Prognostic Efficacy of the RTN1 Gene in Patients with Diffuse Large B-Cell Lymphoma

Mohamad Zamani-Ahmadmahmudi (zamani_2012@ut.ac.ir)
Fatemeh Soltani-Nezhad
Shahid Bahonar University of Kerman
Amir Asadabadi
Shahid Bahonar University of Kerman

Research article

Keywords: DLBCL, RTN1, Prognosis, Survival

Posted Date: July 16th, 2019

DOI: https://doi.org/10.21203/rs.2.11441/v1

Abstract

Background Gene expression profiling has been vastly used to extract genes that can predict the clinical outcome in patients with diverse cancers, including diffuse large B-cell lymphoma (DLBCL). With the aid of bioinformatics and computational analysis on gene expression data, various prognostic gene signatures for DLBCL have been recently developed. The major drawback of the previous signatures is their inability to correctly predict survival in external data sets. In other words, they are not reproducible in other datasets. Hence, in this study, we sought to determine the gene(s) that can reproducibly and robustly predict survival in patients with DLBCL. Methods Gene expression data were extracted from 7 datasets containing 1636 patients (GSE10846 [n=420], GSE31312 [n=470], GSE11318 [n=203], GSE32918 [n=172], GSE4475 [n=123], GSE69051 [n=157], and GSE34171 [n=91]). Genes significantly associated with overall survival were detected using the univariate Cox proportional hazards analysis with a P value <0.001 and a false discovery rate (FDR) <5%. Thereafter, significant genes common between all the datasets were extracted. Additionally, chromosomal aberrations in the corresponding region of final common gene(s) were evaluated as copy number alterations using the single nucleotide polymorphism (SNP) data of 570 patients with DLBCL (GSE58718 [n=242], GSE57277 [n=148], and GSE34171 [n=180]). The results were experimentally confirmed using the quantitative real-time PCR (qRT-PCR) analysis. Results Our results indicated that reticulon family gene 1 (RTN1) was the only gene that met our rigorous pipeline criteria and associated with a favorable clinical outcome in all the datasets (P<0.001, FDR<5%). In the multivariate Cox proportional hazards analysis, this gene remained independent of the routine international prognostic index components (i.e., age, stage, lactate dehydrogenase level, Eastern Cooperative Oncology Group [ECOG] performance status, and number of extranodal sites) (P<0.0001). Our experimental step confirmed the results and revealed that the expression of RTN1 in the long-survival group was significantly higher than that in the short-survival group. Furthermore, no significant chromosomal aberration was found in the RTN1 genomic region (14q23.1: Start 59,595,976/ End 59,870,966). Conclusion In light of the results of present study, RTN1 can be considered a potential prognostic gene that can robustly predict survival in patients with DLBCL.

Background

Reticulon family gene 1 (*RTN1*) (formerly termed “neuroendocrine-specific protein” [NSP]) is a reticulon-encoding gene that is associated with the endoplasmic reticulum. Reticulons play critical roles in membrane trafficking or neuroendocrine secretion in neuroendocrine cells. *RTN1* encodes 3 variants—namely NSP-A, -B, and –C—which are attached to the endoplasmic reticulum by means of 2 putative transmembrane domains in the homologous C-terminal region [1-4].

Previous investigations have introduced *RTN1* as a potential diagnostic/therapeutic marker of neurological diseases and cancers [2, 5-7]. RTN1 was proposed as a potential marker for carcinomas with neuroendocrine characteristics [2]. It has been shown that RTN1 reduces the anti-apoptotic activity of a protein encoded by BCL2-like 1 (BCL2L1) (ie, B-cell lymphoma-extra large [Bcl-xL]). Indeed, RTN1 can

change the subcellular localization of the Bcl-xL protein from the mitochondria to the endoplasmic reticulum, which disrupts its anti-apoptotic action [5].

Because of the major shortages of previous prognostic gene signatures developed based on gene expression profiling [8-13], we sought to find the gene(s) that can reproducibly predict the clinical outcome in patients with diffuse large B-cell lymphoma (DLBCL). Some of the shortcomings of the previous signatures hindering their clinical utility include the infeasibility to reproduce a prognostic signature in external datasets, negligible overlaps between the developed signatures, and large numbers of genes in the developed prognostic genes (180 genes, 90 genes, and 27 genes in signatures developed by Lenz et al [2008], Alizadeh et al [2000], and Wright et al [2003], respectively). In our efforts to find the gene(s) reproducibly associated with survival via bioinformatics and computational approaches, we obtained a surprising result: The *RTN1* gene was robustly and reliably associated with a favorable outcome in 1636 patients with DLBCL (including 7 gene expression data sets). Furthermore, the *RTN1* gene remained as one of the most powerful independent prognostic factors in comparison with the international prognostic index (IPI) components.

Methods

Datasets

The Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) database was searched to find the gene expression profiling datasets of patients with DLBCL. Only datasets containing clinical metadata (especially overall survival) (7 datasets) were retained, and the rest were excluded. Additionally, every effort was made to select expression datasets from all types of microarray chips such as Affymetrix and Illumina, if possible. The datasets were downloaded in SOFT file format and were subsequently transformed logarithmically using tools provided in *geWorkbench 2.5.1* package [14], if necessary. More details on the clinical characteristics of the studied datasets are provided in Table 1. The datasets included GSE10846 (n= 420), GSE31312 (n=470), GSE32918 (n=172), GSE69051 (n=157), GSE4475 (n=123), GSE11318 (n=203), and GSE34171 (n=91). Since GSE32918 and GSE69051 have originated from a similar research study [15] and have some common samples, they were merged as a single data set and termed “GSE32918/69051”. The number of samples for these datasets was determined after corrections were made based on the common samples (172 samples for GSE32918 and 157 samples for GSE69051). In addition, the genetic aberrations of the desired gene(s) at the genome level were evaluated by employing 3 single nucleotide polymorphism (SNP) array data sets—namely GSE58718 (n=242), GSE57277 (n=148), and GSE34171 (n=180). GSE34171 contains both gene expression and SNP data (Table 1).

Identification of the common gene(s) associated with survival in the gene expression data sets

The association between gene expression and overall survival was examined using the univariate Cox proportional hazards analysis. In this analysis, the association between a group of covariates (genes) and the response variable (overall survival) was evaluated. The univariate Cox analysis was performed

using *BRB-Array tools*, developed by Richard Simon and the BRB-ArrayTools Development Team. In this analysis, the findings were strengthened by employing rigorous pipeline criteria and retaining only genes with a *P* value <0.001 and a false discovery rate (FDR) <5%. Subsequently, the common gene(s) significantly associated with overall survival between all the data sets was/were extracted. For this purpose, only common gene(s) with consistent associations were selected, while genes with inconsistent associations (negatively associated with overall survival in a dataset and positively associated with overall survival in another) were excluded. Moreover, the patients were categorized into 2 risk groups (high-risk vs. low-risk) based on the median of the selected common gene expression values (> median value vs. < median value), and overall survival was compared between the groups using the Kaplan–Meier analysis and log-rank test at a *P* value <0.01. The Kaplan–Meier analysis and the log-rank test were performed in *SPSS 16.0* package (Chicago, USA). Since *RTN1* was the only gene that fulfilled the criteria and was selected as the final gene, the subsequent analyses were exclusively performed on this gene.

As a confirmatory step, we checked whether *RTN1* was differentially expressed between the 2 predefined survival classes using the significance analysis of microarray (SAM) analysis. In this analysis, 2 classes (long survival [≥5 y] vs. short survival [<5 y]) were created and, thereafter, the genes that were differentially expressed were detected. The SAM analysis was performed using the method added in *BRB-Array tools*. In this analysis, the FDR and the number of permutations were set at 5% and 1000, respectively.

Prognostic efficacy of the *RTN1* gene in a multivariate model

The prognostic efficacy of *RTN1* was also evaluated in a multivariate Cox proportional-hazards regression analysis, where the *RTN1* gene expression and all the individual components of the international prognostic index (IPI) (ie, age, stage, lactate dehydrogenase level, Eastern Cooperative Oncology Group [ECOG] performance status, and number of extranodal sites) [16] were entered as covariate variables. Additionally, the molecular subtype (ie, ABC-like, GCB-like, and type 3) and sex were incorporated as another 2 variables into the model. The multivariate analysis was performed solely on the datasets with the clinical IPI data (i.e., GSE10846 and GSE31312). This analysis was carried out using *Survival* package (http://cran.r-project.org/package=survival) and *SPSS 16.0* package (Chicago, USA).

Prognostic efficacy of the *RTN1* gene in molecular subtypes of DLBCL

We also checked the prognostic worth of *RTN1* in molecular subtypes of DLBCL (i.e., ABC-like, GCB-like, and type 3) using similar strategy described above. In brief, in each subtypes two risk groups constituted based on median of the *RTN1* expression values (> median value vs. < median value) and then overall survival was compared between the groups using the Kaplan–Meier analysis and log-rank test at a *P* value <0.01.

Correlation between *RTN1* and *BCL2L1* expressions

One of the main targets of *RTN1* is *BCL2L1* via the inhibition of its anti-apoptotic activity [5]. Moreover, BCL2L1 is a member of BCL-2 protein families deregulated in lymphoma tumors, especially DLBCL [17-19]. Accordingly, the associations between 2 probe-sets of *RTN1* (i.e., 203485_at and 210222_s_at) and 4 probe-sets of *BCL2L1* (i.e., 206665_s_at, 212312_at, 215037_s_at, and 231228_at) were evaluated using correlation analysis. The correlations were graded based on the classification proposed by Papasouliotis et al (2006) [20](i.e., r=0.93 to 0.100 as excellent, r=0.80 to 0.92 as good, r=0.59 to 0.79 as fair, and r<0.59 as poor correlations). The correlation analysis was performed using *SPSS 16.0* package (Chicago, USA) in all the datasets, and a *P* value <0.05 was considered significant.

Evaluation of *RTN1* at the genome level in the DLBCL samples

For the assessment of the chromosomal aberrations of the *RTN1* gene, 3 datasets—namely GSE58718 (n=242), GSE57277 (n=148), and GSE34171 (n=180)—were used to extract copy number variations (CNVs) from the SNP data. GSE58718 was generated based on Illumina HumanCNV370-Duov1 DNA Analysis BeadChip, while GSE57277 and GSE34171 were generated using Affymetrix Mapping 250K SNP Arrays. In brief, *PennCNV* package [21] was used to call and analyze the CNV data. For the Illumina datasets, signal intensity data in the form of log R ratios (LRRs) and B allele frequencies (BAFs) were directly generated from the downloaded raw file. For the Affymetrix datasets, LRRS and BAFs were calculated by processing raw intensity (.CEL) files in *Affymetrix Power Tools* (https://www.affymetrix.com/support/developer/powertools/changelog/index.html), followed by *PennCNV-Affy* package. Finally, these LRRS and BAFs were used to generate CNV calls. CNVs with lengths <1 kb, confidence scores <10, or containing <5 SNPs were discarded. A CNV was considered to be a recurring acquired copy number alteration (rCNA) if it occurred in more than 2.5% of the patients and was not reported in the Database of Genomic Variants, build 36 (hg18) (DGV, http://projects.tcag.ca/variation/) [22]. The location of *RTN1* was explored on chromosome 14 (14q23.1: Start 59,595,976/End 59,870,966) for chromosomal aberrations.

Verification of the results by quantitative real-time PCR (qRT-PCR)

Thirty patients (20 males and 10 females) with DLBCL were recruited. Formalin-fixed paraffin-embedded (FFPE) tissue specimens were used to assay *RTN1* and BCL2L1 expressions via the qRT-PCR technique. Samples were retrieved from tissue banks at Shahid Bahonar Hospital and Shafa Hospital (Kerman University of Medical Sciences). The diagnosis of DLBCL was established by 2 expert pathologists according to the revised European/American lymphoma classification. The mean age of the patients was 60 years (range 22-81 years). The median follow-up was 49 months (range: 6–80 month). All patients were treated with a regimen that included an anthracycline (cyclophosphamide, doxorubicin, vincristine, and prednisone [CHOP] or CHOP-like regimens. The patients were divided into two survival groups (long survival [≥3 y] vs. short survival [<3 y]). The study protocol was approved by our institutional review board, and written informed consent was obtained from all the patients. We also used 7 normal lymph nodes as controls.

The qRT-PCR procedure was carried out as previously described [23-24] with some modifications. Total RNA was extracted from the FFPE tissues according to a method described elsewhere (https://bio-protocol.org/e161). In brief, paraffin was first removed from the FFPE tissues using 100% xylene and then total RNA was extracted using Tripure isolation reagent (Sigma-Aldrich, USA). The quality and the quantity of the isolated RNA were examined with a NanoDrop spectrophotometer (Roche, Germany). Complementary DNA (cDNA) was synthesized with a Maxime RT PreMix Kit (Parstous, Iran) according to the manufacturer's instructions. The cDNA synthesis reaction was run at 47 °C for 60 minutes, followed by 85 °C for 5 minutes. SYBR green-based qRT-PCR was performed on the synthesized cDNA with the Roche Real-Time PCR System. Cycle conditions were 95 °C for 10 minutes, followed by 40 cycles of 95 °C for 15 seconds, 52 °C for 45 seconds, and 72 °C for 1 minute. Relative quantifications values were calculated as described previously, where for each gene, relative quantifications values > 1 and < 1 indicated the up-regulation and down-regulation of that gene, respectively [25]. The data are presented as fold changes in the gene expression level of the target gene. The fold changes in the gene expressions were compared between the 2 groups (long-survival vs. short-survival) using the Student *t*-test at a *P* value <0.05. The primers designed for *RTN1*, *BCL2L1*, and *HPRT* are presented in Additional Table 1. *HPRT* was employed as the reference gene for normalization.

Results

RTN1 as the most robust and reproducible prognostic gene in all the data sets

First, a univariate Cox proportional hazards analysis was run so as to find genes significantly associated with overall survival in all the datasets. The analysis revealed that 3 genes—namely *APOC1*, *RTN1*, and *PLAU*—fulfilled the criteria and were significantly associated with the clinical outcome at an FDR <10% and a *P* value <0.001. When the FDR cutoff value was set at 5%, only *RTN1* met our rigorous pipeline criteria and was significantly associated with survival at a *P* value <0.001 in the 1636 patients encompassing all the 7 data sets (Table 2). Meanwhile, this gene showed a consistent positive association with survival in all the data sets (hazard ratio range [HR]: 0.41 to 0.79) (Table 2). Overall survival was significantly different between the low-risk and high-risk groups reconstructed based on the median of the *RTN1* expression values (> median value vs. < median value) at a *P* value < 0.0001. The rates of overall survival at 5 years in the high-risk and low-risk groups in the different datasets were as follows: GSE31312 (51% vs. 72%), GSE10846 (48% vs. 67%), GSE32916/69051 (43% vs. 60%), GSE4475 (18% vs. 58%), GSE34171 (60% vs. 87%), and GSE11318 (36% vs. 60%) (Fig. 1). Moreover, *RTN1* was differentially expressed between the 2 classes (i.e., long survival vs. short survival) in all the datasets in the SAM analysis (*P*<0.001).

Our verification experiment confirmed these results and verified that the expression of *RTN1* was significantly higher in the long-survival group than in the short-survival group. By contrast, the difference in *BCL2L1* expression between the 2 risk groups was not significant (Fig. 2).

Additionally, the multivariate Cox proportional hazards analysis indicated that *RTN1* remained independent of routine IPI components in both GSE10846 (HR: 0.78 [0.67 to 0.90]) and GSE31312 (HR: 0.77 [0.66 to 0.88]) (*Ps*<0.0001). Nonetheless, among the IPI parameters, only age remained an independent predictor in both datasets. Some other the IPI factors were only significant in 1 dataset (Table 3).

Expression of *RTN1* in the molecular subtypes of DLBCL

The expression of *RTN1* was compared between the different molecular subtypes of DLBCL (i.e., ABC-like, GCB-like, and type 3) using the *one-way ANOVA* test. The examination indicated that the expression of *RTN1* was significantly higher in the subtype with the better overall survival (i.e., GCB-like) than in the subtype with the inferior survival (i.e., ABC-like) (*Ps*<0.05) in both GSE10846 and GSE31312. We also checked whether overall survival was significantly different between the groups based on *RTN1* in the different molecular subtypes of DLBCL. Our analysis revealed that overall survival was significantly different between the 2 risk groups in the GCB-like subtype (*Ps*<0.05), whereas there was no significant association between the 2 risk groups in the other subtypes (i.e., ABC-like and type 3) (Fig. 3).

Correlation between *RTN1* and *BCL2L1* expressions

Tagamei et al (2000) indicated that RTN1 only changes the subcellular localization of Bcl-xL from mitochondria to the endoplasmic reticulum and does not alter the expression level of the corresponding gene (i.e., *BCL2L1*). Our analysis revealed no significant and consistent correlations between the *RTN1* probe-sets (n=2) and the *BCL2L1* probe-sets (n=4), where some inconsistent (a mix of positive and negative results) and poor correlation coefficients (r<0.59) were obtained in the different analyses (Fig. 4 and Additional Table 2). Nevertheless, an elevation in *RTN1* expression did not suppress *BCL2L1* expression. There was a good and significant correlation between the 2 *RTN1* probe-sets (r=0.82, *P*<0.01) (Fig. 5 and Additional Table 2). These findings were confirmed in the experimental validation procedure, where the expression level of *BCL2L1* was not significantly different between the long-survival and short-survival groups, while the long-survival group had a higher level of *RTN1* expression than did the short-survival group (Fig. 2).

RTN1 at the genome level

Possible CNAs of the *RTN1* gene (14q23.1: Start 59,595,976/End 59,870,966) were checked through an analysis on the SNP data of 570 patients with DLBCL. The upstream and downstream regions that might partly include the *RTN1* region were also explored in order to detect possible aberrations. The analysis indicated no significant chromosomal aberrations in the region of the *RTN1* gene (14q23.1: Start 59,595,976/End 59,870,966). Nonetheless rare chromosomal gains (CN=3, amplification of 59,526,724-60,901,663 segment) and losses (CN=1, loss of 59,580,209-60,388,423 segment) were detected in 2 samples of GSE58718 and GSE34171, correspondingly, where these abnormalities were not recurrent (frequency <1%). No chromosomal instability was found in the region of interest in GSE57277.

Discussion

Several studies have proposed various prognostic signatures comprising different numbers of genes via gene expression analysis [8-13]. Indeed, there is a minimum of overlap between these signatures and in many situations, there is no common gene in the suggested signatures. In addition, we found no genes of the previous signatures that were reproducibly associated with the clinical outcome in our data sets. Another disadvantage of these signatures is their large number of genes. Indubitably, the use of such large signatures is impractical in routine clinical practice [26-27].

In the current study, we assessed the prognostic efficacy of *RTN1* in several large cohorts of patients with DLBCL. This gene was consistently associated with a favorable outcome in all the datasets, comprised of 1636 patients with DLBCL. The association between the *RTN1* gene expression and a favorable clinical outcome in all the datasets was significant at an FDR <5%, which means that the probability of a false positive was extremely low. Furthermore, it remained as a one of the most powerful independent prognostic factors in comparison with the IPI components. Although *RTN1* was not the most powerful gene associated with overall survival, it was the only gene that reproducibly predicted a favorable clinical outcome. This gene was previously reported as a member of the stromal-1 signature in a 108-gene model developed by Lenz et al (2008). Additionally, the upregulation of *RTN1* in CXCR4- DLBCL versus CXCR4+ DLBCL was indicated, where CXCR4- and CXCR4+ subtypes were associated with better and poorer overall survival, respectively [28].

Various roles of RTN1 in the biology of cancers have been previously investigated. As was previously described in the introduction, RTN1 induces its antitumor activity through interaction with Bcl-xL on the endoplasmic reticulum and reduces its anti-apoptotic activity [5]. A previous investigation revealed that a member of the RTN family (ie, RTN-1C) sensitizes neuroepithelioma cells to fenretinide-induced apoptosis through interaction with glucosylceramide synthase [29]. Moreover, *RTN1*-encoded proteins have been proposed as a category criterion for human lung cancer [30]. Previous research has also demonstrated that NSP-reticulon expression is restricted to lung carcinoma cells with a neuroendocrine phenotype [6-7, 30]. Another RTN1 paralog (i.e., RTN3) has a similar antitumor activity through the enrichment of TRAIL-mediated apoptosis via the downregulation of c-FLIP and the upregulation of death receptor 5 [31].

Conclusions

In light of the results of the present study, *RTN1* can be considered a potential prognostic gene that can robustly predict survival in patients with DLBCL. Further studies can explore the prognostic efficacy of *RTN1* in depth. In this way, the prognostic efficacy of this gene can be well again compared with the regular prognostic parameters.

Abbreviations

ABC: activated B cell-like; BAF: B allele frequency; Bcl-xL: B-cell lymphoma-extra large; CNV: copy number variations; DLBCL: diffuse large B-cell lymphoma; ECOG: Eastern Cooperative Oncology Group; FDR: false discovery rate; FFPE: Formalin-fixed paraffin-embedded; GCB: germinal center B cell-like; GEO: Gene Expression Omnibus; HR: hazard ratio range; IPI: international prognostic index; LRR: log R ratio; NSP: neuroendocrine-specific protein; qRT-PCR: quantitative real-time PCR; rCNA: recurring acquired copy number alteration; RTN1: reticulon family gene 1; SAM: significance analysis of microarray; SNP: single nucleotide polymorphism.

Declarations

Acknowledgements

We wish to thank Mr. Pedram Amouzadeh who assisted in the proof-reading of the manuscript. We also thank personals of pathology labs of the Bahonar hospital and Shafa hospital for providing tumor samples.

Funding

This study had no funding support.

Availability of data and materials

The datasets in the manuscript weredeposited in GEO database (http://www.ncbi.nlm.nih.gov/geo/) with the accession numbers provided in Table 1. Other supporting data are included as supplementary files.

Authors' Contributions

MZA, FS, and AA participated in the study design and analysis of the data. MZA and FS wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The study protocol was approved by our institutional review board, and written informed consent was obtained from all the patients entered into the study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Roebroek AJ, Contreras B, Pauli IG, Van de Ven WJ: **cDNA cloning, genomic organization, and expression of the human RTN2 gene, a member of a gene family encoding reticulons**. *Genomics* 1998, **51**(1):98-106.
2. Roebroek AJ, van de Velde HJ, Van Bokhoven A, Broers JL, Ramaekers FC, Van de Ven WJ: **Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product**. *J Biol Chem* 1993, **268**(18):13439-13447.
3. Senden NH, Timmer ED, Boers JE, van de Velde HJ, Roebroek AJ, Van de Ven WJ, Broers JL, Ramaekers FC: **Neuroendocrine-specific protein C (NSP-C): subcellular localization and differential expression in relation to NSP-A**. *Eur J Cell Biol* 1996, **69**(3):197-213.
4. GrandPré T, Nakamura F, Vartanian T, Strittmatter SM: **Identification of the Nogo inhibitor of axon regeneration as a Reticulon protein**. *Nature* 2000, **403**(6768):439-444.
5. Tagami S, Eguchi Y, Kinoshita M, Takeda M, Tsujimoto Y: **A novel protein, RTN-XS, interacts with both Bcl-XL and Bcl-2 on endoplasmic reticulum and reduces their anti-apoptotic activity**. *Oncogene* 2000, **19**(50):5736-5746.
6. Senden N, Linnoila I, Timmer E, van de Velde H, Roebroek A, Van de Ven W, Broers J, Ramaekers F: **Neuroendocrine-specific protein (NSP)-reticulons as independent markers for non-small cell lung cancer with neuroendocrine differentiation. An in vitro histochemical study**. *Histochem Cell Biol* 1997, **108**(2):155-165.
7. Senden NH, Timmer ED, de Bruïne A, Wagenaar SS, Van de Velde HJ, Roebroek AJ, Van de Ven WJ, Broers JL, Ramaekers FC: **A comparison of NSP-reticulons with conventional neuroendocrine markers in immunophenotyping of lung cancers**. *J Pathol* 1997, **182**(1):13-21.
8. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X *et al*: **Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling**. *Nature* 2000, **403**(6769):503-511.
9. Lossos IS, Czerwinski DK, Alizadeh AA, Wechser MA, Tibshirani R, Botstein D, Levy R: **Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes**. *N Engl J Med* 2004, **350**(18):1828-1837.
10. Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM *et al*: **The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma**. *N Engl J Med* 2002, **346**(25):1937-1947.
11. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM: **A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma**. *Proc Natl Acad Sci U S A* 2003, **100**(17):9991-9996.
12. Lenz G, Wright G, Dave SS, Xiao W, Powell J, Zhao H, Xu W, Tan B, Goldschmidt N, Iqbal J *et al*: **Stromal gene signatures in large-B-cell lymphomas**. *N Engl J Med* 2008, **359**(22):2313-2323.
13. Shipp MA, Ross KN, Tamayo P, Weng AP, Kutok JL, Aguiar RCT, Gaasenbeek M, Angelo M, Reich M, Pinkus GS *et al*: **Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning**. *Nat Med* 2002, **8**(1):68-74.

14. Floratos A, Smith K, Ji Z, Watkinson J, Califano A: **geWorkbench: an open source platform for integrative genomics**. *Bioinformatics* 2010, **26**(14):1779-1780.
15. Barrans SL, Crouch S, Care MA, Worrillow L, Smith A, Patmore R, Westhead DR, Tooze R, Roman E, Jack AS: **Whole genome expression profiling based on paraffin embedded tissue can be used to classify diffuse large B-cell lymphoma and predict clinical outcome**. *Br J Haematol* 2012, **159**(4):441-453.
16. Sehn LH, Gascoyne RD: **Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity**. *Blood* 2015, **125**(1):22-32.
17. Tsuyama N, Sakata S, Baba S, Mishima Y, Nishimura N, Ueda K, Yokoyama M, Terui Y, Hatake K, Kitagawa M *et al*: **BCL2 expression in DLBCL: reappraisal of immunohistochemistry with new criteria for therapeutic biomarker evaluation**. *Blood* 2017, **130**(4):489-500.
18. Iqbal J, Neppalli VT, Wright G, Dave BJ, Horsman DE, Rosenwald A, Lynch J, Hans CP, Weisenburger DD, Greiner TC *et al*: **BCL2 expression is a prognostic marker for the activated B-cell-like type of diffuse large B-cell lymphoma**. *J Clin Oncol* 2006, **24**(6):961-968.
19. Iqbal J, Sanger WG, Horsman DE, Rosenwald A, Pickering DL, Dave B, Dave S, Xiao L, Cao K, Zhu Q *et al*: **BCL2 translocation defines a unique tumor subset within the germinal center B-cell-like diffuse large B-cell lymphoma**. *Am J Pathol* 2004, **165**(1):159-166.
20. Papasouliotis K, Cue S, Crawford E, Pinches M, Dumont M, Burley K: **Comparison of white blood cell differential percentages determined by the in-house LaserCyte hematology analyzer and a manual method**. *Vet Clin Pathol* 2006, **35**(3):295-302.
21. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SFA, Hakonarson H, Bucan M: **PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data**. *Genome Res* 2007, **17**(11):1665-1674.
22. Bouska A, McKeithan TW, Deffenbacher KE, Lachel C, Wright GW, Iqbal J, Smith LM, Zhang W, Kucuk C, Rinaldi A *et al*: **Genome-wide copy-number analyses reveal genomic abnormalities involved in transformation of follicular lymphoma**. *Blood* 2014, **123**(11):1681-1690.
23. Zamani-Ahmadmahmudi M, Aghasharif S, Ilbeigi K: **Prognostic efficacy of the human B-cell lymphoma prognostic genes in predicting disease-free survival (DFS) in the canine counterpart**. *BMC Vet Res* 2017, **13**(1):17.
24. Klopfleisch R, Lenze D, Hummel M, Gruber AD: **Metastatic canine mammary carcinomas can be identified by a gene expression profile that partly overlaps with human breast cancer profiles**. *BMC Cancer* 2010, **10**(1).
25. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method**. *Methods* 2001, **25**(4):402-408.
26. Alizadeh AA, Gentles AJ, Alencar AJ, Liu CL, Kohrt HE, Houot R, Goldstein MJ, Zhao S, Natkunam Y, Advani RH *et al*: **Prediction of survival in diffuse large B-cell lymphoma based on the expression of 2 genes reflecting tumor and microenvironment**. *Blood* 2011, **118**(5):1350-1358.

27. Haibe-Kains B, Desmedt C, Sotiriou C, Bontempi G: **A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all?** *Bioinformatics* 2008, **24**(19):2200-2208.
28. Chen J, Xu-Monette ZY, Deng L, Shen Q, Manyam GC, Martinez-Lopez A, Zhang L, Montes-Moreno S, Visco C, Tzankov A *et al*: **Dysregulated CXCR4 expression promotes lymphoma cell survival and independently predicts disease progression in germinal center B-cell-like diffuse large B-cell lymphoma**. *Oncotarget* 2015, **6**(8):5597-5614.
29. Di Sano F, Fazi B, Citro G, Lovat PE, Cesareni G, Piacentini M: **Glucosylceramide synthase and its functional interaction with RTN-1C regulate chemotherapeutic-induced apoptosis in neuroepithelioma cells**. *Cancer Res* 2003, **63**(14):3860-3865.
30. van de Velde HJ, Senden NH, Roskams TA, Broers JL, Ramaekers FC, Roebroek AJ, Van de Ven WJ: **NSP-encoded reticulons are neuroendocrine markers of a novel category in human lung cancer diagnosis**. *Cancer Res* 1994, **54**(17):4769-4776.
31. Lee JT, Lee T-J, Kim C-H, Kim N-S, Kwon TK: **Over-expression of Reticulon 3 (RTN3) enhances TRAIL-mediated apoptosis via up-regulation of death receptor 5 (DR5) and down-regulation of c-FLIP**. *Cancer Lett* 2009, **279**(2):185-192.

Tables

Table 1. Clinical characteristics of the microarray datasets used in our study

Dataset	Number of patients	Chip manufacturer	Platform	Usage in our study
GSE10846	420	Affymetrix	GPL570	Survival analysis
GSE31312	470	Affymetrix	GPL570	Survival analysis
GSE32918	172	Illumina	GPL8432	Survival analysis
GSE69051	157	Illumina	GPL14951	Survival analysis
GSE4475	123	Affymetrix	GPL96	Survival analysis
GSE11318	203	Affymetrix	GPL570	Survival analysis
GSE34171	91	Affymetrix	GPL570	Survival analysis
GSE58718	242	Illumina	GPL6986	Chromosomal aberration
GSE57277	148	Affymetrix	GPL3720	Chromosomal aberration
GSE34171	180	Affymetrix	GPL6801	Chromosomal aberration

Table 2. Statistics of univariate Cox proportional hazard analysis of the *RTN1* gene in the various datasets. Significant *P* values were bolded.

Dataset	HR[1]	SE	95% CI[2]	*P* value
GSE10846	0.73	0.07	0.64-0.84	**0.000**
GSE31312	0.79	0.07	0.69-0.90	**0.000**
GSE32918/69051	0.78	0.07	0.68-0.90	**0.000**
GSE4475	0.41	0.27	0.24-0.69	**0.000**
GSE34171	0.53	0.17	0.38-0.73	**0.000**
GSE11318	0.69	0.10	0.57-0.83	**0.000**

1. Hazard ratio, 2. Hazard ratio 95% confidence interval

Table 3. Multivariate analysis of the *RTN1* and common prognostic variables in DLBCL (the IPI components). The *RTN1* gene was remained as a one of the most powerful independent prognostic factor. Significant *P* values were bolded.

	GSE10846			
Variable	HR[1]	SE	95% CI[2]	*P* value
RTN1	0.78	0.08	0.67-0.90	**0.000**
Molecular subtype				
GCB-like vs. type 3	0.98	0.30	0.55-1.76	0.96
ABC-like vs. type 3	1.36	0.29	0.78-2.39	0.28
Age (≥60 vs. <60 years)	1.88	0.18	1.32-2.68	**0.000**
Sex (male vs. female)	1.23	0.17	0.89-1.72	0.22
Stage (III/IV vs. I/II)	1.81	0.20	1.23-2.67	**0.000**
NES[3] (≥2 vs. <2)	1.62	0.19	1.12-2.33	**0.01**
ECOG[4] (≥2 vs. <2)	1.50	0.20	1.03-2.20	**0.04**
LDH[5]	1.58	0.20	1.07-2.35	**0.02**
	GSE31312			
RTN1	0.77	0.07	0.66-0.88	**0.000**
Molecular subtype				
GCB-like vs. type 3	0.75	0.30	0.41-1.35	0.33
ABC-like vs. type 3	1.63	0.28	0.95-2.81	0.08
Age (≥60 vs. <60 years)	2.23	0.20	1.51-3.31	**0.000**
Sex (male vs. female)	1.04	0.18	0.73-1.50	0.82
Stage (III/IV vs. I/II)	1.30	0.20	0.88-1.92	0.20
NES (≥2 vs. <2)	1.16	0.34	0.59-2.27	0.67
ECOG (≥2 vs. <2)	2.23	0.20	1.52-3.27	**0.000**
LDH	1.12	0.03	1.06-1.18	**0.000**

1. Hazard ratio, 2. Hazard ratio 95% confidence interval, 3. No. of extranodal sites, 4. ECOG performance status, 5. Lactate dehydrogenase,

Figures

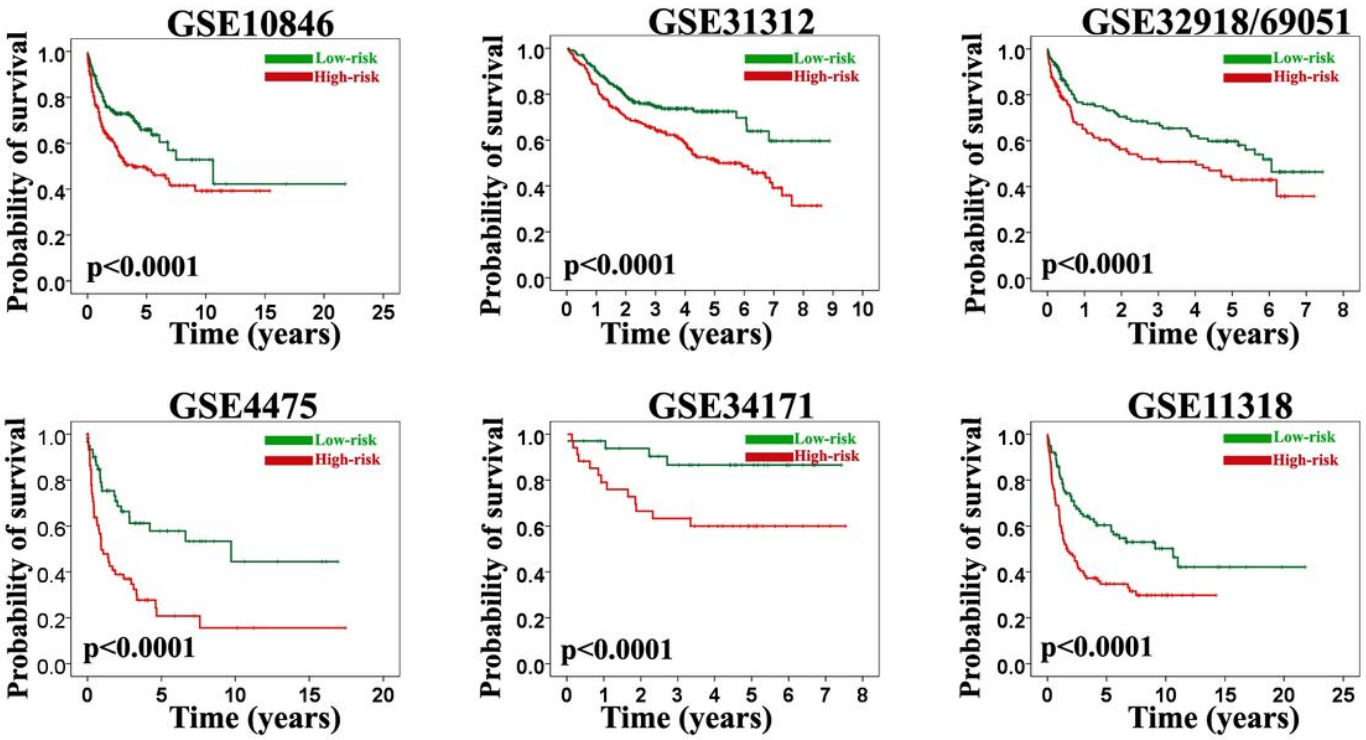

Figure 1

Kaplan-Meier survival analysis of the RTN1 in various gene expression datasets. This gene was found to be significantly associated with the overall survival at a P value <0.0001 in all datasets.

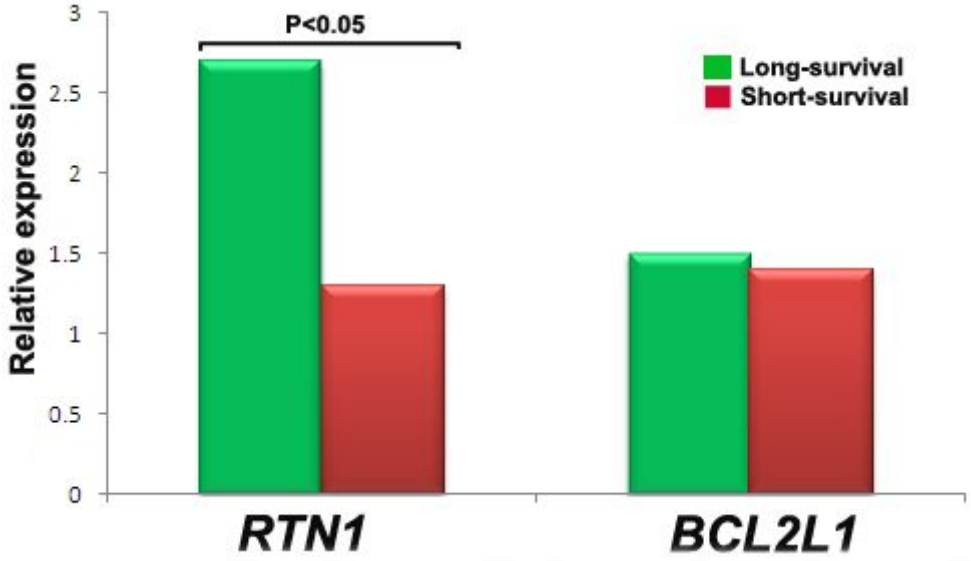

Figure 2

Quantitative real-time PCR (qRT-PCR) analysis of the RTN1 and BCL2L1 in DLBCL patients with the different survival. Gene expression level of RTN1 was significantly greater in long-survival group than

short-survival group (P<0.05). Expression level BCL2L1 wasn't statistically different between two groups (P>0.05).

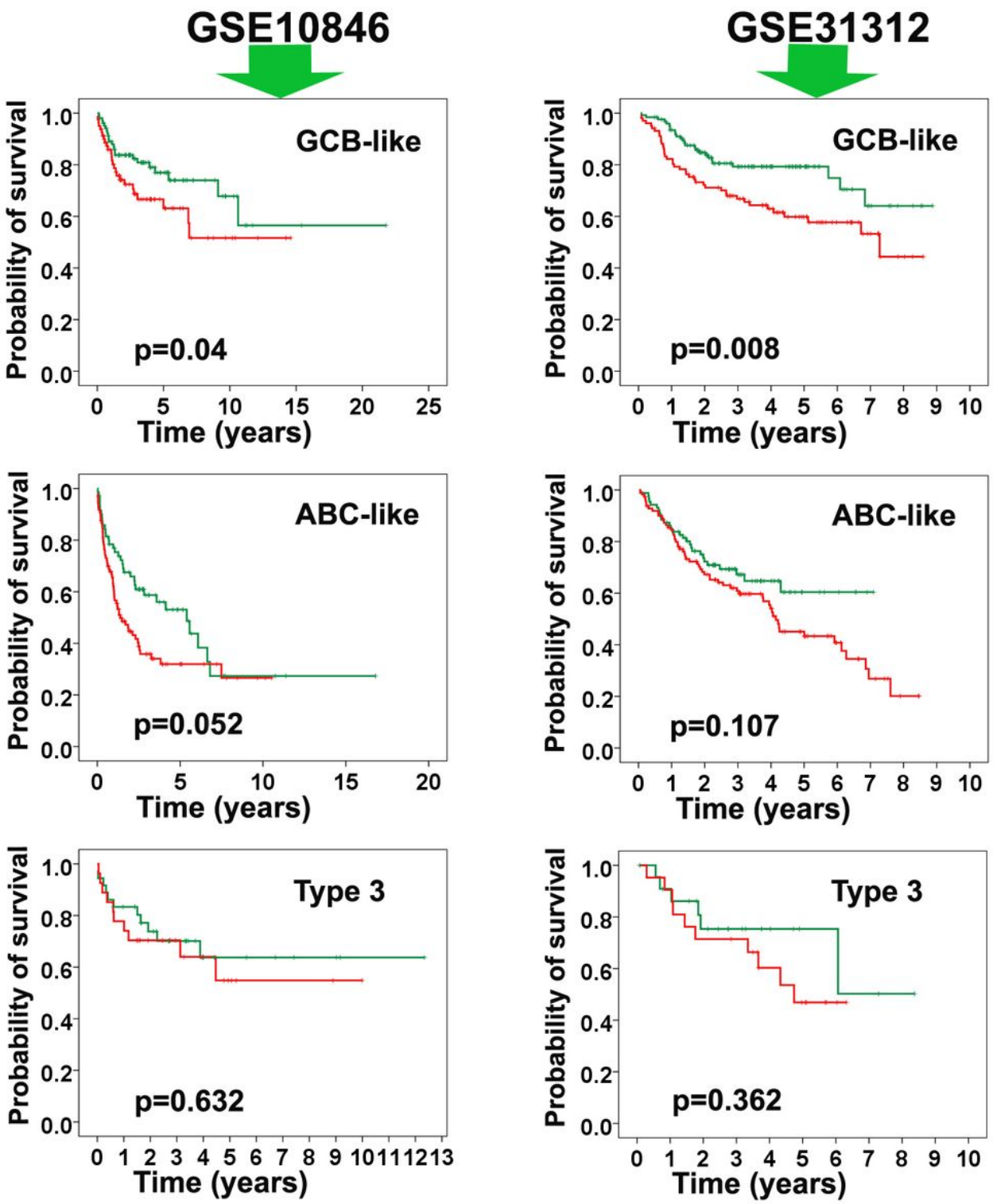

Figure 3

Comparison of expression of our predictor gene (RTN1) in three molecular subtypes of DLBCL (i.e. ABC-like, GCB-like, and type 3). Left and right panels indicate GSE10846 and GSE31312 datasets, respectively. The survival time in GCB-like was significantly different between two risk groups (Ps<0.05).

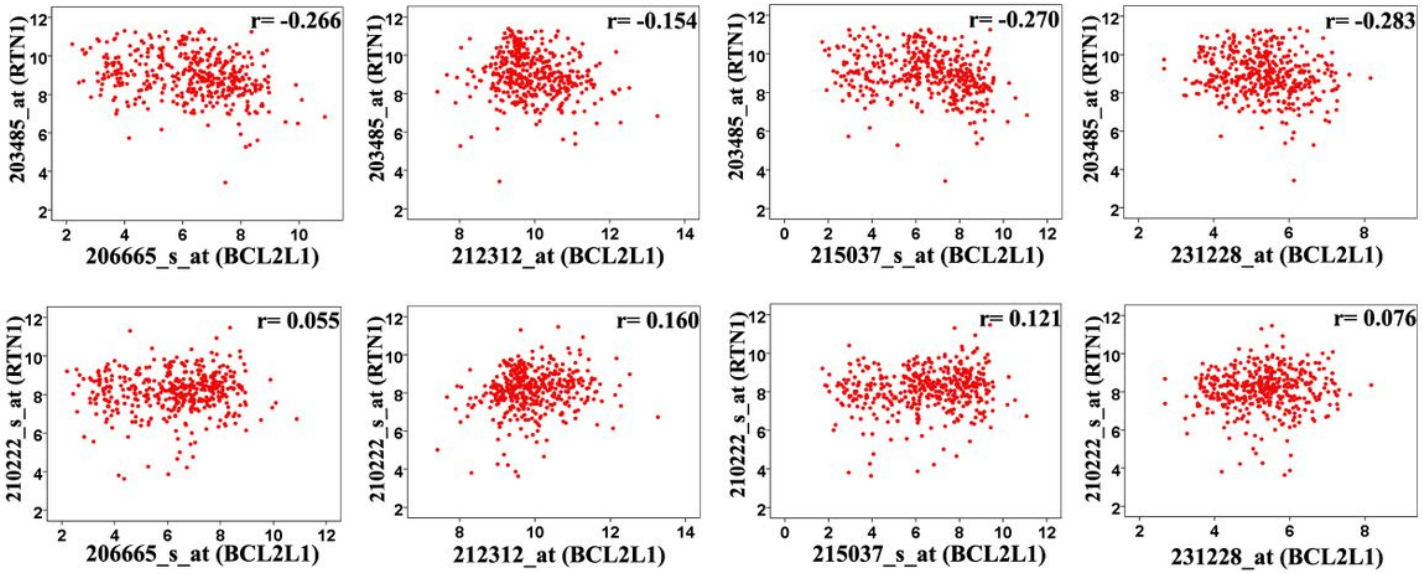

Figure 4

Representing scatterplot depicting the correlation between various BCL2L1 probe-sets (x-axis) and RTN1 probe-sets (y-axis) in GSE10846. The correlations between pairs of probes were poor and inconsistent (negative and positive correlations). Correlation analysis for other datasets was presented in Additional Table 2.

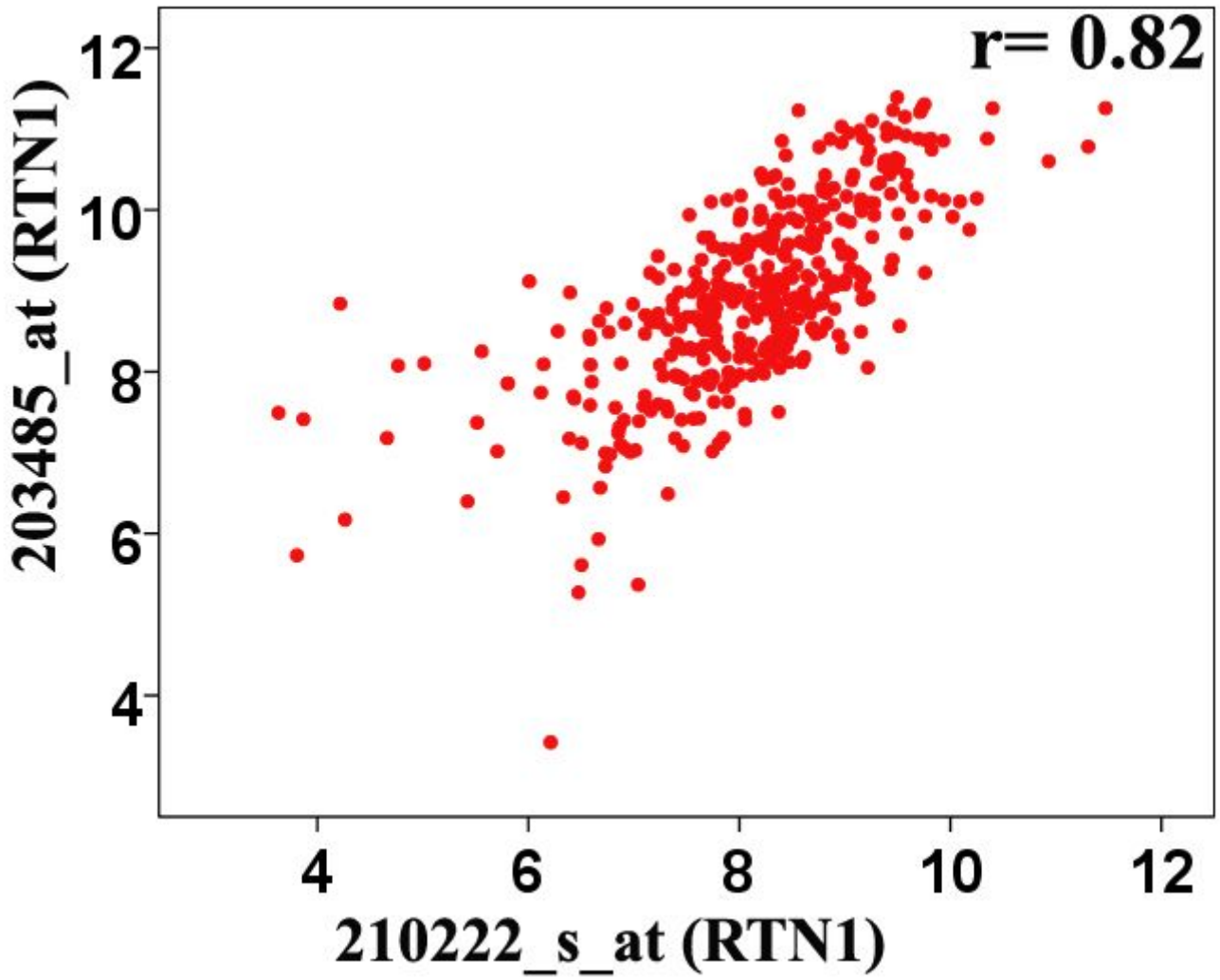

Figure 5

Representing scatterplot depicting the correlation between two RTN1 probe-sets (203485_at and 210222_s_at). Correlation coefficient between two probe-sets was statistically significant (r=0.82) (P<0.05).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementryTables.docx