wong, M.; Sloan, C.A.; Rosenbloom, K.R.; Roe, G.; Rhead, B.; et al. The UCSC Genome Browser database: Extensions and updates 2013. *Nucleic Acids Res.* **2013**, *41*, D64–D69. [[CrossRef](#)] [[PubMed](#)]
53. Kozomara, A.; Griffiths-Jones, S. miRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **2014**, *42*, D68–D73. [[CrossRef](#)] [[PubMed](#)]
54. Sai Lakshmi, S.; Agrawal, S. piRNABank: A web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Res.* **2008**, *36*, D173–D177. [[CrossRef](#)] [[PubMed](#)]
55. Lestrade, L.; Weber, M.J. snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res.* **2006**, *34*, D158–D162. [[CrossRef](#)] [[PubMed](#)]
56. Chan, P.P.; Lowe, T.M. GtRNAdb: A database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.* **2009**, *37*, D93–D97. [[CrossRef](#)] [[PubMed](#)]
57. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [[CrossRef](#)]
58. Vivian, J. Rapid and efficient analysis of 20,000 RNA-seq samples with Toil. Available online: <https://xenabrowser.net/datapages/?host=https://toil.xenahubs.net> (accessed on 15 July 2016).
59. Honsi, T. Highcharts JS. Available online: <http://www.highcharts.com/> (accessed on 1 November 2016).

