

# Differentially regulated splice variants and systems biology analysis of Kaposi's sarcoma-associated herpesvirus-infected lymphatic endothelial cells

Ting-Yu Chang<sup>1</sup>, Yu-Hsuan Wu<sup>1</sup>, Cheng-Chung Cheng<sup>2</sup> and Hsei-Wei Wang<sup>1,3,4,\*</sup>

<sup>1</sup>Institute of Microbiology and Immunology, National Yang-Ming University, 11221, <sup>2</sup>Division of Cardiology, Department of Medicine, Tri-Service General Hospital, 11490 <sup>3</sup>Veteran General Hospital-Yang Ming Genome Research Center, National Yang-Ming University, 11221 and <sup>4</sup>Department of Education and Research, Taipei City Hospital, 10341, Taipei, Taiwan

Received January 18, 2011; Revised and Accepted May 6, 2011

## ABSTRACT

**Alternative RNA splicing greatly increases proteome diversity, and the possibility of studying genome-wide alternative splicing (AS) events becomes available with the advent of high-throughput genomics tools devoted to this issue. Kaposi's sarcoma associated herpesvirus (KSHV) is the etiological agent of KS, a tumor of lymphatic endothelial cell (LEC) lineage, but little is known about the AS variations induced by KSHV. We analyzed KSHV-controlled AS using high-density microarrays capable of detecting all exons in the human genome. Splicing variants and altered exon-intron usage in infected LEC were found, and these correlated with protein domain modification. The different 3'-UTR used in new transcripts also help isoforms to escape microRNA-mediated surveillance. Exome-level analysis further revealed information that cannot be disclosed using classical gene-level profiling: a significant exon usage difference existed between LEC and CD34<sup>+</sup> precursor cells, and KSHV infection resulted in LEC-to-precursor, dedifferentiation-like exon level reprogramming. Our results demonstrate the application of exon arrays in systems biology research, and suggest the regulatory effects of AS in endothelial cells are far more complex than previously observed. This extra layer of molecular diversity helps to account for various aspects of endothelial biology, KSHV life cycle and disease pathogenesis that until now have been unexplored.**

## INTRODUCTION

Alternative pre-mRNA processing greatly increases the complexity of eukaryotic transcriptomes and hence the

diversity of proteome. Such processing may thereby contribute to species-specific and tissue-specific functions (1). The roles of spliced RNA isoforms in tumorigenesis and development have long been documented, in this context, genome-wide analysis of alternative splicing (AS) events have been impeded by difficulties in systematically identifying and validating transcript isoforms. Initially, most gene expression arrays were designed to target their probes to only a specific part of genes and thus were unable to identify specific isoform changes. Methods of studying alternative splicing using microarrays and other related bioinformatics tools have appeared only recently, and this advance in technology has greatly aided our understanding of AS events. Microarrays using exon-junction probes or exon-specific probes were used first for genome-wide analysis in yeast (2) and then to detect alternative splicing differences in mammalian tissues; they have been found to have a validity rates ranging from ~50% (3) to 70–85% (4,5). Genome-wide analysis of AS events also led to the new discoveries; for example, brain-specific splicing that was researched using microarrays has revealed how a neuronal splicing factor, Nova, shapes the synapse (6,7).

Kaposi's sarcoma (KS) is the most common cancer in individuals with HIV or AIDS and is one of the most common cancers in patients who have had their immune system suppressed (8). The epidemiology of KS, in particular its geographical distribution as well as its prevalence in gay men, indicated that there was an infectious agent closely related to this neoplasm. In 1994, a new type of human  $\gamma$ -herpesvirus, Kaposi's sarcoma-associated herpesvirus (KSHV), also known as human herpesvirus 8 (HHV8), was identified. This was done by representational difference analysis whereby KS tissue was compared with normal tissue from the same individual (9). Genomics tools have been widely applied to KSHV research and have unmasked novel KS tumorigenesis mechanisms that cannot be readily disclosed by classical hypothesis-driven studies (10–12). For example, transcriptome analysis

\*To whom correspondence should be addressed. Tel: +886 2 2826 7109; Fax: +886 2 2821 2880; Email: hwwang@ym.edu.tw

showed that the expression profile of KS lesions most closely resembles that of lymphatic endothelial cells (LEC), and KSHV infection induces unexpected LEC-to-blood vessel endothelial cell (BEC) transcriptome reprogramming (13). Many other novel discoveries have also been obtained by array-based screening using KSHV-infected LECs, BECs and Human umbilical vein endothelial cells (HUVECs) (14). However, little is known at present about whether KSHV is able to influence the global RNA splicing pattern in infected cells, as well as the potential outcomes of AS on endothelial cells (ECs) activities.

Another approach to understand how KSHV may contribute to sarcomagenesis is to organize KSHV-regulated genes or AS events into functional modules or signaling pathways. Although gene expression profiling can reveal differentially expressed genes, the challenge remains the assignment of the biological significance of these genes within the complete biological system. Systems biology helps to interpret the interactions between the components of biological systems (such as KSHV-infected ECs), and how these interactions give rise to the functional behavior of the system. Disrupted crosstalk among biological modules can be a hallmark of cancer (15). Functional network analysis has been widely applied to analyze the transcriptomes of complex biological systems such as cancer and stem cells (16,17). However, there has been little research that has used genetic network tools to reveal the system changes caused by KSHV (14). As yet there has been no global genetic network/systems biology picture published for KSHV-infected cells, particularly involving virus-induced RNA isoforms. For example, how KSHV-regulated AS events may produce profound changes in the predicted protein domain/motif composition that ought to affect protein function and interaction, such as changes in the genetic networking, and how exon inclusion/exclusion may modify the predicted microRNA binding site composition of transcripts, remains unelucidated.

Here, we investigate KSHV-regulated endothelial transcripts across their entire length using the Affymetrix GeneChip Human Exon 1.0 ST Array, which can detect splicing differences between various types of samples (18,19).

## MATERIALS AND METHODS

### Cell culture and KSHV infection

Human dermal LECs were purchased from Promo Cell (Heidelberg, Germany; product No. C-12217) and maintained for up to 10 passages in endothelial cell growth medium MV (Promo Cell C-22020) in a humidified incubator supplemented with 5% CO<sub>2</sub>. KSHV BAC36 was harvested as described previously (20). Briefly, BCBL-1 cells bearing with KSHV BAC36 were maintained in RPMI-1640 supplemented with 10% fetal calf serum (FCS) and 150 µg/ml hygromycin (Sigma-Aldrich, St. Louis, MO USA). Recombinant Vero cells harboring JSC-1-derived rKSHV.219 viruses, which express green fluorescent protein (GFP) during infection but red fluorescent protein during lytic cycle (21), were maintained in Dulbecco's Modified Eagle Medium (DMEM)

supplemented with 10% FCS and 5 µg/ml puromycin (Calbiochem, San Diego, CA, USA). KSHV BAC36 was induced into the lytic cycle by adding 20 ng/ml 12-*O*-tetradecanoyl phorbol-13-acetate (TPA; Sigma-Aldrich, St. Louis, MO, USA), and 1.25 mM sodium butyrate and culture continued for 5 days. rKSHV.219 was induced into lytic cycle by infecting with the baculovirus harboring KSHV/RTA (BacK50) and following by a treatment of 1.25 mM sodium butyrate as described previously (21). Culture supernatant was then filtered through a 0.45 µm filter and concentrated by centrifugation at 18 000g. Virus titer was determined by infecting 293 T cells with serially diluted culture supernatant and counting the percentage of GFP positive cells. For KSHV infection, a >10 MOI viral concentrate was overlaid onto LECs and centrifuged at 3000g for 30 min at room temperature to enhance the infection rate. The cells were then incubated at 37°C for 2 h and washed with phosphate buffered saline (PBS) followed by a change to fresh medium. RNA was harvested two days post-infection using a Qiagen<sup>TM</sup> (Valencia, CA, USA) RNeasy<sup>®</sup> kit by following the manufacturer's suggestions.

### Microarray hybridization

Total RNA from KSHV infected LECs or mock infections with biological replicates were checked for RNA integrity by Bioanalyzer (Agilent, Santa Clara, CA, USA). Two µg of total RNA served as the starting material and mRNA was purified using a RiboMinus human transcriptome isolation kit (Invitrogen, Carlsbad, CA, USA). cDNA was then synthesized, labeled and fragmented using a GeneChip WT cDNA synthesis kit and a WT terminal labeling kit (Affymetrix, Santa Clara, CA, USA). Hybridization and array scanning were carried out according to the manufacturer's instructions. The data has been deposited in NCBI's Gene Expression Omnibus (GEO) database and has the accession number of GSE26341.

### Additional microarray data

Raw CEL files of Affymetrix Human Exon 1.0 ST arrays for CD34<sup>+</sup> hematopoietic stem/precursor cells and peripheral blood mononuclear cells (PBMCs) were downloaded from NCBI GEO website (series accession number GSE15207). CEL files of Affymetrix Human Genome U133 Plus 2.0 arrays for CD34<sup>+</sup> cells and PBMC were from our previous publication (22), and KSHV-infected LECs (KLECs) and uninfected LECs were from GEO (series accession number GSE16354).

### Gene expression analysis

For the gene expression analysis based on the Affymetrix Human Genome U133 Plus 2.0 platform, we performed RMA normalization and summarization using the 'affy' package in the Bioconductor suite as described previously (13). We then identified differential gene expression between KLEC and LEC by the 'limma' package and calculated the positive false discovery rate (pFDR) by the 'q-value' package in the Bioconductor suite (13). Principal components analysis (PCA) and multidimensional scaling (MDS) plots were processed by the Partek Genomics Suite 6.5 (Partek, St. Louis, MO, USA).

For gene expression analysis based on the exon array data, we used AltAnalyze (23) with Ensembl build 54, human genome build 19 (GRCh37) and Affymetrix annotation files. Analysis was carried out at the default settings. Briefly, RMA summarized probe set intensities and DABG ('detection above background') *P*-values were calculated using the Affymetrix Power Tools (APT), which was automatically called from AltAnalyze. Probe sets uniquely mapped to the each Ensembl gene accession (ENSG) were considered constitutive and were used for calculating gene expression values. For each ENSG, gene expression was determined by the mean expression intensities of all constitutive probe sets. The Bioconductor R package 'q-value' was used to calculate pFDR from the raw *P*-values for KLEC and LEC data. The complete gene expression results are presented in File S1 in Supplementary Data.

### Alternative splicing analysis

In order to identify alternative expressed exons, we used both AltAnalyze and easyExon (24) to conduct the AS analysis. Exon array raw data (CEL files) were normalized and summarized using Affymetrix Power Tools (APT) with the RMA-sketch algorithm. AltAnalyze was run using the default parameters. The default signal filtration only retained probe sets with an intensity of more than 70 and a DABG *P*-value of smaller than 0.05. The same database (Ensembl release 54) that had been used for gene expression analysis was used for mapping the probe sets to annotated exons. Splicing-index (SI) and MiDAS statistical algorithms were used for AS identification (24). Probe sets with a MiDAS *P*-value smaller than 0.05 were considered significant for exon inclusion/exclusion. The complete AS analysis results are presented in File S2 in Supplementary Data. The degree of AS was generally represented by normalized intensity (NI), which means exon level intensity divided by gene level intensity. NI, rather than probe set intensity, was used for plotting exon level heat maps and PCA/MDS figures in order to obtain a better representation of the AS degrees.

### Alternative protein domains and microRNA binding sites analysis

To identify protein domains and motifs potentially modified by AS, a database containing aligning and non-aligning proteins for all probe sets was built using AltAnalyze. AltAnalyze aligns probe sets to protein domains using 'inferred comparison' (default option) and reports the domains to which the differentially expressed probe sets align (23,25).

In addition to protein annotations, putative microRNA binding sites predicted by PicTar (<http://pictar.mdc-berlin.de/>), miRanda (26), miRbase (27) and TargetScan (<http://www.targetscan.org/>) within probe selection region were also pre-aligned and stored in an AltAnalyze database. Differentially expressed probe sets aligned to protein domain or microRNA binding sites were subjected to gene set enrichment analysis. Gene sets with three or more associated genes, whose domains or motifs had *z* scores >2, microRNA binding sites had *z* scores >2

and had a Benjamini-Hochberg *P* < 0.05, were reported. The AltAnalyze export file for graphic presentation of domain architecture by Domaingraph is provided as File S6 in Supplementary Data.

### Gene ontology, pathway and network analysis

The filtered AS genes were subjected to gene ontology enrichment analysis using AltAnalyze bundled module GO-Elite ([http://www.genmapp.org/go\\_elite](http://www.genmapp.org/go_elite)). GO-Elite implements an over-representation statistical inference that can identify significantly enriched gene ontology (GO) categories with AS exons. GO terms with a *z*-score >2, a permutation of *P* < 0.01 and three or more regulated genes for each GO term were reported as significant. Canonical pathway and gene interaction network analysis were conducted with Ingenuity Pathway Analysis web tool (<http://www.ingenuity.com/>).

### Confirmation of alternative spliced mRNA and differential gene expression

To confirm the AS events identified by the MiDAS algorithm, we designed primer pairs flanking the AS exon and quantitatively analyzed the RT-PCR amplified fragments produced by these primers. The forward primer sequence for CLDND1 AS detection is 5'-gctggcggagcaggaggat-3' and the reverse primer is 5'-gtcctttctggtgggctataccaa-3'. One microgram total RNA was used as starting material for the RT-PCR. The cDNA was synthesized by Superscript III (Invitrogen Carlsbad, CA, USA) using random hexamer. The subsequent PCR reaction was carried out by GoTaq Master Mixes (Promega, Madison, WI, USA) for 30 cycles at annealing temperature of 50°C. The amplified DNA fragments were separated on an ethidium bromide stained agarose gel and quantified by ImageQuant software (GE Healthcare, Piscataway, NJ, USA). Quantitative PCR detection of AS and differential gene expression was performed using Maxima<sup>®</sup> SYBR Green qPCR Master Mix (Fermentas, Glen Burnie, MD, USA) for 40 cycles at annealing temperature of 58°C with StepOne<sup>®</sup> Plus Real-Time PCR system (Applied Biosystems, Carlsbad, CA, USA). Primer sequences for qPCR detection of EGFR, CASP9, BAIAP2L1, DVL2, RAPGEF5, MLPH and CTCL1 are listed in Supplementary Table S1.

### Knocking down RAPGEF5 by RNA interference

For the lentivirus-expressed shRNA against RAPGEF5 (NM\_012294), two 21-nt sequences corresponding coding sequences 1145–1165 (5'-gctaagaagtatcaaggcaaa-3') and 1376–1396 (5'-cgtcacactgtagatgaatat-3') inserted into pLKO.1 were obtained from the National RNAi Core Facility at Academia Sinica, Taiwan (<http://rna.genmed.sinica.edu.tw/en/>).

### Transwell cell migration assay

Cell migration ability was evaluated using Costar Transwell<sup>®</sup> Polycarbonate Permeable Supports (Corning, NY, USA). Briefly, 5 × 10<sup>4</sup> cells in 400 μl of culture medium were overlaid on the upper chamber of the

Transwell® device, while 600 µl of medium supplemented with 10 ng/ml human VEGF (R&D Systems, Minneapolis, MN, USA) was added to the lower chamber. The pore size of the polycarbonate membrane situated between these two chambers is 8 µm. After 6 h of 37°C incubation, the membrane was fixed in 4% paraformaldehyde for 20 min at room temperature and then stained with Hoechst 33342 solution (Sigma-Aldrich, St. Louis, MO, USA) for 30 min. Migrated cells on the bottom side of the membrane were counted using a fluorescent microscope.

## RESULTS

### Global identification of alternative splicing events in KSHV-infected LECs

To uncover the specific traits of the exome expression induced by KSHV, we used the Affymetrix exon arrays to determine the exon expression profiles of infected and uninfected primary LEC. A BCBL-1 cell-derived recombinant KSHV that expresses green fluorescent protein (GFP) (20) was used to infect a pure population of LEC. Uninfected LECs showed a characteristic contact-inhibited cobble stone appearance, but the infected LECs (KSHV-LECs) became elongated and spindle shaped (Figure 1A). KSHV latent nuclear antigen (LANA) was expressed in the infected cells and showed a typical nuclear stippling staining pattern (Figure 1A, right panel). Total RNA was collected 48 h after KSHV infection. Pilot quality control PCR test proved that the transcripts of KSHV latent genes were present and intact in the KSHV-LECs (Figure 1A). KSHV-LEC RNA was then subjected to hybridization to Human Exon 1.0 ST GeneChip arrays that consist of 5 362 207 different oligonucleotides, corresponding to more than 1 400 000 probe sets (PS). Exons or parts of exons within a gene are represented on the microarray by individual PS, and were considered discrete units for our analysis of the transcript isoform-processing differences. Of note, indirect paracrine effects may also contribute in observed AS events as we reached ~80–90% infectivity but using the whole LEC culture (composed of infected and uninfected cells) for exon array analysis.

Initially, a global analysis was carried out at gene level (i.e. individual transcript level) by summarizing all PS signal values from a given Ensembl gene into one unique value. A PCA based on genes differentially expressed between sample groups (pFDR < 0.1) showed that the KSHV-infected LECs and parental LECs segregated in a different spatial locations, substantiating that a very specific gene expression profile was stimulated by KSHV infection, as has been previously reported (13) (Figure 1B, left panel).

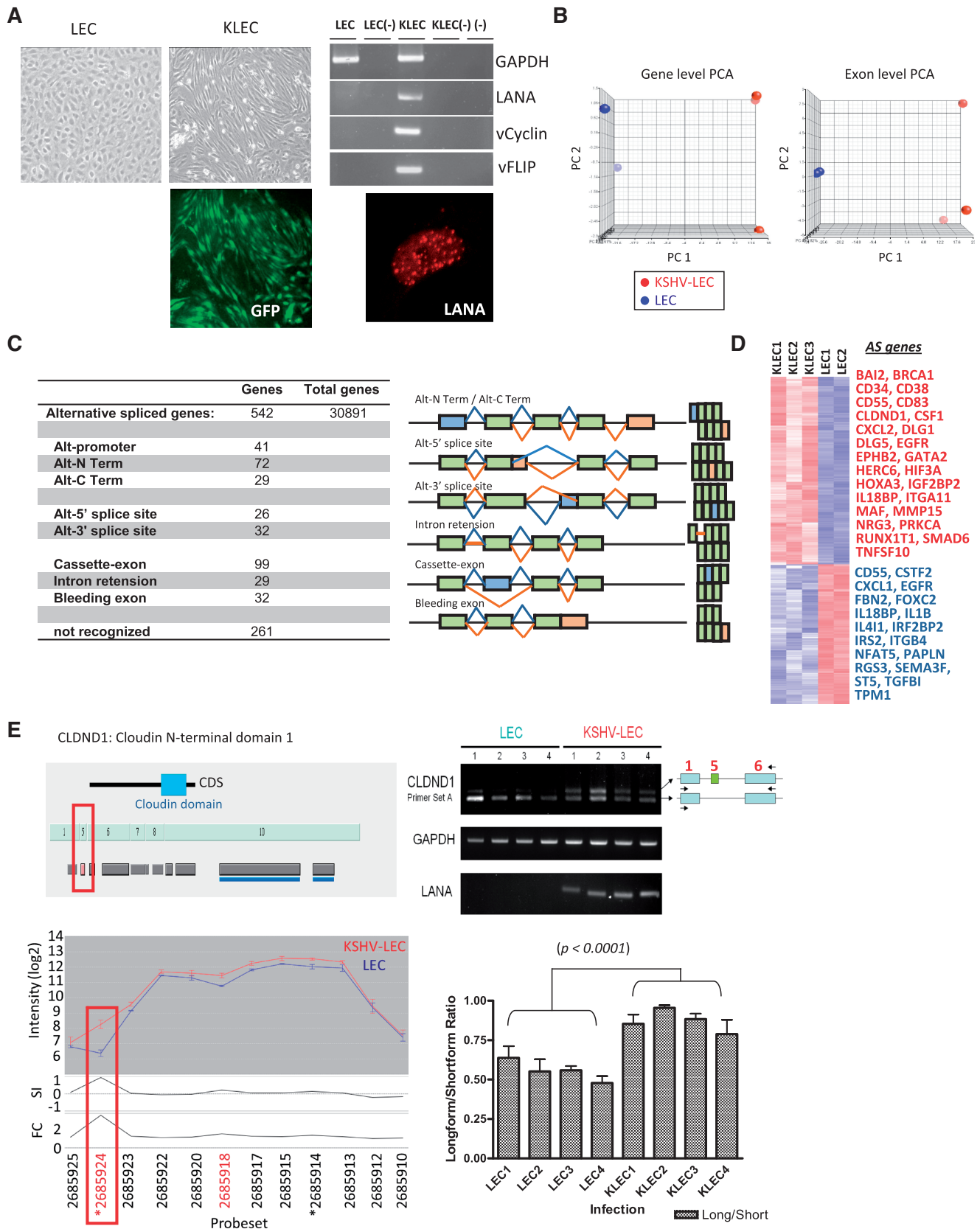
At the exome level, probe select regions (PSRs, ~exons) were recognized as alternatively spliced exons using two free open source applications: AltAnalyze (23) and easyExon (24). Both applications utilize the MiDAS algorithm, which calculates a splicing index (SI) first and then performs statistical inference on the SI to obtain a MiDAS *P*-value. To identify alternative exons, AltAnalyze was run using the default parameters. The CEL files of three

KSHV infected LECs (KLECs) and two mock infection controls (LECs) were summarized by RMA and filtered by the SI and MiDAS statistical algorithms (*P*-value < 0.05). The corresponding genomic library files were based on Ensembl Release 54 (GRCh37/hg19). Probe sets with large cross-hybridization scores were removed to ensure the fidelity of the microarray results. To identify high-confidence alternative exons, only regulated probe sets overlapping with exons in annotated mRNAs (i.e. categorized as ‘core’ PSRs by Affymetrix) were used for the further analyses. In the end, a total of 805 core PSRs were obtained by filtering and thus identified as differentially expressed on KSHV infection (MiDAS *P* < 0.05; see File S2 in Supplementary Data). The PCA plot produced using these normalized intensities of PSRs illustrates the segregation of the infected and uninfected LEC at the isoform levels (Figure 1B, right panel) and shows clearly that the exome pattern in KSHV-infected LEC is also quite different from that of uninfected cells.

### AS is a key feature of KSHV infection

Having defined the entire expression map of the human exome in KSHV-LEC, we examined in depth the exons and genes that might be involved in alternative splicing events. The above core PSRs belong to 542 alternatively spliced genes and the nature of the isoform changes at the transcript level is illustrated in Figure 1C. These categories were automatically classified on the basis of the positions of the variable probe sets, followed by manual examination based on visualization of the entire transcript. A significant number (261/542, or 48.2%) of genes showed very complex patterns of isoform variation that were difficult to group (Figure 1C). Other transcript variations (*n* = 281) were identified at the level of transcript termination (29, or 10.3%), transcript initiation (72, or 25.6%) and promoter alteration (41, or 14.6%) (Figure 1C; theoretical transcripts with distinct exon compositions are also shown). The sum of the differential splicing isoforms is more than that of the alternative spliced genes since multiple alternative splicing events may occur in the same gene. On average, there were 1.49 (805/542) alternatively spliced exons per gene. Notably, some of the genes that showed changes at the level of whole gene expression also showed further changes in terms of splicing, transcript termination and/or transcript initiation. Examples of this include IL1B and EGFR (28) and suggest that transcript isoform variation constitutes a large part of the genetic variation we have observed in this study and in previous gene level-based KSHV transcriptomic studies (13).

The exon expression profiles of the above 542 genes are presented in Figure 1D as a heat map. The top 50 most significant in terms of their MiDAS *P*-value alternative spliced exons and their corresponding transcript names are also listed (Figure 1D). GATA2, a key stemness factor in CD133<sup>+</sup> endothelial stem cells, which plays a critical role in vascular system development (29) is one of them (Figure 1D). Overexpressing GATA2 in matured endothelial cells induces dedifferentiation-like transcriptome drift (22). Also present in this group are MAF and



**Figure 1.** KSHV induces alternative RNA splicing in infected LEC. (A) Confirmation of KSHV infection in LECs. The cell morphology of mock infection LECs at 2 days (LEC panel) shows a cobble stone appearance, whereas LECs infected by BCBL1-derived rKSHV at 2 days have become elongated and spindle shaped (KLEC panels). The GFP positive cells indicate successful KSHV infection. KSHV infection was further confirmed by RT-PCR of KSHV genes (upper-right panel). LEC and KLEC indicate cDNA from these two samples was used as PCR-template. LEC- and

(continued)

IGFBP2, which are two oncofetal proteins that are essential for normal embryonic growth and development (30–33). Various angiogenesis related genes, such as EPHB2, HIF3A and RUNX1T1, were also found to be altered by KSHV, which is consistent with the clinical observation that KS is a highly angiogenic neoplasm (34). Exons of various genes involved in cell growth regulation or DNA repair, including BRCA1, EGFR and SMAD6, were also in this group (Figure 1D). Among KSHV-regulated isoforms, exons 4–8 of ST5 (Suppression of tumorigenicity 5) are down-regulated (Supplementary Figure S1). On the other hand, exon 5 of CLDND1 (Cludin N-terminal domain 1) was found included in KSHV-infected LEC whose AS pattern is a typical simple exon inclusion (Figure 1D and E). As shown in Figure 1E, the AS of CLDND1 was validated by RT-PCR using exon-body primers present in the two exons flanking the candidate probe set.

### Functional module and pathway analysis as a framework for interpreting the biology of KSHV-induced AS

The exon/gene list gave us preliminary insights into the functional consequences of the detected differential isoform expression. To understand more about how AS might be correlated with sarcomagenesis and the KSHV life cycle, as well as to provide quantitative evidence, the 542 spliced variants identified were subjected to a GO database analysis by GO-Elite program implemented in AltAnalyze (35). The aim was to pinpoint statistically overrepresented functional groups within this AS list. The GO categories of the biological processes that were statistically overrepresented ( $P < 0.05$ ) among genes with different exons are shown in Figure 2A. Genes regulating the cell cycle, cell migration, cell adhesion and cell junction assembly were significantly affected by KSHV, implying that KSHV is actively involved in the regulation of LEC motility and proliferation, that is promoting lymphangiogenesis. It has been shown that KSHV, as well as viral proteins, are able to induce cell migration in HUVECs and immortalized microvascular endothelial cells (MVECs) (36–38). We confirmed by Transwell assay that KSHV also stimulates the motility of primary LECs (Figure 2A, histogram).

The ‘insulin receptor’ and ‘smoothened signaling’ pathways were also both affected by KSHV (Figure 2A, Supplementary Figure S2A). Insulin receptor (INSR) has been shown to be essential for the virus-induced tumorigenesis of Kaposi’s sarcoma (39). Further KSHV-regulated canonical pathways were revealed when the AS genes were subjected to analysis by the Ingenuity Pathway Analysis (IPA) web tool. KSHV-induced AS events were found to involve cancer pathways (such as the HER-2/neu signaling pathway, Figure 2B, Supplementary Figure S2B) and active cell cycle genes (such as the ‘G2/M DNA Damage Checkpoint Regulation’ one; Figure 2B, Supplementary Figure S2C). The top-ranked canonical pathway is ‘Virus Entry via Endocytic Pathways’, especially the ‘Macropinocytosis Signaling’ (Figure 2B, Supplementary Figure S2D and S2E). It has been reported that KSHV utilizes an actin polymerization-dependent macropinocytic pathway to enter both HUVECs and dermal MVECs (28). The results from the exon array analysis support the hypothesis that KSHV applies a similar macropinocytic pathway when entering primary LECs. Finally, a number of pathways that have not been linked with KSHV infection, such as the GNRH (gonadotrophin releasing hormone) and RAR (retinoic acid receptor) pathways, were also pinpointed (Figure 2B, Supplementary Figure S2F and S2G).

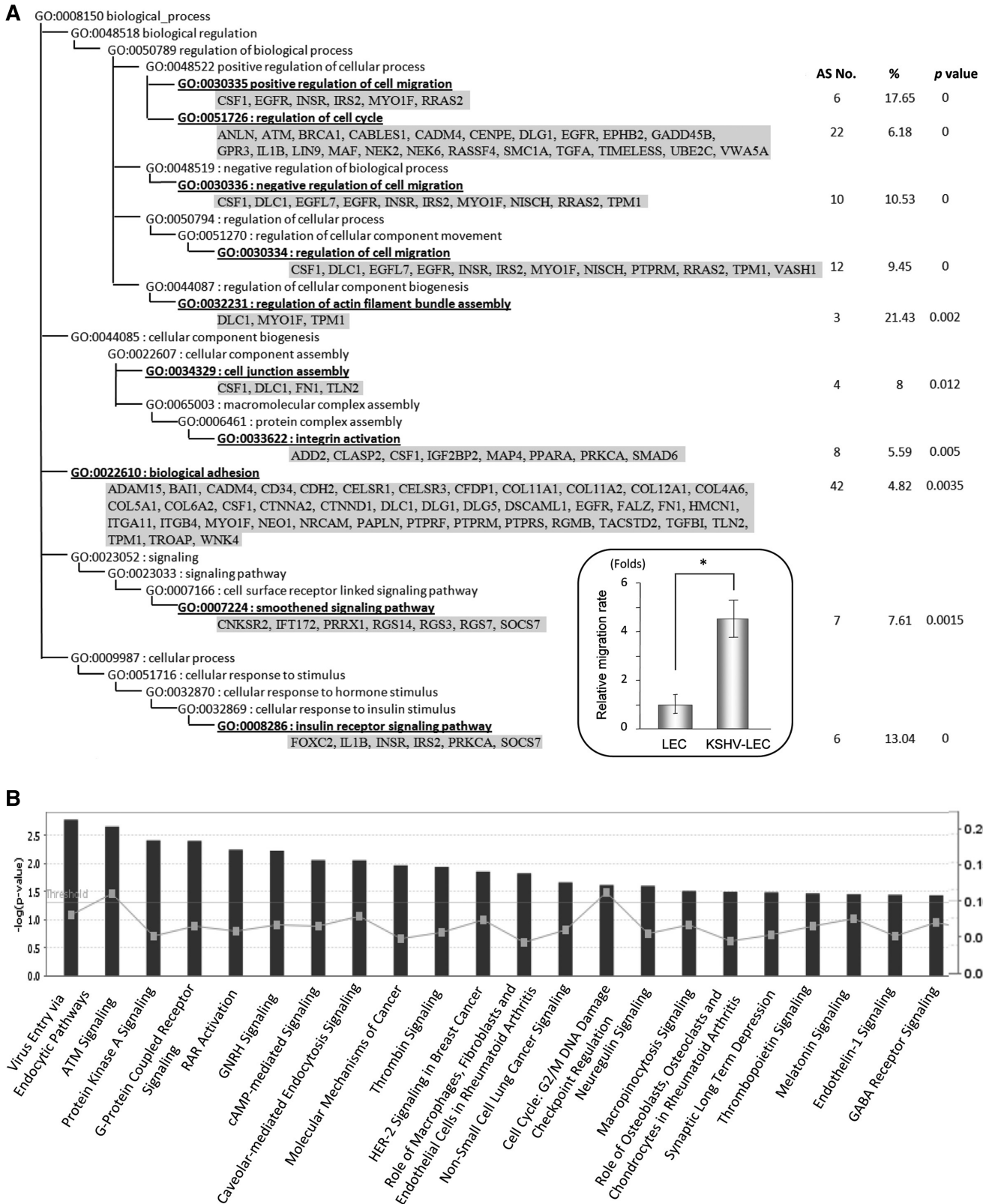
### Genetic networks and submodules in KSHV-infected LECs

Increasing evidence shows that genes do not act individually but collaborate in genetic networks. To better understand how the AS genes that are enriched in KSHV-LECs are related to each other, we performed genetic network analysis. AS transcripts were inputted into the IPA software package in order to construct network modules. The knowledge base behind IPA summarizes the known molecular interactions found in the published literature. The term ‘network’ in IPA is not the same as a biological or canonical pathway with a distinct function (i.e. angiogenesis) but rather is a reflection of all the interactions of a given protein as defined in the literature.

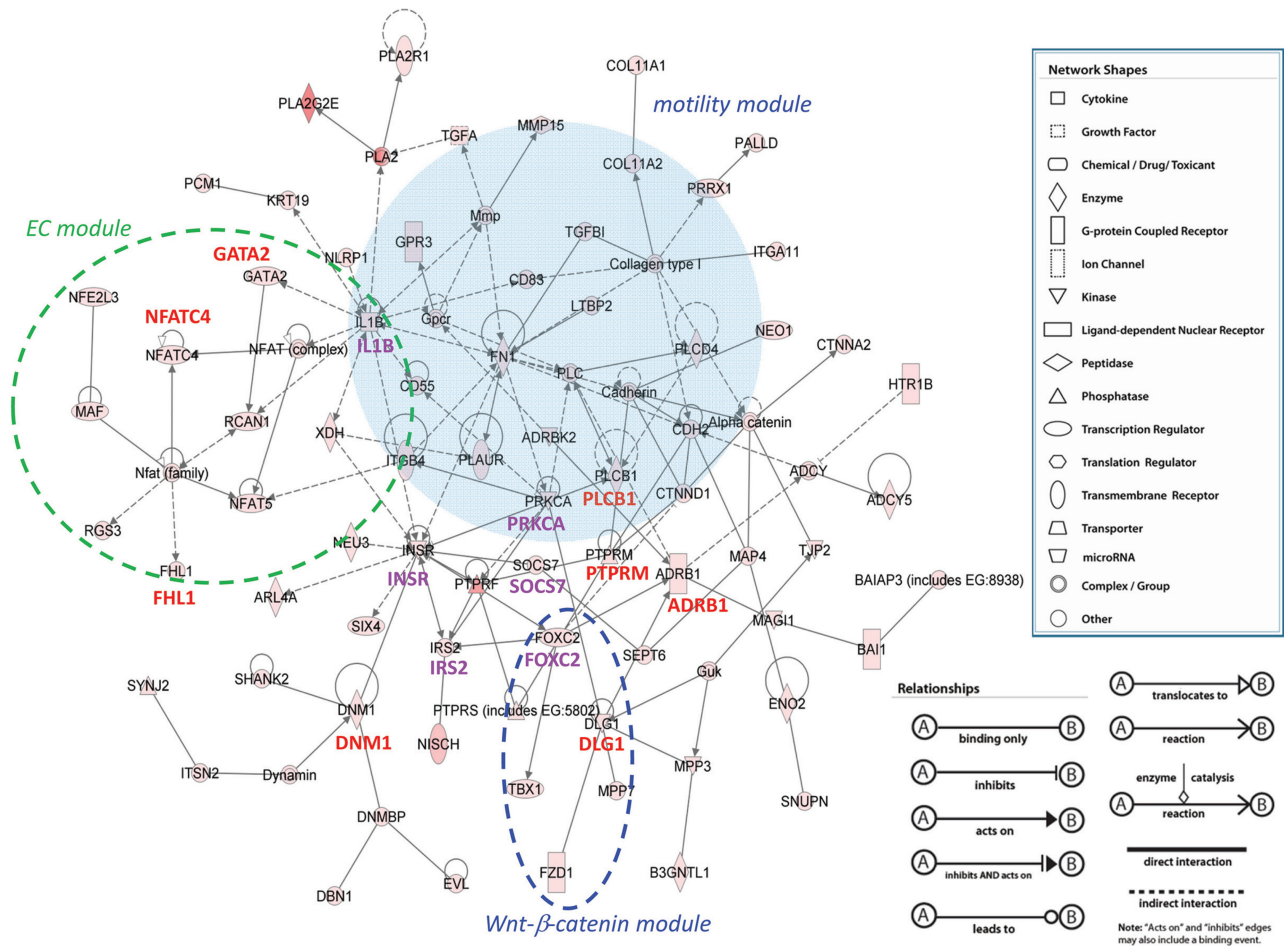
Among KSHV-regulated RNA isoforms, a major network consisting of 85 genes was identified (Figure 3 and File S3 in Supplementary Data). This network includes known angiogenesis or cell motility genes. Among these

#### Figure 1. Continued

KLEC- mean no-RT control, and the (–) set is a no template control. Three KSHV genes: LANA, vCyclin and vFLIP were all expressed in the KLECs. Immunofluorescent staining of KSHV LANA protein is demonstrated in the lower right panel, which shows a typical nuclear stippling staining pattern. (B) PCA plots of both gene and exon levels between KLEC and LEC. The PCA plot, based on the differential expressed genes with a pFDR < 0.1 is shown (Left). (Right) A PCA plot of normalized intensities from probe sets with MiDAS  $P < 0.05$ . KSHV-LEC samples are labeled in red while LEC samples are labeled in blue. (C) Statistics of AS events. AS events detected by AltAnalyze are summarized into the categories indicated. (D) A heat map of AS probe sets. The color of heat map is labeled according to the NI of each probe set. In red, inclusive exons; in blue, exclusive. The affected AS genes are also listed. (E) Verification of CLDND1 AS event. (Upper-left) The DomainGraph output for domain annotation. The top row labels the protein coding sequence (CDS), and the blue box indicates the Cludin domain. The middle row shows the composition of mRNA. Each exon is labeled with a number. The lower row designates the exon array probe select regions (PSRs). Probe sets upregulated in KLECs are in red. Blue lines below the last two PSRs indicate that these regions contain microRNA binding sites. The red rectangle emphasizes that exon 5 is upregulated in KLECs. (Lower-left) The analysis results from easyExon. The upper polygonal graph demonstrates the log<sub>2</sub>-scaled intensities across the whole transcript cluster (red: KLEC; blue: LEC). The SI and FC plots indicate the splicing index and fold-change across each probe set, respectively. The red box again highlights a differential expressed probe set. (Upper-right) PCR confirmation on AS events. A primer pair designed based on exon 1 and 6 amplifies two fragments that are different in size. The larger form is enriched in KLECs, which is consistent with the array data. (Lower-right) A histogram of the quantified long form/short form ratio. The Student’s *t*-test indicates the ratio is different between KLECs and LECs ( $P < 0.0001$ ).



**Figure 2.** Functional analysis of KSHV induced AS genes. (A) Gene ontology (GO) tree of KSHV induced AS genes in LECs. AS No.: AS gene number. %: percentage of AS genes in this GO category. P-value: GO-Elite calculated P-value. The subfigure demonstrates the difference in migration rate of LECs upon KSHV infection. \* $P < 0.05$  (B) Ingenuity canonical pathway analysis of AS genes. The P-values were calculated by the Fisher's exact test. Central line: the P-value threshold (0.05). Connected grey dots: changed gene ratio in each canonical pathway.



**Figure 3.** A genetic network containing products of KSHV-induced AS genes. Protein nodes with AS are indicated by a red color. Solid lines designate binding and interaction, and dashed lines designate an indirect interaction. Proteins involved in cell motility are highlighted by a solid blue circle. The green-dashed elliptical region contains genes involved in endothelial cell activities (EC module), while the blue-dashed region represents the Wnt- $\beta$ -catenin module. Genes mentioned in the main text are in red, and those in purple belong to the INSR pathway. The gene function of each node is denoted in different shapes.

genes, MMP15, PLAUR (plasminogen activator, urokinase receptor), ITGB4, CD83 and CDH2 (N-cadherin) are able to induce cell migration and/or invasion, which may in turn contribute to KSHV-induced LEC migration/invasion (highlighted in the ‘motility module’, Figure 3). Other genes in this main network have no previous known involvement with KSHV biology, but known to be associated with angiogenesis; these include DLG1, IL1B, XDH (xanthine dehydrogenase), NFATC4, FHL1 (four and a half LIM domains 1) and ADRB1 (adrenergic,  $\beta$ -1, receptor). Of these genes, DLG1 (hDLG, human homologue of the *Drosophila* discs large gene) contributes to tumor angiogenesis and neoangiogenesis via scaffolding G-protein-coupled receptors (GPCRs) such as tumor endothelial marker 5 (TEM5) and other TEM5-like proteins at the plasma membrane (40). Similarly, NFATC4 signaling is able to initiate cross-talk between blood vessels and the surrounding tissues, which helps to pattern the vasculature (41). FHL1 has been found to be up-regulated in the pulmonary vasculature of pulmonary hypertension patients (42). The KSHV-induced

angiogenic AS transcripts mentioned above form a sub-module in the main network (the ‘EC module’, Figure 3), with IL1B being the hub gene (genes with higher connectivity to others or resided in a position among submodules in the major network). This network structure suggests that the pro-inflammatory cytokine IL1B plays a critical role in KSHV-related lymphangiogenesis and sarcomagenesis.

Hub genes in a biological system very often play more important roles than spoke genes, and the dysregulation of a hub can eventually lead to disruption of the genetic network and the malfunction of the cell (43). Systems biology analysis also revealed other KSHV-regulated hub genes. These included ADRB1, FOXC2, DLG1, DNMT1 (DNA (cytosine-5-)-methyltransferase 1), PLCB1 and INSR (Figure 3). FOXC2 is one of the downstream regulators of the canonical Wnt signaling pathway. FZD1 (receptor for Wnt proteins) and TBX1, a downstream effector of the Wnt- $\beta$ -catenin signaling pathway (44), were also included in KSHV-LECs (Figure 3). The genes involved in the INSR pathway, as revealed by GO analysis



in Figure 2A, were all included and form an important part of the main network (in purple, Figure 3).

#### Dedifferentiation-like AS drift in KSHV-infected LEC

From the present study, it can be clearly seen that KSHV is able to influence the splicing patterns of critical stemness genes and oncofetal proteins [such as GATA2 (22) and NFATC4 (41)] in LECs. The splicing pattern of the CD34 precursor marker was also altered by KSHV (File S2 in Supplementary Data). Since re-expression of GATA2 stemness genes is able to induce reprogramming and dedifferentiation of matured endothelial cells (45), and it has been recognized that aggressive and poor prognostic tumors acquire characters reminiscent of embryonic stem cells (ESCs) (46), we hypothesized that KSHV infection may induce dedifferentiation-like transcriptome reprogramming of mature LECs. When published mRNA expression arrays (14,22) were used to examine our hypothesis, no dedifferentiation could be found (Figure 4A). However, when transcriptome reprogramming was examined on an exome level, MDS analysis clearly showed that the exomes of the infected LECs had moved away from the parental pattern toward a pattern closer to that of CD34<sup>+</sup> endothelial stem/precursor cells (Figure 4B, left panel). This change was quantified by measuring the average linkage Euclidean distances between sample groups and KSHV-LECs were found to be closer to CD34<sup>+</sup> precursor cells than to uninfected LECs (Figure 4B, right panel).

A total of 14 AS transcripts with an exon-exclusion pattern, as well as another 19 exon-inclusion patterns, were found to be common between KSHV-infected LECs and CD34<sup>+</sup> cells (Figure 4C). CD34<sup>+</sup> exons, such as those in EGFR, CCDC85A, INCENP, ITGB4, KCNMB4, LIMK2, LTBP4, LY75, MLPH, RGS3, TPM1 and TXNRD2, were included more often in KSHV-LECs. In contrast, exons in CD55 and CD83 etc. were excluded (Figure 4D). In addition, EGFR and CASP9 were also in common between CD34<sup>+</sup> precursor cells and KSHV-LECs (Figure 4E). Both AS events include an additional exon into their spliced RNA isoforms in regions encoding important protein domains (red circles, Figure 4E), which suggest an alternation in the protein's activity or a change in the protein's regulatory mechanism. To support the reproducibility of our work, we used another rKSHV.219 viruses derived from JSC-1 cells (21) for RT-qPCR validation. The AS of EGFR and CASP9 could be validated by RT-qPCR using primer sets specific for exons 25 or 26 of EGFR, as well as primers for exons 4b or 5 of CASP9. The expression ratios between altered and constant exons were compared (Figure 4E). The consistency between exon array and RT-qPCR data suggests the reliability of our results.

#### Modification of domain and motif composition by KSHV-induced isoforms

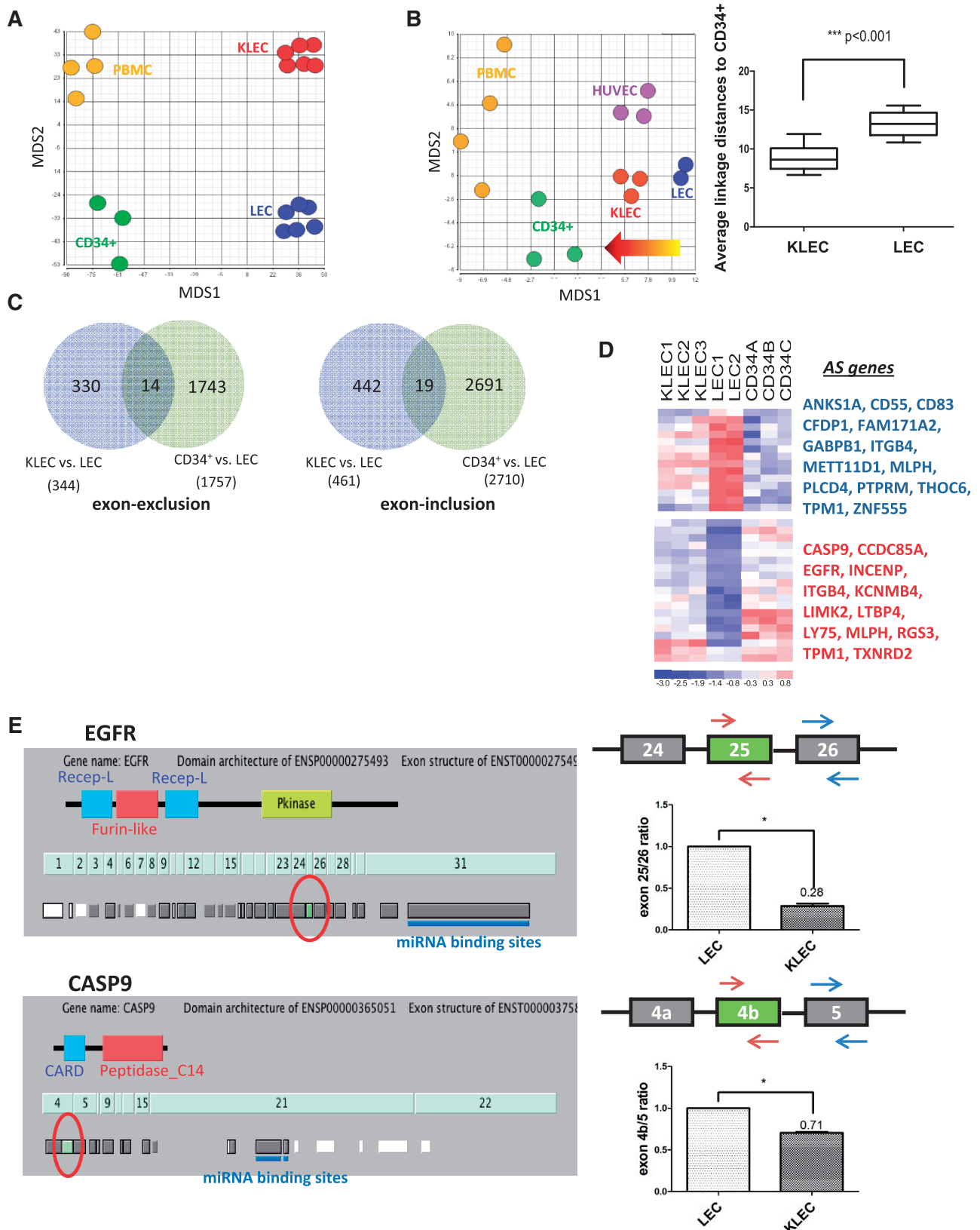
It is well known that splicing not only provides feedback that affects transcription, but also it can feed forward by modulating protein function. The above EGFR and CASP9 isoforms analysis suggested that it would be

useful to explore protein domains and motifs globally for possibly modification/disruption in KSHV-LECs. AltAnalyze, which contains a database consisting of protein domain and motif information obtained from multiple protein annotation databases (UniProt, InterPro) (25), was used for this task. A domain database search revealed 1394 protein domains in 36.2% (196/542) of genes that have alternatively spliced exons. Of these, 17 domain/motif types affected by KSHV were involved in adhesion and oncogenic signal transduction. These included SH3, fibronectin, collagen, Ras GEF, RasGRF and pleckstrin/G-protein interact domains (Figure 5A, File S4 Supplementary Data). Several AS isoforms, such as BAIAP2L1 (BAI1-associated protein 2-like 1, involved in the formation of clusters of actin bundles) and DVL2 (dishevelled, dsh homolog 2; a component of the Wnt signaling pathway), have alternations in SH3 and DEP protein domains, respectively (Figure 5B). The AS of both transcripts was validated by RT-qPCR in rKSHV.219 viruses infected LECs using primer sets specific for exons 11 and 12 of BAIAP2L1, or for exons 12, 13 of DVL2. The expression ratios between altered and constant exons were compared (Figure 5B). The tendencies of exon inclusion/exclusion of these two genes were similar between exon array and qRT-PCR results.

#### AS modifies microRNA binding site architecture

Another consequence of alternative RNA splicing is the gain or loss of microRNA binding sites in the AS isoforms. We next explored the possibility that KSHV-induced AS may cause the expressed isoforms to escape from or be newly included in microRNA-mediated regulation. We first determined which microRNAs (miRNAs) are present in KSHV-LEC. The overall miRNA expression signals in KSHV-LEC was determined by the Agilent v2 miRNA array, and a miRNA was assigned as 'Present' if its hybridization intensity was by a factor  $>16$  ( $2^4$ ). This cutoff value was determined by taking advantage of the fact that there are also probes detecting KSHV and EBV viral miRNAs on the Agilent miRNA chips. For most KSHV viral miRNAs, but not the EBV ones, their normalized array signals were more than  $2^4$  (Supplementary Figure S3). Using this criterion we found there were 216 cellular miRNAs expressed in KSHV-LECs and the top 40 miRNAs are shown (Figure 6A, File S5 Supplementary Data). The most abundant miRNA is miR-21, which can be induced by the KSHV K15 protein, a viral angiogenic latent protein (47). Another abundant KSHV-induced miRNA is miR-221, which is an anti-angiogenic miRNA present in KS tissues (48,49) (Figure 6A).

We then again applied AltAnalyze, which also contains detailed information on miRNA binding sites as predicted by multiple algorithms (PicTar, miRbase, miRanda and TargetScan) (25). Overall, 7.45% (60/805) of KSHV-induced AS events were found to harbor a predicted gain/loss of at least one miRNA binding site (Figure 6B). Taking into account whether miRNAs exist in KEHV-LECs or not, 85/302 (28.15%) miRNA binding motifs would seem to be significantly altered in KSHV-LECs (Figure 6B).



**Figure 4.** Dedifferentiation-like AS drift in KSHV-infected LECs. (A) A gene-level MDS plot using KLEC-LEC differential expressed genes (Affymetrix Human Genome U133 Plus 2.0 array). In this figure, no dedifferentiation could be found. (B) An exon-level MDS plot using NI of KSHV-induced AS probe sets. An arrow indicates that the exomes of the infected LECs have moved away from those of control LECs and toward to those of CD34<sup>+</sup> endothelial stem/precursor cells. (right) Average linkage distances between sample groups. (C) Venn diagrams showing AS events common between KSHV-LEC and CD34<sup>+</sup> cells. (D) A heat map showing the 33 AS events common between KSHV-LEC and CD34<sup>+</sup> cells. The

(continued)

The third parameter applied was the correlation of the expression levels of the candidate mRNA-miRNA pairs in LECs and KSHV-LECs. It has been reported that the expression levels of a miRNA and its downstream mRNA targets are negatively correlated (50). So, when AS events with inclusive miRNA sites were checked, it was found that the levels were down in KSHV-LECs compared with those in parental LECs control. A total of 16 genes passed this third criterion (Figure 6C, upper panel). ATM, HOXA3 and AURKA were three typical genes (Supplementary Figure S1). Another four genes, IL4L, HSD17B7P2, RAPGEF5 and RGS3, with exon inclusion were up in KSHV-LEC (Figure 6C, lower panel). Some miRNAs targeted more than one AS event in KSHV-LECs and may therefore be more crucial. miR-23a and miR-27a binding sites were over-represented among KSHV-regulated AS events and both miRNA expression levels were on the top 10 in KSHV-infected LECs (Figure 6A and C). Neither miRNA has been previously associated with the KSHV life cycle or KS tumorigenesis.

The decreasing expression of MLPH and CTCL1 (both with inclusive miRNA binding sites) in rKSHV.219 viruses infected LEC was verified by RT-qPCR, while that of RAPGEF5 (with exclusive miRNA binding sites) was increased (Figure 6D). The role of RAPGEF5 in KSHV-induced pathogenicity was explored by examining its function in LEC motility: when the expression of endogenous RAPGEF5 was knocked down by shRNA, a reduced cellular motility was observed (Figure 6E), suggesting RAPGEF5 is crucial for LEC motility.

## DISCUSSION

In this study, we did exon level expression profiling on KSHV(+) LECs. Understanding the added complexity of transcript-processing variations and the potential outcome of these differences will greatly enhance our knowledge of KS pathogenesis and KSHV life cycle. Very few studies have investigated splicing regulation in endothelial cells. Elucidating the underlying mechanisms associated with KSHV-induced phenomena, such as LEC motility, is critical to a better understanding of vascular biology under normal and pathological conditions.

In this study we applied only the 'Core' exon panel from the Affymetrix annotation. Since the exons in the 'Core' annotation set are well recognized ones, our results are therefore reliable but are somewhat reliant on prior knowledge. It should be possible to identify more AS events and get broader insights into KSHV and LECs biology if the 'Extended' or even the 'Full' exon panels are used. In addition, it ought to be possible to find novel KSHV-regulated exons and RNA isoforms if we extend our

analysis further. On the other hand, the RNA-Seq application in relation to next-generation sequencing (NGS) can also be used to discover novel exons and RNA isoforms [reviewed by (51)]; in these circumstances even, more KSHV-regulated AS events should be found when RNA-Seq is applied to the study of KSHV and KS biology. A very recent article used whole-exome sequencing to decipher the genetic basis of fatal classic Kaposi sarcoma and discovered a homozygous splice-site mutation in the STIM1 gene (52). This proof-of-principle similar case to KS also supports the critical role of AS in KS tumorigenesis.

The AS induced by KSHV infection may either be caused specifically by KSHV or by general virus-host interaction. To clarify these issues, we analyzed another exon array dataset regarding EBV-infected primary B cells (GEO dataset GSE20200). In this analysis, 773 genes harbored alternative splicing upon early EBV infection, in which 49 AS events were also found in KSHV-infected LEC (Supplementary Figure S4 and Supplementary Table S2). These 49 AS events represent the common AS patterns that can be induced by the infection of  $\gamma$ -herpesvirus.

Gene expression profiles for KSHV-infected cells have been reported by a few groups (13,53,54). Nevertheless, assigning biological significance and identifying potential targets is always a daunting task. Analyzing gene signatures using systems biology approaches, such as dividing the genes into functional subgroups or network modules, is an efficient way of obtaining more insights from the gene lists (55). A systemic strategy is mandatory when trying to view the overall molecular events as the biological system for a given biological process; this allows important genes to be pinpointed as controllers (55). Those key genes often serve as hubs that maintain the stability of the genetic module or act as go-between connecting modules within a major network. By applying systems biology tools we were able to identify a major functional network that consisted of KSHV-infected AS events. Hub genes, such as IL1B, FOXC2 and INSR, were then identified using this systemic approach. It should be noted, however, because we applied a knowledge-based strategy to construct the genetic networks and to separate the various functional groups, only genes with known interactions are included. In such circumstances, the enrichment of AS events and/or the appearance of hubs may be a consequence of previous intense research in these areas. There are probably many more potential key genes currently without known interactions that are still waited to be unveiled.

Six insulin receptor (INSR) pathway genes (IL1B, INSR, IRS2, FOXC2, PRKCA and SOCS7) were found

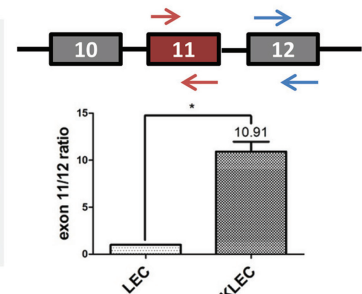
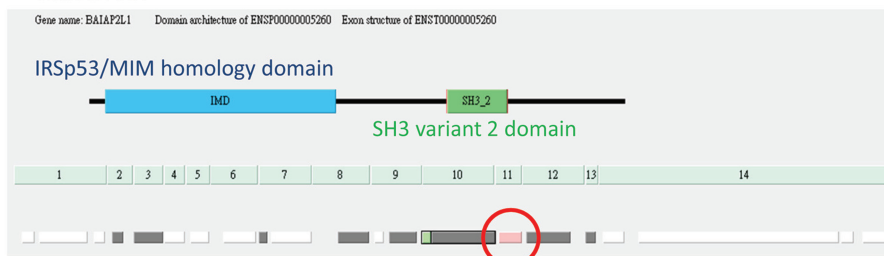
### Figure 4. Continued

Gene symbols of AS genes are also listed. (E) DomainGraph 'Probeset view' of EGFR and CASP9. EGFR exon 25 is downregulated upon KSHV infection, and this KSHV-regulated isoform would seem to harbor an altered protein kinase domain. Part of CASP9 exon 4 is downregulated in infected LECs, and this AS will affect the integrity and function of the CARD domain. Green boxes: probe sets with exon exclusion; White boxes: probe sets that do not meet the minimum expression threshold in AltAnalyze; Gray boxes: no alternative expression event detected. (right) Validation of AS events by RT-qPCR. Primer sets for EGFR exon 25 and CASP9 exon 4b were designed for detecting differential expressed exons, whereas primer sets for EGFR exon 26 and CASP9 exon 5 for constantly expressed exons. Total RNA was from primary LECs infected with JSC-1-derived rKSHV.219 viruses. Results are expressed as the mean  $\pm$  SD. Ratios between AS exons and constant exons are shown. \* $P < 0.05$  by Student's *t*-test.

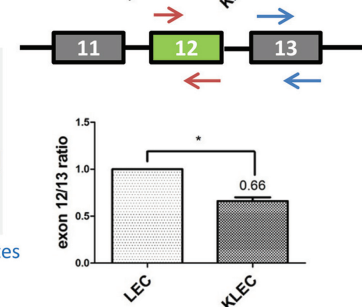
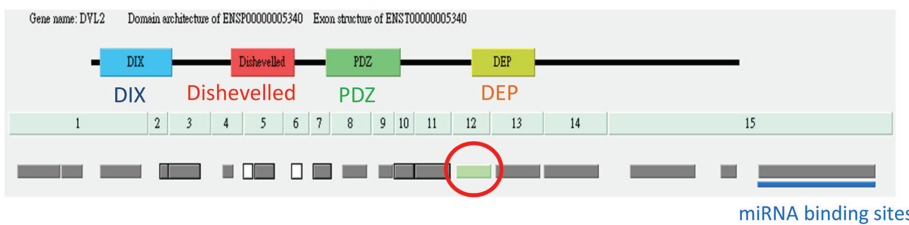
**A**

Domain-Name	IPR	# Changed	% Changed	Zscore	Changed GeneSymbols
PDZ/DHR/GLGF	IPR001478	10	8.196721	3.936505	<b>DVL2</b> , DLG1, MAST2, PARD3B, RGS3, MPP7, DLG5, AC121493.1, TIAM1, PDLIM5
SH3_domain	IPR001452	10	5.681818	2.61786	<b>BAIAP2L1</b> , DLG1, DNMBP, TJP2, MPP7, DLG5, MPP3, SHANK2, PACSIN3, ITSN2
Ig_I-set	IPR013098	9	6.666667	3.015143	MYLK, PAPLN, PTPRS, PALLD, PTPRF, HMCN1, PTPRM, DSCAML1, ADAMTSL1
Fibronectin_typ-III-like_fold	IPR008957	9	5.696203	2.490175	NRCAM, PTPRS, COL12A1, FN1, ITGB4, EPHB2, PTPRF, PTPRM, DSCAML1
FN_III	IPR003961	8	5.673759	2.333978	PTPRS, COL12A1, FN1, ITGB4, EPHB2, PTPRF, PTPRM, DSCAML1
FnIII_subd	IPR003962	7	9.333333	3.704071	PTPRS, FN1, ITGB4, EPHB2, PTPRF, PTPRM, DSCAML1
SH3_2	IPR011511	7	4.605263	1.588064	<b>BAIAP2L1</b> , DLG1, DNMBP, MPP7, SHANK2, PACSIN3, ITSN2
Collagen	IPR008160	6	9.375	3.441352	COL12A1, COL5A1, C1QTNF6, COL6A2, COL4A6, COL11A2
VWF_A	IPR002035	6	8.695652	3.216812	COL12A1, CPNE5, ITGA11, COL6A2, HMCN1, VWA5B1
Guanylate_kin	IPR008144	5	27.77778	6.753377	DLG1, TJP2, MPP7, AC121493.1, MPP3
DAG_PE_bd	IPR002219	5	7.246377	2.454976	TRIM24, CHN1, NSMCE1, PRKCE, FALZ
C2_membr_targeting	IPR018029	5	5.747126	1.872967	BAIAP3, PLCD4, CPNE5, RGS3, PRKCE
C2_Ca-dep	IPR000008	5	4.424779	1.245632	BAIAP3, PLCD4, CPNE5, RGS3, PRKCE
Pleckstrin/G-protein_interact	IPR000591	4	23.52941	5.456523	<b>DVL2</b> , DEPDC5, RAPGEF5, RGS7
Guanylt/Ca	IPR008145	4	18.18182	4.623966	DLG1, TJP2, AC121493.1, MPP3
RasGRF_CDC25	IPR001895	4	15.38462	4.126384	RAPGEF4, RALGPS2, RAPGEF5, RGL1
Ras_GEF	IPR008937	4	14.81481	4.018074	RAPGEF4, RALGPS2, RAPGEF5, RGL1

**B BAIAP2L1**



**DVL2**

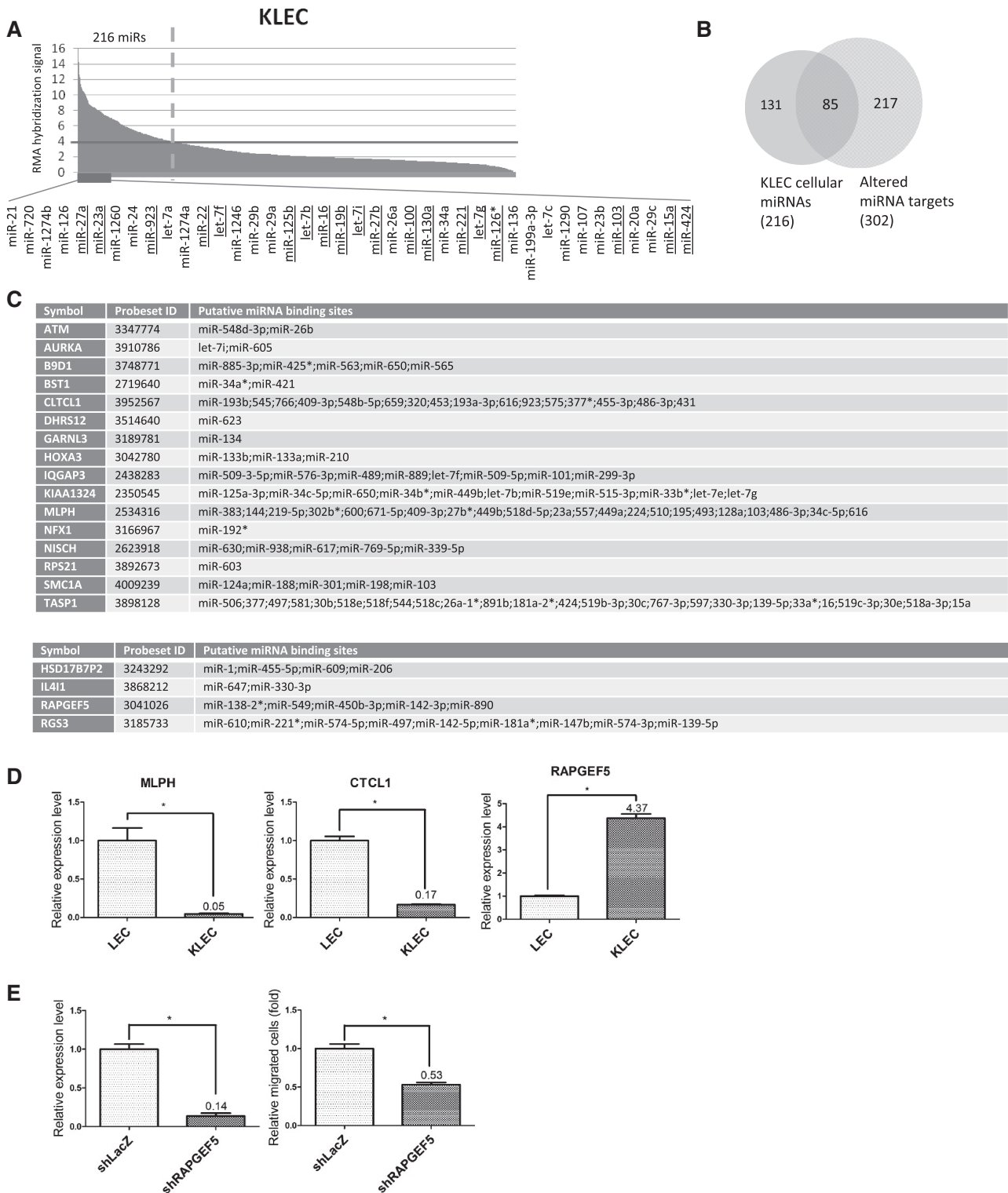


**Figure 5.** Modification of domain and motif compositions within KSHV-induced isoforms. (A) Table of protein domains in genes with AS exons. In total, 17 domains/motifs that are involved in adhesion and oncogenic signaling transduction are affected on KSHV infection. (B) Part of the SH3\_2 domain is altered in BAIAP2L1 due to the upregulated exon 11. Similarly, part of the DEP domain (Domain found in Dishevelled, Egl-10, and Pleckstrin) in DVL2 is affected on KSHV infection due to the downregulation of exon 12. Probe sets with exon inclusion are in red, whereas those with exon exclusion in green. (right) Validation of AS events by RT-qPCR. Total RNA was from LECs infected with JSC-1-derived rKSHV.

to be differentially spliced in KSHV-infected LECs (Figure 3). Systems biology analysis further revealed that INSR is a hub gene in this KSHV-regulated genetic network (Figure 3). This further suggests a critical role for this pathway in KS tumorigenesis and the KSHV life cycle. IL1B expression is detrimental to viral infection (56). It has recently been shown that KSHV lytic gene Orf63 is a viral homolog of NLR (nucleotide binding and oligomerization, leucine-rich repeat) and can subvert

the function of cellular NLRs (57). As a result, KSHV Orf63 hinders inflammasome formation and IL1B and IL18 secretion during KSHV lytic replication (57). IL1B and NLR-mediated innate immunity is suggested to be important for the lifelong persistence of herpesviruses. Our data weight the importance of IL1B in KSHV life cycle and imply possible applications.

The exact biological roles of KSHV-induced protein isoforms still need more investigation. We showed by



**Figure 6.** AS modifies miRNA binding site architecture. (A) The miRNA expression profile of KSHV infected LECs. In total, 216 miRNAs that have summarized probe set intensities  $>4$  on a  $\log_2$  scale can be considered to be expressed in the cells. The y-axis scale is the RMA-normalized hybridization signal ( $\log_2$  transformed) (B) Venn diagram of expressed miRNAs and altered miRNA binding sites identified by AltAnalyze. Overall, 85 miRNA-altered miRNA binding sites can be identified. (C) Correlation between differential gene expression and miRNA-miRNA binding sites regulation. (Upper) Genes downregulated in KSHV-LEC, and their miRNA binding sites are in inclusive exons. (Lower) Genes up-regulated in KSHV-LEC, and their miRNA binding sites are in exclusive exons. (Some 'miR-' prefixes were omitted for better demonstration.) (D) Relative expression levels of genes with miRNA binding site alteration. Potential miRNA binding sites are upregulated in MLPH and CTCL1, and their gene expression levels are significantly down. In contrast, the miRNA binding sites in RAPGEF5 are downregulated, and its gene expression level is up. Total RNA was from LECs infected with JSC-1-derived rKSHV. (E) RAPGEF5 is crucial for LEC motility. Primary LECs transduced with shRNA against LacZ or RAPGEF5 were subjected to Transwell cell migration assays (right). The knock-down effect was measured by RT-qPCR (left). Mean expression levels of RAPGEF5 were compared with that of GAPDH control.

RT-qPCR and Transwell assays that the expression of RAPGEF5 (with exclusive miRNA binding sites) is increased in KSHV-LEC, and this gene is crucial in LEC motility (Figure 6E). KSHV may exploit RAPGEF5 splicing for enhancing the migration and spreading of infected cells. Another AS product is CASP9 isoform (with reduced exon 4 and defected CARD domain; Figure 4E). The product from KSHV-induced AS CASP9 is expected to be similar to that encoded by Casp9b, which is due to the exclusion of an exons 3–6 cassette (58). Casp9b protein acts as an endogenous inhibitor of Casp9a (with intact CARD) by competing with the full-length Casp9a for binding to the apoptosome (58–60). Casp9b protein also directly interacts with Casp9a (full-length) for blocking the autoproteolysis of the enzyme (59). We anticipate that KSHV-induced AS CASP9 may also inhibit wild-type CASP9 for KSHV to escape from apoptosis.

It is known that KSHV is able to induce 'trans-differentiation'-like transcriptome drift and this will influence the cellular identity of infected ECs, thereby contributing to KSHV-induced oncogenesis (13). On the other hand, the possibility that KSHV may induce dedifferentiation in infected cells has not been explored. It has been recognized that during malignant transformation, cancer cells acquire genetic mutations that override the normal mechanisms that control cellular proliferation. Cancer cells further acquire characters reminiscent of normal stem cells during disease progression, and clinically cancer cells with poor differentiated pathological grading usually have a worse therapeutic response than those with well-differentiated morphology (61). The degree of embryonic gene re-expression correlates with pivotal tumor features and patient prognosis (62,63). In poorly differentiated tumors, stem cell-like gene expression signatures are exhibited and the degree of stem cell program recapitulation correlates with tumor stages and patient survival (46,64). Identifying genes shared between transformed cells, especially the more malignant ones, and stem cells will help to unmask the pathogenic mechanisms at work in tumors, as well as provide us with novel therapeutic targets and prognosis biomarkers. In this study, it has been revealed by exon array that KSHV can also induce 'dedifferentiation' like exome drift, which is also readily disclosed by classical gene expression-based transcriptome analysis (Figure 4). It has been shown recently that KSHV is able to induce the expression of an endothelial precursor cell marker (CD133) and various mural cell markers (calponin, desmin and smooth muscle alpha actin) (65), which support both dedifferentiation and trans-differentiation may occur in KSHV-infected cells.

In summary, our present results demonstrate that the nature of KSHV-induced transcriptome changes is not only quantitative but is also qualitatively. AS is a key component of cellular reprogramming upon viral infection, and exon array analysis could identify previously undisclosed potential oncogenic pathways in an unbiased manner. Our results imply that detecting exon changes is a more sensitive tactic than analyzing mRNA levels only. Careful studies of the functional variants arising from AS in the context of *in vitro* and *in vivo* models of

oncogenesis are still necessary to further support the clinical relevance of these exciting findings.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors acknowledge the technical services provided by Microarray & Gene Expression Analysis Core Facility, which was supported by National Research Program for Genomic Medicine, National Science Council (NSC), of the National Yang-Ming University VGH Genome Research Center and the shRNA knockdown plasmids provided by National RNAi Core Facility, Taiwan.

## FUNDING

This work was supported by National Health Research Institutes, Taiwan (NHRI-EX99-9704BI), National Science Council, Taiwan (NSC98-3111-B-010-004, NSC98-2320-B-010-020-MY3, NSC 99-2911-I-009-101 and NSC99-2627-B-010-010), Department of Health, Taiwan (CCMP99-RD-063), Taipei Veteran General Hospital (V100E2-011 and DOH100-TD-C-111-007, Center of Excellence for Cancer Research at Taipei Veterans General Hospital) and National Yang-Ming University, Taiwan (Ministry of Education, Aim for the Top University Plan). Funding for open access charge: National Yang-Ming University, Taiwan (Ministry of Education, Aim for the Top University Plan).

*Conflict of interest statement.* None declared.

## REFERENCES

- Black,D.L. (2000) Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. *Cell*, **103**, 367–370.
- Clark,T.A., Sugnet,C.W. and Ares,M. Jr. (2002) Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science*, **296**, 907–910.
- Johnson,J.M., Castle,J., Garrett-Engele,P., Kan,Z., Loerch,P.M., Armour,C.D., Santos,R., Schadt,E.E., Stoughton,R. and Shoemaker,D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, **302**, 2141–2144.
- Le,K., Mitsouras,K., Roy,M., Wang,Q., Xu,Q., Nelson,S.F. and Lee,C. (2004) Detecting tissue-specific regulation of alternative splicing as a qualitative change in microarray data. *Nucleic Acids Res.*, **32**, e180.
- Pan,Q., Shai,O., Misquitta,C., Zhang,W., Saltzman,A.L., Mohammad,N., Babak,T., Siu,H., Hughes,T.R., Morris,Q.D. *et al.* (2004) Revealing global regulatory features of mammalian alternative splicing using a quantitative microarray platform. *Mol. Cell*, **16**, 929–941.
- Ule,J., Stefani,G., Mele,A., Ruggiu,M., Wang,X., Taneri,B., Gaasterland,T., Blencowe,B.J. and Darnell,R.B. (2006) An RNA map predicting Nova-dependent splicing regulation. *Nature*, **444**, 580–586.
- Ule,J., Ule,A., Spencer,J., Williams,A., Hu,J.S., Cline,M., Wang,H., Clark,T., Fraser,C., Ruggiu,M. *et al.* (2005) Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.*, **37**, 844–852.

8. Hengge, U.R., Ruzicka, T., Tying, S.K., Stuschke, M., Roggendorf, M., Schwartz, R.A. and Seeber, S. (2002) Update on Kaposi's sarcoma and other HHV8 associated diseases. Part 2: pathogenesis, Castleman's disease, and pleural effusion lymphoma. *Lancet Infect Dis.*, **2**, 344–352.
9. Chang, Y., Cesarman, E., Pessin, M.S., Lee, F., Culpepper, J., Knowles, D.M. and Moore, P.S. (1994) Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science*, **266**, 1865–1869.
10. Ganem, D. (2010) KSHV and the pathogenesis of Kaposi sarcoma: listening to human biology and medicine. *J. Clin. Invest.*, **120**, 939–949.
11. Luan, S.L., Boulanger, E., Ye, H., Chanudet, E., Johnson, N., Hamoudi, R.A., Bacon, C.M., Liu, H., Huang, Y., Said, J. *et al.* (2010) Primary effusion lymphoma: genomic profiling revealed amplification of SELPLG and CORO1C encoding for proteins important for cell migration. *J. Pathol.*, **222**, 166–179.
12. Toth, Z., Maglinte, D.T., Lee, S.H., Lee, H.R., Wong, L.Y., Brulois, K.F., Lee, S., Buckley, J.D., Laird, P.W., Marquez, V.E. *et al.* (2010) Epigenetic analysis of KSHV latent and lytic genomes. *PLoS Pathog.*, **6**, e1001013.
13. Wang, H.W., Trotter, M.W., Lagos, D., Bourbouli, D., Henderson, S., Makinen, T., Elliman, S., Flanagan, A.M., Alitalo, K. and Boshoff, C. (2004) Kaposi sarcoma herpesvirus-induced cellular reprogramming contributes to the lymphatic endothelial gene expression in Kaposi sarcoma. *Nat. Genet.*, **36**, 687–693.
14. Lagos, D., Trotter, M.W., Vart, R.J., Wang, H.W., Matthews, N.C., Hansen, A., Flore, O., Gotch, F. and Boshoff, C. (2007) Kaposi sarcoma herpesvirus-encoded vFLIP and vIRF1 regulate antigen presentation in lymphatic endothelial cells. *Blood*, **109**, 1550–1558.
15. Hanahan, D. and Weinberg, R.A. (2000) The hallmarks of cancer. *Cell*, **100**, 57–70.
16. Chang, W.Y. and Stanford, W.L. (2008) Translational control: a new dimension in embryonic stem cell network analysis. *Cell Stem Cell*, **2**, 410–412.
17. Stites, E.C., Trampont, P.C., Ma, Z. and Ravichandran, K.S. (2007) Network analysis of oncogenic Ras activation in cancer. *Science*, **318**, 463–467.
18. Gardina, P.J., Clark, T.A., Shimada, B., Staples, M.K., Yang, Q., Veitch, J., Schweitzer, A., Awad, T., Sugnet, C., Dee, S. *et al.* (2006) Alternative splicing and differential gene expression in colon cancer detected by a whole genome exon array. *BMC Genomics*, **7**, 325.
19. Clark, T.A., Schweitzer, A.C., Chen, T.X., Staples, M.K., Lu, G., Wang, H., Williams, A. and Blume, J.E. (2007) Discovery of tissue-specific exons using comprehensive human exon microarrays. *Genome Biol.*, **8**, R64.
20. Zhou, F.C., Zhang, Y.J., Deng, J.H., Wang, X.P., Pan, H.Y., Hettler, E. and Gao, S.J. (2002) Efficient infection by a recombinant Kaposi's sarcoma-associated herpesvirus cloned in a bacterial artificial chromosome: application for genetic analysis. *J. Virol.*, **76**, 6185–6196.
21. Vieira, J. and O'Hearn, P.M. (2004) Use of the red fluorescent protein as a marker of Kaposi's sarcoma-associated herpesvirus lytic gene expression. *Virology*, **325**, 225–240.
22. Huang, T.S., Hsieh, J.Y., Wu, Y.H., Jen, C.H., Tsuang, Y.H., Chiou, S.H., Partanen, J., Anderson, H., Jaatinen, T., Yu, Y.H. *et al.* (2008) Functional network reconstruction reveals somatic stemness genetic maps and dedifferentiation-like transcriptome reprogramming induced by GATA2. *Stem Cells*, **26**, 1186–1201.
23. Emig, D., Salomonis, N., Baumbach, J., Lengauer, T., Conklin, B.R. and Albrecht, M. (2010) AltAnalyze and DomainGraph: analyzing and visualizing exon expression data. *Nucleic Acids Res.*, **38**(Suppl.), W755–W762.
24. Chang, T.Y., Li, Y.Y., Jen, C.H., Yang, T.P., Lin, C.H., Hsu, M.T. and Wang, H.W. (2008) easyExon—a Java-based GUI tool for processing and visualization of Affymetrix exon array data. *BMC Bioinformatics*, **9**, 432.
25. Salomonis, N., Nelson, B., Vranizan, K., Pico, A.R., Hanspers, K., Kuchinsky, A., Ta, L., Mercola, M. and Conklin, B.R. (2009) Alternative splicing in the differentiation of human embryonic stem cells into cardiac precursors. *PLoS Comput. Biol.*, **5**, e1000553.
26. John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C. and Marks, D.S. (2004) Human MicroRNA targets. *PLoS Biol.*, **2**, e363.
27. Griffiths-Jones, S., Saini, H.K., van Dongen, S. and Enright, A.J. (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res.*, **36**, D154–D158.
28. Raghu, H., Sharma-Walia, N., Veetil, M.V., Sadagopan, S. and Chandran, B. (2009) Kaposi's sarcoma-associated herpesvirus utilizes an actin polymerization-dependent macropinocytic pathway to enter human dermal microvascular endothelial and human umbilical vein endothelial cells. *J. Virol.*, **83**, 4895–4911.
29. Bresnick, E.H., Lee, H.Y., Fujiwara, T., Johnson, K.D. and Keles, S. (2010) GATA switches as developmental drivers. *J. Biol. Chem.*, **285**, 31087–31093.
30. Busund, L.T., Richardsen, E., Busund, R., Ukkonen, T., Bjornsen, T., Busch, C. and Stalsberg, H. (2005) Significant expression of IGFBP2 in breast cancer compared with benign lesions. *J. Clin. Pathol.*, **58**, 361–366.
31. Poupponot, C., Sii-Felice, K., Hmitou, I., Rocques, N., Lecoin, L., Druillenec, S., Felder-Schmittbuhl, M.P. and Eyche, A. (2006) Cell context reveals a dual role for Maf in oncogenesis. *Oncogene*, **25**, 1299–1310.
32. Zhang, W. and Fuller, G. (2007) IGFBP2 as a brain tumor oncogene. *Cancer Biol. Ther.*, **6**, 995–996.
33. Zhang, C.C., Kaba, M., Iizuka, S., Huynh, H. and Lodish, H.F. (2008) Angiopoietin-like 5 and IGFBP2 stimulate ex vivo expansion of human cord blood hematopoietic stem cells as assayed by NOD/SCID transplantation. *Blood*, **111**, 3415–3423.
34. Boshoff, C. (1998) Kaposi's sarcoma. Coupling herpesvirus to angiogenesis. *Nature*, **391**, 24–25.
35. Krishnan, K., Salomonis, N. and Guo, S. (2008) Identification of Spt5 target genes in zebrafish development reveals its dual activity in vivo. *PLoS One*, **3**, e3621.
36. Bais, C., Van Geelen, A., Eroles, P., Mutlu, A., Chiozzini, C., Dias, S., Silverstein, R.L., Rafii, S. and Mesri, E.A. (2003) Kaposi's sarcoma associated herpesvirus G protein-coupled receptor immortalizes human endothelial cells by activation of the VEGF receptor-2/ KDR. *Cancer Cell*, **3**, 131–143.
37. Wang, L., Dittmer, D.P., Tomlinson, C.C., Fakhari, F.D. and Damania, B. (2006) Immortalization of primary endothelial cells by the K1 protein of Kaposi's sarcoma-associated herpesvirus. *Cancer Res.*, **66**, 3658–3666.
38. Qian, L.W., Xie, J., Ye, F. and Gao, S.J. (2007) Kaposi's sarcoma-associated herpesvirus infection promotes invasion of primary human umbilical vein endothelial cells by inducing matrix metalloproteinases. *J. Virol.*, **81**, 7001–7010.
39. Rose, P.P., Carroll, J.M., Carroll, P.A., DeFilippis, V.R., Lagunoff, M., Moses, A.V., Roberts, C.T. Jr. and Fruh, K. (2007) The insulin receptor is essential for virus-induced tumorigenesis of Kaposi's sarcoma. *Oncogene*, **26**, 1995–2005.
40. Yamamoto, Y., Irie, K., Asada, M., Mino, A., Mandai, K. and Takai, Y. (2004) Direct binding of the human homologue of the Drosophila disc large tumor suppressor gene to seven-pass transmembrane proteins, tumor endothelial marker 5 (TEM5), and a novel TEM5-like protein. *Oncogene*, **23**, 3889–3897.
41. Graef, I.A., Chen, F., Chen, L., Kuo, A. and Crabtree, G.R. (2001) Signals transduced by Ca(2+)/calcineurin and NFATc3/c4 pattern the developing vasculature. *Cell*, **105**, 863–875.
42. Kwapiszewska, G., Wygrecka, M., Marsh, L.M., Schmitt, S., Trosser, R., Wilhelm, J., Helmus, K., Eul, B., Zakrzewicz, A., Ghofrani, H.A. *et al.* (2008) Fhl-1, a new key protein in pulmonary hypertension. *Circulation*, **118**, 1183–1194.
43. Wang, L., Tang, H., Thayanyithy, V., Subramanian, S., Oberg, A.L., Cunningham, J.M., Cerhan, J.R., Steer, C.J. and Thibodeau, S.N. (2009) Gene networks and microRNAs implicated in aggressive prostate cancer. *Cancer Res.*, **69**, 9490–9497.
44. Huh, S.H. and Ornitz, D.M. (2010) Beta-catenin deficiency causes DiGeorge syndrome-like phenotypes through regulation of Tbx1. *Development*, **137**, 1137–1147.
45. Ho, P.J., Yen, M.L., Lin, J.D., Chen, L.S., Hu, H.I., Yeh, C.K., Peng, C.Y., Lin, C.Y., Yet, S.F. and Yen, B.L. (2010) Endogenous KLF4 expression in human fetal endothelial cells allows for reprogramming to pluripotency with just OCT3/4 and SOX2—brief report. *Arterioscler. Thromb. Vasc. Biol.*, **30**, 1905–1907.

46. Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A. and Weinberg, R.A. (2008) An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.*, **40**, 499–507.
47. Tsai, Y.H., Wu, M.F., Wu, Y.H., Chang, S.J., Lin, S.F., Sharp, T.V. and Wang, H.W. (2009) The M type K15 protein of Kaposi's sarcoma-associated herpesvirus regulates microRNA expression via its SH2-binding motif to induce cell migration and invasion. *J. Virol.*, **83**, 622–632.
48. Poliseño, L., Tuccoli, A., Mariani, L., Evangelista, M., Citti, L., Woods, K., Mercatanti, A., Hammond, S. and Rainaldi, G. (2006) MicroRNAs modulate the angiogenic properties of HUVECs. *Blood*, **108**, 3068–3071.
49. O'Hara, A.J., Wang, L., Dezube, B.J., Harrington, W.J. Jr., Damania, B. and Dittmer, D.P. (2009) Tumor suppressor microRNAs are underrepresented in primary effusion lymphoma and Kaposi sarcoma. *Blood*, **113**, 5938–5941.
50. Gregory, R.I. and Shiekhattar, R. (2005) MicroRNA biogenesis and cancer. *Cancer Res.*, **65**, 3509–3512.
51. Wang, Z., Gerstein, M. and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.*, **10**, 57–63.
52. Byun, M., Abhyankar, A., Lelarge, V., Plancoulaine, S., Palanduz, A., Telhan, L., Boisson, B., Picard, C., Dewell, S., Zhao, C. *et al.* (2010) Whole-exome sequencing-based discovery of STIM1 deficiency in a child with fatal classic Kaposi sarcoma. *J. Exp. Med.*, **207**, 2307–2312.
53. Hong, Y.K., Foreman, K., Shin, J.W., Hirakawa, S., Curry, C.L., Sage, D.R., Libermann, T., Dezube, B.J., Fingerth, J.D. and Detmar, M. (2004) Lymphatic reprogramming of blood vascular endothelium by Kaposi sarcoma-associated herpesvirus. *Nat. Genet.*, **36**, 683–685.
54. Chandriani, S., Xu, Y. and Ganem, D. (2010) The lytic transcriptome of Kaposi's sarcoma-associated herpesvirus reveals extensive transcription of noncoding regions, including regions antisense to important genes. *J. Virol.*, **84**, 7934–7942.
55. Whichard, Z.L., Sarkar, C.A., Kimmel, M. and Corey, S.J. (2010) Hematopoiesis and its disorders: a systems biology approach. *Blood*, **115**, 2339–2347.
56. Johnston, J.B., Rahaman, M.M. and McFadden, G. (2007) Strategies that modulate inflammasomes: insights from host-pathogen interactions. *Semin. Immunopathol.*, **29**, 261–274.
57. Gregory, S.M., Davis, B.K., West, J.A., Taxman, D.J., Matsuzawa, S., Reed, J.C., Ting, J.P. and Damania, B. (2011) Discovery of a viral NLR homolog that inhibits the inflammasome. *Science*, **331**, 330–334.
58. Shultz, J.C., Goehe, R.W., Wijesinghe, D.S., Murudkar, C., Hawkins, A.J., Shay, J.W., Minna, J.D. and Chalfant, C.E. (2010) Alternative splicing of caspase 9 is modulated by the phosphoinositide 3-kinase/Akt pathway via phosphorylation of SRp30a. *Cancer Res.*, **70**, 9185–9196.
59. Seol, D.W. and Billiar, T.R. (1999) A caspase-9 variant missing the catalytic site is an endogenous inhibitor of apoptosis. *J. Biol. Chem.*, **274**, 2072–2076.
60. Srinivasula, S.M., Ahmad, M., Guo, Y., Zhan, Y., Lazebnik, Y., Fernandes-Alnemri, T. and Alnemri, E.S. (1999) Identification of an endogenous dominant-negative short isoform of caspase-9 that can regulate apoptosis. *Cancer Res.*, **59**, 999–1002.
61. Kaiser, S., Park, Y.K., Franklin, J.L., Halberg, R.B., Yu, M., Jessen, W.J., Freudenberg, J., Chen, X., Haigis, K., Jegga, A.G. *et al.* (2007) Transcriptional recapitulation and subversion of embryonic colon development by mouse colon tumor models and human colon cancer. *Genome Biol.*, **8**, R131.
62. Hu, M. and Shivdasani, R.A. (2005) Overlapping gene expression in fetal mouse intestine development and human colorectal cancer. *Cancer Res.*, **65**, 8715–8722.
63. Kho, A.T., Zhao, Q., Cai, Z., Butte, A.J., Kim, J.Y., Pomeroy, S.L., Rowitch, D.H. and Kohane, I.S. (2004) Conserved mechanisms across development and tumorigenesis revealed by a mouse development perspective of human cancers. *Genes Dev.*, **18**, 629–640.
64. Chang, S.J., Wang, T.Y., Tsai, C.Y., Hu, T.F., Chang, M.D. and Wang, H.W. (2009) Increased epithelial stem cell traits in advanced endometrial endometrioid carcinoma. *BMC Genomics*, **10**, 613.
65. Liu, R., Li, X., Tulpule, A., Zhou, Y., Sehnet, J.S., Zhang, S., Lee, J.S., Chaudhary, P.M., Jung, J. and Gill, P.S. (2010) KSHV-induced notch components render endothelial and mural cell characteristics and cell survival. *Blood*, **115**, 887–895.