Identification of B52-dependent Gene Expression Signature and Alternative Splicing Using a *D. melanogaster* B52-null Mutant

Sunwoo Hong,[a] Misun Jung,[†,a] Eunkyung Kim,[†] Dong-ki Lee,[*] and Soyoun Kim[†,*]

Global Research Laboratory for RNAi Medicine, Department of Chemistry and BK21 School of Chemical Materials Science, Sungkyunkwan University, Suwon 440-746, Korea. [*]E-mail: dklee@skku.edu
[†]Department of Biomedical Engineering, Dongguk University, Seoul 100-715, Korea. [*]E-mail: skim@dongguk.edu
Received December 12, 2008, Accepted December 17, 2008

SR proteins are essential splicing regulators and also modulate alternative splicing events, which function both as redundant and substrate-specific manner. The Drosophila B52/SRp55, a member of the SR protein family, is essential for the fly development in vivo, as deletion of B52 gene results in lethality of animals at the second instar larval stage. Identification of the splicing target genes of B52 thus should be crucial for the understanding of the specific developmental role of B52 in vivo. In this study, we performed whole-genome DNA microarray experiments with a B52- knock-out animal. Analysis of the microarray data not only provided the B52-dependent gene expression signature, but also revealed a larval-stage specific, alternative splicing target gene of B52. Our result thus provides a starting point to understand the essential function of B52 at the organismal level.

Key Words: B52/SRp55, DNA microarray, Alternative splicing

Introduction

Gene splicing in higher eukaryotes is regulated by a battery of protein factors that facilitates generation of appropriate splice variants. This spacial and temporal regulation not only increases the complexity of the proteome but is also required to achieve developmental stage and tissue specific gene expression. One of these regulatory protein factors involved in splicing are the SR proteins, which act as both essential splicing factors and alternative splicing regulators that function in the early stages of spliceosome assembly.[1] SR proteins share a similar structure, with RNA recognition motifs (RRM) at the *N*-terminus and have a region rich in Arg-Ser dipeptides (RS domain) at the C-terminus.

B52/SRp55, which was originally identified as a bracketing factor of heat shock puffs in polytene chromosomes,[2] later turned out to be a member of the SR protein family. B52 is an essential splicing factor in *Drosophila melanogaster* both *in vitro* and *in vivo*. Further, it has also been shown to be essential for ftz splicing in splicing-deficient S100 extract in vitro.[3] Also, B52 deficiency in animals is lethal at the second instar larval stage, suggesting the essential, non-redundant function of B52 *in vivo*.[4] Targeted disruption of other SR proteins such as ASF/SF2[5] and SC35[6] in mice also supports the idea of non-overlapping role of SR proteins *in vivo*.

Recently studies in this field have uncovered the specific alternative splicing events regulated by different splicing regulators. Blanchette *et al*.[7] combined RNA interference (RNAi) and DNA microarray techniques to identify splicing target genes of several splicing regulators including B52 in *Drosophila* SL2 cells. Another study also utilized RNAi and microarray to identify 11 target genes of hnRNP L.[8] Recently, splice site-specific DNA microarray analysis of human U2OS cells treated with siRNA targeting SRp55 revealed KSR1, ZAK, and mda7/IL24 as splicing target genes.[9] While these studies are informative in providing a comprehensive list of potential splicing target genes, alternative splicing observed in cell lines might not reflect the physiological importance of splicing regulators at the organismal level. For example, B52 is the predominant splicing factor in *Drosophila* Kc cells, but not in whole embryos.[3]

In vivo studies, mostly carried out in flies, have identified several *in vivo* splicing target genes of SR proteins. B52 was identified as a splicing regulator of dE2F2 in flies.[10] Overexpression studies identified several alternative splicing target genes of B52 and ASF/SF2 in developing fly eye,[11] and a recent study[12] showed that the alternative splicing of eyeless gene is specifically regulated by B52, suggesting that B52 might play a critical role in normal eye development in flies. While these studies have been performed in specific tissue types, they clearly demonstrate the specific role of splicing factors *in vivo*.

To gain an insight into the essential role of B52 in fly embryonic development, we have been studying the B52 knock-out flies.[4] Using genomic SELEX technology, we have previously identified several genes whose alternative splicing pattern was changed in B52-null mutant flies.[1] The expression levels of B52-target genes identified were similar in both the B52 wild-type and mutant animals. In this study, to test whether B52 depletion affects the global gene expression signature, we explored the gene expression changes triggered by the depletion of B52 in fly embryos by using the whole genome DNA microarray. From this experiment, we were able to identify a number of genes whose expression level is affected by B52 depletion. We also report the finding of a novel gene whose alternative splicing is regulated by B52 during the larval development.

Experimental Section

Fly preparations. B52 depleted mutant flies were described previously.[4] In brief, homozygous B52 mutants ($B52^{28}/B52^{28}$)

[a]These authors contributed equally.

were picked 7 days after laying eggs. Homozygous B52 larvae were distinguished from heterogeneous B52 wild type by non-tubby and tubby shapes respectively. About 150 non-tubby and 10 tubby larvae were hand-picked and frozen with TRI reagent (SIGMA, USA) in -70 ^{o}C. Secondary- and third-instar larvae of Oregon R were used to test for the larval stage-dependent gene expression.

RNA isolations. Frozen larvae were homogenized in TRI reagent using Pellet Pestle® Motor (KONTES, USA) and disposable grinder. Total RNAs from larvae were purified according to manufacturer's protocol and treated by DNase I to remove genomic DNAs. The concentration of total RNA prepared were determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, USA) and RNA integrity was verified by 1% agarose gel electrophoresis. The RNA samples for microarray analysis were cleaned up using RNeasy Mini Kit (Qiagen, Germany) and stored at -70 °C before use.

Oligonucleotide microarray, hybridization, image acquisition and data analysis. NimbleGen *Drosophila melanogaster* 1-Plex Array (Roche NimbleGen, USA) was used for gene expression profiling. Sample labeling and hybridization were done as described in manufacturer's protocol. The prepared microarrays were scanned with an Axon GenePix 4000B scanner (Molecular Devices Corp, USA). The data were extracted from the raw images using NimbleScan software (Roche NimbleGen, USA). Relative signal intensities (log_2) for each feature were generated using the Robust Multi-Array Average (RMA) algorithm.

Differentially-Expressed Genes (DEGs) selection. We converted the normalized microarray data of each sample into the signal ratio (B52del/WT) which represents the relative expression level of a target transcript. A difference of more than 2xSD in the average of log2(signal ratio) was considered to be accounting for the up-regulated genes in B52 mutants, while a difference of less than 2xSD was defined as representing a down-regulated gene in B52 mutants. Among the 15,634 fruit fly transcripts present on the NimbleGen oligonucleotide array, 13.3% changed their expression in both samples.

Gene Ontology analysis. GO analysis was performed through Database for Annotation, Visualization, and Integrated Discovery (DAVID) platform. Selected DEGs list was uploaded onto DAVID website (http://niaid.abcc.ncifcrf.gov/home.jsp) and using the total genes of *Drosophila melongaster* as background, enrichment of specific GO term was identified.

cDNA synthesis and PCR reactions. 1 μg of total RNAs per 20 μl reaction were reverse-transcribed with ThermoScript Reverse Transcriptase (Invitrogen, USA). RT reactions were done in 50 ^{o}C for 2 hours and treated with RNase H (Invitrogen) at 37 ^{o}C for 30 minutes. 1 μl of cDNA was used as a template for each PCR. PCR (in 25 μl) was performed as follows: 30 seconds at 95 ^{o}C, 30 seconds at 58 ^{o}C, 30 seconds at 72 ^{o}C for 35 cycles, followed by 72 ^{o}C for 7 minutes. The PCR was performed using primer sets as shown in Supplementary Table 1.

Results and Discussion

Genome-wide gene expression analysis of B52 deletion mutant. To understand the comprehensive gene expression profile in B52 deletion mutant animals, we performed a genome-wide DNA microarray experiment. Total RNAs from B52 mutant and wild type were extracted and the microarray experiment was conducted by using the NimbleGen *Drosophila melanogaster* 1-Plex Array platform, which contains multiple probes of about 15,000 *Drosophila* transcripts. We obtained the expression ratio of each gene and extracted differentially expressed genes (DEGs) in B52 deletion mutant flies. For the selection of DEGs, we calculated standard deviation of gene expression ratio and defined genes which showed higher or lower gene expression ratio than 2 S.D. as DEG. As a result, we identified 278 up-regulated and 854 down-regulated genes in B52 deletion mutant flies.

To characterize these DEGs, we performed gene ontology analysis. First we analyzed up-regulated genes and found that only two terms were significantly enriched ('polysaccharide metabolic process' and 'chitin metabolic process', $p < 0.001$). For down-regulated genes, we found that diverse gene ontology terms were enriched (Fig. 1). For biological process category, ontology analysis reported significantly enriched terms such as 'spermatogenesis', 'microtubule-based process', and 'ATP biosynthetic process'. We also found that cellular component analysis showed enrichment of 'microtubule cytoskeleton' and 'mitochondrion'. These results suggest that the impairment of B52 function affects expression of genes engaged in specific cell lineage differentiation and inner cellular structure organization. In particular, we found reduced expression of genes which function in crucial processes for sustaining cell viability. For example, some downregulated genes were found

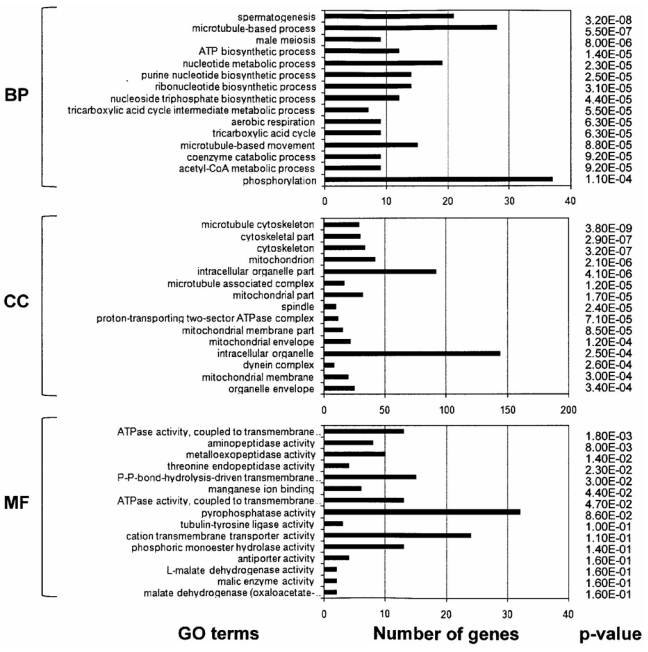

Figure 1. Gene ontology analysis of down regulated genes in B52 deletion mutant. List of 854 down regulated genes in B52 deletion mutant analyzed by web-based gene ontology tool. Enriched GO terms were sorted by reported p-values and 15 most enriched terms of three categories [Biological process (top panel), Cellular component (middle panel), and Molecular function (bottom panel)] are presented.

to be involved in synthetic process of ATP which is a universally important coenzyme and enzyme regulator. In addition, genes involved in tricarboxylic acid cycle were also found as B52-dependent genes. These results indicate that the impairment of B52 causes deleterious effects on the organism and may explain the lethality of B52 deletion in animal development.

B52 depletion affects alternative splicing of CG32548 gene at the organismal level. Multiple probes for individual gene in NimbleGen array format allowed us to examine the differential expression of individual exons within a specific gene. Each probe set of NimbleGene *Drosophila* expression microarray is composed of 12 different oligonucleotide probes designed for targeting a single transcript. Some of these probes are designed for targeting common exons and the others for transcript-specific exons. We inspected the array data from B52 null mutants and found that, the expression of transcript variants of a gene named CG32548, was differentially regulated in B52 null mutants. CG32548 gene has two alternatively spliced transcript variants (Fig. 2A). Microarray data revealed that CG32548-RB transcript, which consists of exon 1 and exon 3, was up-regulated (Fig. 2B). In contrast, CG32548-RC, another transcript variant which consists of exon 2 and exon 3, was down-regulated in B52 null mutants (Fig. 2B). CG32548 was previously identified as one of the conserved polyglutamine proteins in *Drosophila* species with unknown function.[13]

To validate the microarray data, we designed the primer sets specific for each exon and transcript variant as shown in Figure 2A. RT-PCR analysis was performed to measure the expression of each CG32548 exon (Fig. 3, panels Exon 1, Exon 2 and Exon 3) in wild-type (Fig.3. lane 1) and B52-null mutant (Fig.3, lane 2). As shown in Fig. 3, RT-PCR data confirmed the expression patterns of each exon from microarray analysis. Whereas the expression of exon 3, a common exon of CG32548, was not significantly changed, the expression of CG32548-RB specific exon 1 was up-regulated and that of CG32548-RC specific exon 2 was down-regulated by the loss of B52.

As previously observed, the homozygous B52-null mutants used for this study showed growth arrest and died in 2nd instar

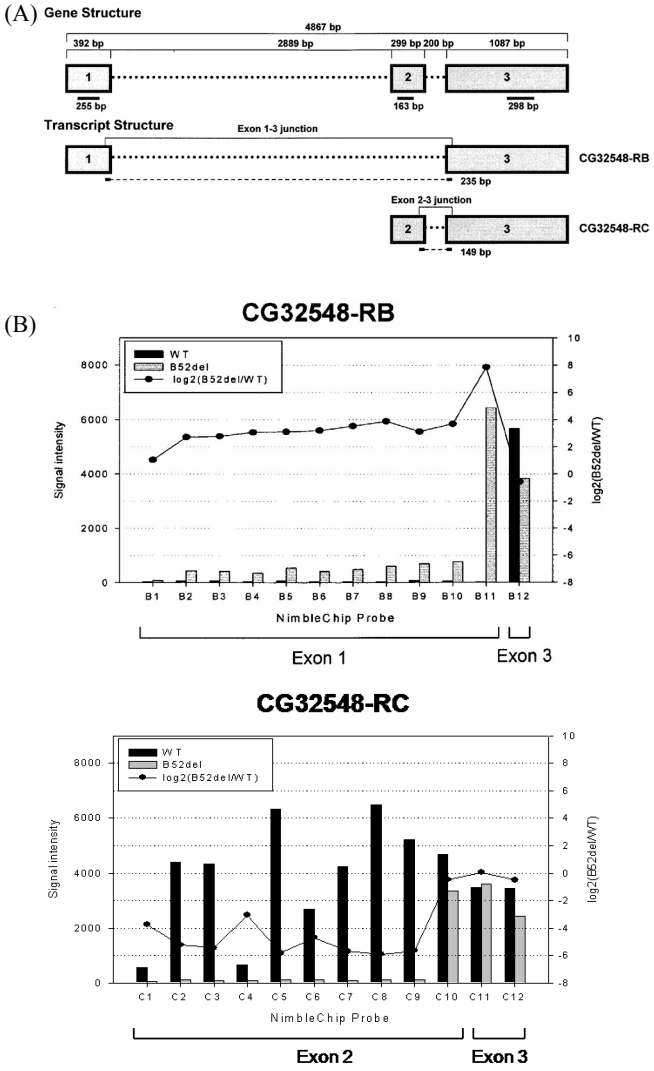

Figure 2. Expression level of CG32548 in microarray analysis and primer design for RT-PCR validation. (A) Primer design for CG32548. Target site and PCR product of each exon-specific primer set are represented on the gene structure (upper panel). Target site and PCR product of each transcript variant-specific primer set are indicated on the transcript structure (lower panel). The same reverse primer was used for the RT-PCR analysis of CG32548-RB and RC. (B) Expression level of two CG32548 transcript variants in Nimble Gen expression microarray analysis. Expression level of each transcript variant is represented by the signal intensity (left Y-axis) of each probe for target transcript variant. Fold change of each transcript expression is indicated by the signal ratio[log2-(B52del/ WT)] (right Y-axis). Target exon of each probe is shown underneath the graph.

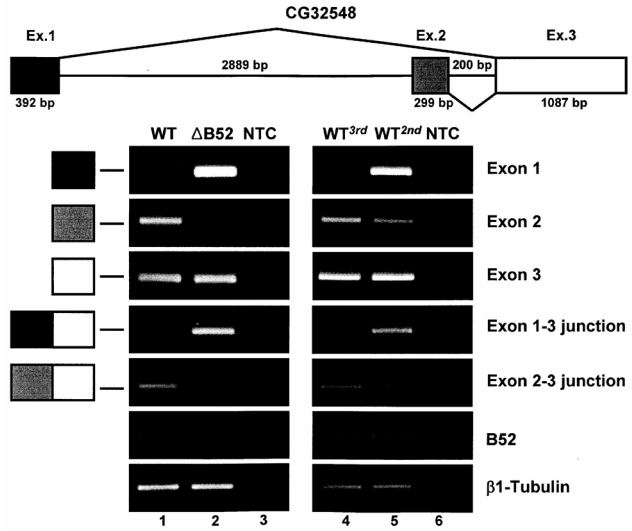

Figure 3. Changes in expression of CG32548 transcript variants under the loss of B52 and during the larval life of *D. melanogaster*. RT-PCR of CG32548 was performed with primer sets specific for each exon or transcript variant as shown in Fig. 2A. The genomic organization and splicing scheme of CG32548 gene are summarized in the top panel. Three exons are represented by black, gray, and white boxes, respectively, and targets of each RT-PCR are also represented by these exon-specific boxes in the left panel. To test the effect of B52 deletion in CG32548 expression, wild-type (tubby 3rd instar larvae of $B52^{28}/B52^{28}$, WT, lane 1) and B52-null mutant (non-tubby 2nd instar larvae of $B52^{28}/B52^{28}$, ΔB52, lane 2) were used. 3rd instar larvae (WT^{3rd}, lane 4) and 2nd instar larvae (WT^{2nd}, lane 5) of Oregon R were used to analyze the pattern of CG32548 expression during the larval stage. β1-Tubulin was used as an internal control and NTC (No Template Control, lane 3 and 6) was used as a negative control.

larval stage.[4] To confirm whether the expression patterns of CG32548 transcript variants was caused by the lack of B52 splicing activity or by larval growth arrest, we performed RT-PCR analysis on total RNA from the wild-type larvae (Oregon R) collected in two different larval stages. CG32548-RB was up-regulated in 2nd instar larvae, and then down-regulated in 3rd instar larvae (Fig. 3, panel Exon 1-3 junction, lanes 4~5). Considering that CG32548-RB was also up-regulated at the 2nd instar larval stage in the B52-null mutant although the mutant cannot survive beyond this stage, these results suggest that the loss of B52 affected the decrease of CG32548-RB transcript level during the transition stage between 2nd and 3rd instar larval stage. In contrast, CG32548-RC transcript, which was down-regulated in B52 null mutant, was consistently expressed during the larval life of wild-type animals (Fig. 3, panel Exon 2-3 junction, lanes 4 and 5). This result clearly demonstrates that the loss of B52 affects the exon 2 usage of CG32548 gene during the *Drosophila* larval development, and thus reduces the expression of CG32548-RC transcript variant in 3rd instar larval stage.

In this work, we reported that the expression of CG32548 transcript variants during *Drosophila* larval development, are different in B52 null mutant as compared to that of the wild-type larvae. In line with our previous report,[1] the current study provides additional evidences supporting the idea that B52 regulates the alternative splicing of genes during the larval development. In the present study, CG32548-RB was expressed differentially during larval stages, while CG32548-RC was expressed consistently. These results strongly suggest that the differential expression of CG32548 transcript variants is associated with *Drosophila* larval development. Furthermore, the loss of B52 interfered with the decrease of CG32548-RB expression during the transition between 2nd and 3rd instar larval stages and induced the decrease of CG32548-RC in 3rd instar larval stage. Although further studies will be needed to verify the correlation between the amount of CG32548 transcript variants and their function during larval stages, our data brings it in focus as one of the main targets of B52 splicing mechanism. Additionally the B52 regulated temporal expression of CG2548 splice variants in Drosophila larval development is intriguing as B52 deficiency has been reported to be lethal at the 2nd instar larval stage.

Altogether, our data shows that loss of B52 affects both the global gene expression pattern and alternative splicing, and thus explains the role of B52 in Drosophila developmental process.

Acknowledgments. S. K. was supported by the National Research Laboratory Program (MOST). D.-k. L was supported by Global Research Laboratory program by KICOS (grant 2008-00582), the Korea Ministry of Environment and the POSTECH BSRI fund.

Supporting Information. Supplementary Table 1 mentioned in the text is available at the bkcs website (*http://www.kcsnet.or.kr/bkcs*).

References

1. Kim, S.; Shi, H.; Lee, D. K.; Lis, J. T. *Nucleic Acids Res.* **2003**, *31*, 1955.
2. Champlin, D. T.; Lis, J. T. *Mol. Biol. Cell* **1994**, *5*, 71.
3. Hoffman, B. E.; Lis, J. T. *Mol. Cell Biol.* **2000**, *20*, 181.
4. Ring, H. Z.; Lis, J. T. *Mol. Cell Biol.* **1994**, *14*, 7499-506.
5. Wang, J.; Takagaki, Y.; Manley, J. L. *Genes Dev.* **1996**, *10*, 2588.
6. Xiao, R.; Sun, Y.; Ding, J. H.; Lin, S.; Rose, D. W.; Rosenfeld, M. G.; Fu, X. D.; Li, X. *Mol. Cell Biol.* **2007**, *27*, 5393.
7. Blanchette, M.; Green, R. E.; Brenner, S. E.; Rio, D. C. *Genes Dev.* **2005**, *19*, 1306.
8. Hung, L. H.; Heiner, M.; Hui, J.; Schreiner, S.; Benes, V.; Bindereif, A. *RNA* **2008**, *14*, 284.
9. Filippov, V.; Schmidt, E. L.; Filippova, M.; Duerksen-Hughes, P. J. *Gene* **2008**, *420*, 34.
10. Rasheva, V. I.; Knight, D.; Bozko, P.; Marsh, K.; Frolov, M. V. *Mol. Cell Biol.* **2006**, *26*, 3468.
11. Gabut, M.; Dejardin, J.; Tazi, J.; Soret, J. *Mol. Cell Biol.* **2007**, *27*, 3087.
12. Fic, W.; Juge, F.; Soret, J.; Tazi, J. *PLoS ONE* **2007**, *2*, e253.
13. Bettencourt, B. R.; Hogan, C. C.; Nimali, M. *J. Biosci.* **2007**, *32*, 537.