

Identifying differential expression genes and single nucleotide variations using RNA-seq in metastatic melanoma

D. Liu¹, Z.G. Zhao¹, Z.L. Jiao² and H.J. Li¹

¹Department of Dermatology, PLA General Hospital, Beijing, China

²Department of Dermatology, Clinic of Navy Headquarters, Beijing, China

Corresponding author: H.J. Li

E-mail: HengjinLi@hotmail.com

Genet. Mol. Res. 13 (4): 8153-8162 (2014)

Received August 12, 2013

Accepted January 13, 2014

Published October 7, 2014

DOI <http://dx.doi.org/10.4238/2014.October.7.10>

ABSTRACT. Melanoma is a malignant tumor and one of the most frequent metastatic cancers. This study was conducted to identify differential expression genes (DEGs) and single nucleotide variations (SNVs) in metastatic melanoma. We analyzed microarray data of GSE23056 downloaded from the Gene Expression Omnibus, including two normal samples (skinN1 and skinN2) and 2 metastatic melanoma samples (skinT and lymphT). We not only compared DEGs in metastatic melanoma samples with normal samples (lymphT_skinN and skinT_skinN), but also compared DEGs between two metastatic melanoma types (lymphT_skinT). SNVs were identified by using Burrows-Wheeler Aligner and Cufflinks in metastatic melanoma samples using RNA-seq. Sequence Alignment/Map tools and the ANNOVAR software were used to analyze and annotate SNVs. We identified 18 significantly common DEGs in lymphT_skinN and skinT_skinN and one common gene, *YBX1*, in lymphT_skinN, skinT_skinN, and lymphT_skinT. We identified 49,534, 48,118, 63,812, and 33,096 SNVs in skinN1, skinN2,

skinT, and lymphT, respectively. Twenty-nine SNVs were located in exonic regions of two DEGs, *HLA-B* and *TSPAN10*. SNVs that exist only in tumors were located in *MARVELD1*, *SLC16A3*, and *VAV3*. The DEGs screened in our study are potential biomarkers for metastatic melanoma therapy.

Key words: Metastatic melanoma; Differential expression genes; RNA-seq; Single nucleotide variations

INTRODUCTION

Metastatic melanoma is an aggressive disease (Hodi et al., 2010) that can spread to the liver, bones, abdomen, or distant lymph nodes. The incidence of metastatic melanoma and associated number of deaths has increased substantially over the past several decades (Lens and Dawes, 2004). Few treatments are available for metastatic melanoma; 2 therapies have been approved by the Food and Drug Administration, including high-dose interleukin-2 and dacarbaine (Comis, 1976; Atkins et al., 1999). However, these therapies are associated with low responses and short survival times (Middleton et al., 2000; Bedikian et al., 2006). Thus, identifying new biomarkers would be useful for developing new metastatic melanoma therapies and to facilitate drug discovery.

In metastatic melanoma, differentially expressed genes were identified between cutaneous melanomas with differing invasive phenotypes by gene expression profiling (Bittner et al., 2000). Comparative oncogenomics enabled identification of *NEDD9*, which is overexpressed in 35-50% of metastatic human melanomas and may enhance invasion and metastasis in human melanoma (Kim et al., 2006). *WNT5A* is expressed in human melanoma and was shown to directly affect cell motility and the invasion of metastatic melanoma (Weeraratna et al., 2002). Clark et al. (2000) demonstrated that *RhoC* is overexpressed in metastatic melanoma and enhances metastasis. During the late stage of many tumors, genomic alterations are so extensive that individual genes show large amounts of variation. For example, *BRAF* and *N-RAS* mutations are common in melanoma metastases (Gorden et al., 2003); the *BRAF* mutation may account for 60% of melanomas (Smalley et al., 2006).

Single nucleotide polymorphisms (SNPs) can be used for high-volume, rapid determination of genome-wide allelic information for a DNA sample (Lindblad-Toh et al., 2000; Kennedy et al., 2003). Whole-genome sequencing is an effective method for identifying SNPs (Ley et al., 2008). However, this method is very costly. RNA-seq, a new sequencing method that has been developed in recent years, uses next-generation sequencing technologies to sequence cDNA in order to determine RNA content in a sample (Nagalakshmi et al., 2010), which can significantly reduce costs and has been widely adopted in studies of cancer (Wang et al., 2009b). RNA-seq analysis can be used to determine differential expression of genes, single nucleotide variation (SNV) information, non-coding RNAs, gene fusions, and mRNA mutations or editing.

In this study, we identified differential expression genes (DEGs) and SNVs in normal, non-metastatic, and metastatic samples using RNA-seq. Next, we analyzed SNVs and identified the DEGs in which the SNVs were located.

MATERIAL AND METHODS

Microarray data

We downloaded microarray and sequencing data from the Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) database. The accession number is GSE23056 (Valsesia et al., 2011). Four samples were selected for analysis, including 2 normal samples and 2 metastatic samples. Sample information is shown in Table 1. Melanoma cell lines were obtained from patients with metastatic melanoma.

Table 1. Detailed information for 4 samples.

Label	Cell type	Source tissue	Accession No.
skinN1, N2	Normal cultured melanocytes	Skin	GSM566260
skinT	Cells derived from metastatic melanoma	Skin	GSM566266
lymphT	Cells derived from metastatic melanoma	Lymph node	GSM566265

DEGs

We identified significantly DEGs in 2 metastatic melanoma samples compared with normal samples when the fragments per kilobase of transcript per million fragments mapped reads (FPKM) was larger than 2 in both samples, and the P value was less than 0.05. We not only searched DEGs between metastatic melanoma samples and normal samples, but also analyzed the expression levels of genes between the 2 metastatic melanoma types.

Alignment and assembly

To prevent high-quality reads from being rejected during the quality filtering or assembly processes, we trimmed bases from the poor-quality ends of reads. We first aligned our reads to the reference hg19 human genome, as well as to a library of synthetic exon junction sequences. The Burrows-Wheeler Aligner (BWA) program was used as an alignment tool to align short reads to long reference sequences with little memory and long sequences, up to 1 Mb, against mammalian-sized genomes (Li and Durbin, 2009).

Next, we used the Cufflinks software (Trapnell et al., 2010) for transcript assembly, which was used to estimate the relative abundances of these transcripts based on default parameters. Aligned RNA-seq reads were accepted and assembled into a parsimonious set of transcripts. Cuffmerge was used to merge transcript assemblies.

Genome-wide SNV analysis

We obtained RefSeq gene annotations tables for the hg19 human genome from the University of California Santa Cruz (UCSC) Genome Bioinformatics website (<http://genome.ucsc.edu>). Sequence Alignment/Map (SAM) tools (Li et al., 2009) were used to parse and manipulate the alignments. SNVs were annotated using the ANNOVAR (functional annotation of genetic variants from high-throughput sequencing data) tool to examine the functional con-

sequence on genes, identify variants in conserved regions, and report functional importance scores (Wang et al., 2010).

RESULTS

DEGs

We not only analyzed DEGs in melanoma cell lines (lymphT and skinT samples) compared with normal samples (skinN sample), but also identified DEGs in the lymphT sample compared with the skinT sample. Next, we compared DEGs among lymphT_skinN, skinT_skinN, and lymphT_skinT (Figure 1). A total of 289 DEGs were identified; 92, 108, and 133 DEGs were found in lymphT_skinN, skinT_skinN, and lymphT_skinT, respectively. Eighteen common DEGs were found between lymphT_skinN and skinT_skinN. The only common gene among lymphT_skinN, skinT_skinN, and lymphT_skinT was *YBX1*. FPKM values for *YBX1* in the 2 normal skin samples were 15.5839 and 12.2385, respectively. The FPKM value was 97.4813 in the skin tumor sample and was 691.043 in the lymph tumor sample (Table 2).

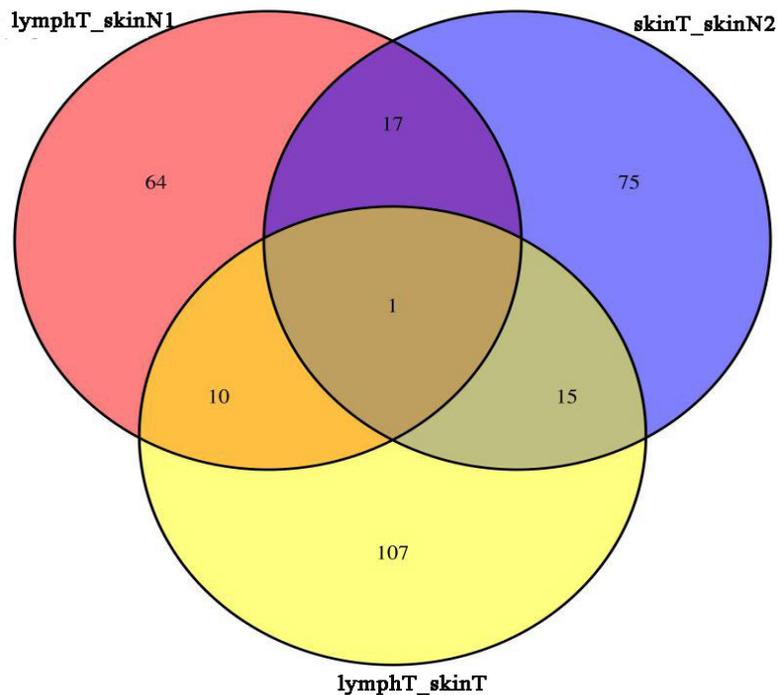


Figure 1. Venn diagram of differentially expression genes (DEGs) in 3 paired samples. LymphT_skinN refers to DEGs in lymph node metastatic melanoma comparing with skin normal cells. SkinT_skinN refers to DEGs in skin tumor comparing with normal skin cells. lymphT_skinT refers to DEGs in lymph node metastatic melanoma comparing with skin tumor cells.

Table 2. Fragments per kilobase of transcript per million fragments mapped reads and log₂ (fold-change) of eighteen common differentially expressed genes between lymphT₂skinN and skinT₂skinN.

Genes	skinN_FPKM	lymphT ₂ skinN_FPKM	log ₂ (fold-change 1)	P value 1	skinN_FPKM	skinT ₂ skinN_FPKM	log ₂ (fold-change 2)	P value 2
<i>CD99</i>	221.722	23.3749	-3.24572	0.033689	378.906	33.6325	-3.49391	0.036382
<i>MARVELD1</i>	41.0166	8.7331	-2.23164	0.045264	54.8552	8.7635	-2.64605	0.025787
<i>P4G1</i>	22.7607	4.07019	-2.48337	0.047157	24.8293	4.58538	-2.43693	0.03401
<i>MMP2</i>	239.792	5.02692	-5.57597	0.018325	342.348	3.51255	-6.6068	0.025526
<i>YBX1</i>	15.5839	691.043	5.47065	0.002083	12.2385	97.4813	2.9937	0.030491
<i>AVPII</i>	48.9394	2.09112	-4.54865	0.00403	44.8479	5.0852	-3.14067	0.031379
<i>HLA-B</i>	169.791	10.8462	-3.96849	0.002254	311.309	10.488	-4.89154	0.014463
<i>IFIT3</i>	114.486	4.415	-4.69661	0.000345	125.333	9.43939	-3.73093	0.028654
<i>SLC18B1</i>	109.539	7.28807	-3.90976	0.017765	94.1017	3.34492	-4.81418	0.00906
<i>SLC16A3</i>	4.97349	70.0353	3.81575	0.002468	6.73548	101.581	3.9147	0.011475
<i>TSPN10</i>	175.499	11.904	-3.88194	0.01832	145.712	5.14502	-4.8238	0.0007
<i>R4B32</i>	107.523	13.3174	-3.01326	0.039851	97.2116	11.1146	-3.12868	0.031072
<i>TAPI</i>	84.3254	4.56255	-4.20806	0.005209	109.447	7.54007	-3.85951	0.041309
<i>C414</i>	59.588	4.3427	-3.77836	0.033433	57.4191	3.67101	-3.96728	0.010681
<i>CHMP1B</i>	48.0368	5.7966	-3.05086	0.029988	59.0335	8.46058	-2.8027	0.029256
<i>IFI27</i>	452.113	9.70971	-5.54111	0.007452	110.221	5.44314	-4.33982	0.026992
<i>FAV3</i>	55.4253	8.17421	-2.76139	0.017914	50.6756	4.83075	-3.39097	0.008433
<i>KCNAB2</i>	559.704	25.1538	-4.47582	0.018574	417.933	13.8699	-4.91325	0.01969

Overview of genome-wide SNVs

We identified 49,534, 48,118, 63,812, and 33,096 SNVs in skinN1, skinN2, skinT, and lymphT, respectively (Figure 2). A Venn diagram shows the filtered sample-specific SNVs. A total of 11,697, 10,606, 11,644, and 26,173 sample-specific SNVs were identified in skinN1, skinN2, skinT, and lymphT, respectively. A total of 3196 SNVs were found only in tumor samples, indicating that these SNVs are critical factors in the development meta-static melanoma.

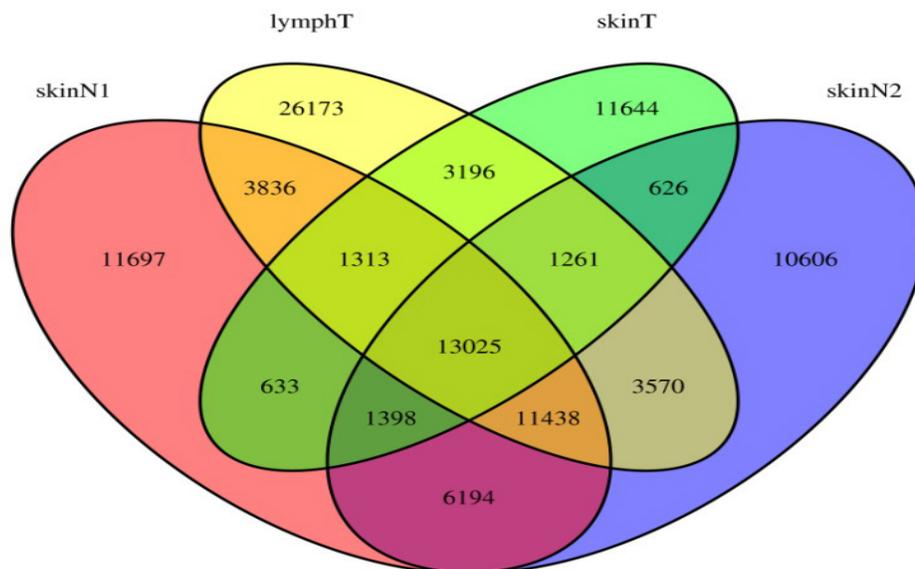


Figure 2. Venn diagram of single nucleotide variations in 4 samples (skinN1, skinN2, skinT, and lymphT).

Differential SNVs in DEG regions

To identify the distributions of SNVs in chromosome, we counted the density of SNVs in each sample and displayed this information using Circos (Figure 3). We found that 98 SNVs were related to the 14 DEGs listed in [Table S1](#). Forty SNVs were located in the exonic region. Twenty-nine SNVs were located in the exonic regions of the 2 DEGs, major histocompatibility complex, class I, B (*HLA-B*) and tetraspanin 10 (*TSPAN10*). SNVs existing only in tumor samples were located in marvel domain containing 1 (*MARVELD1*), solute carrier family 16, member 3 (*SLC16A3*), and *VAV3*. *MARVELD1* possessed a point mutation in the 3'-untranslated region. *SLC16A3* contains a point mutation in its exon, while *VAV3* contains a downstream point mutation.

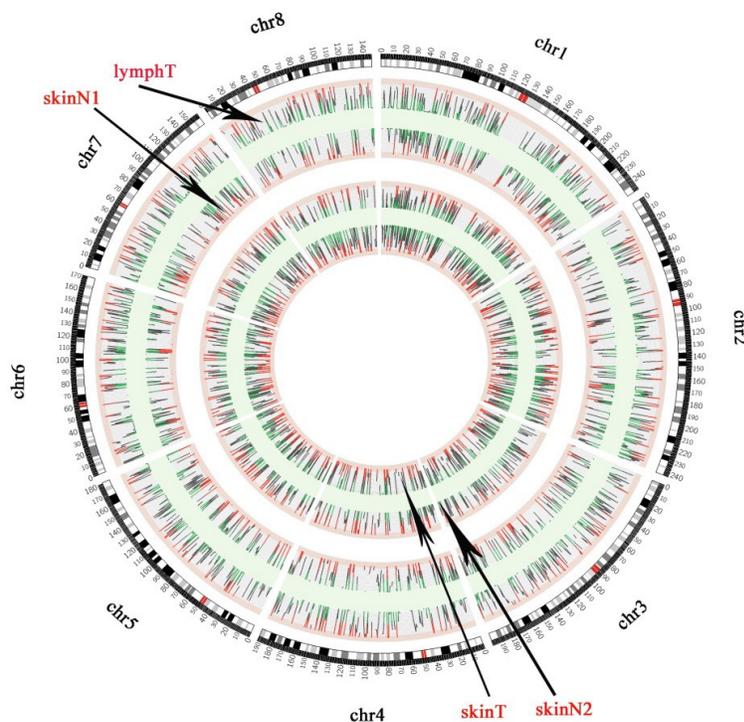


Figure 3. Single nucleotide variation distributions in chromosome 1-8. The red line indicates low density in the region ($<0.003/\text{kb}$) and the green line indicates high density in the region ($>0.024/\text{kb}$). The directions of skinN2 and lymphT are toward the center of the circle, and the directions of skinN1 and skinT are away from the center of the circle.

DISCUSSION

In this study, we identified DEGs in metastatic melanoma and compared these DEGS with normal samples and SNVs in both tumor and normal samples using RNA-seq. In addition, we analyzed the distribution of SNVs in the genome and their relationship with DEGs.

A previous study reported that *YBX1*, which is involved in malignant transformation of melanocytes, could stimulate proliferation, survival, tumor invasion, and chemoresistance (Schitteck et al., 2007). Moreover, downregulation of *YBX1* was hypothesized to lead to an increased rate of apoptotic cell death and a reduced rate of proliferation. Our results correspond well with this hypothesis. In our study, *YBX1* was shown to be expressed in both tumor types, with its expression level varying between metastatic tumor types; the expression level is high in advanced stages, but low in early stages. Reduced *YBX1* expression shows decreased resistance to chemotherapy for melanoma. This information may lead to development of a drug that can reduce the expression of *YBX1* as part of melanoma therapy. Thus, *YBX1* may be a biomarker for metastatic melanoma.

Mutations in noncoding regions may disrupt cis-regulatory elements, which can alter transcript levels; however, the primary causes of genetic diseases were found to be related

to protein coding regions, which can alter protein form (Chepelev et al., 2009). In our study, we identified 29 SNVs located in exonic regions of the 2 DEGs in normal samples, *HLA-B* and *TSPAN10*. *HLA-B* was shown to be downregulated in human melanoma (Griffioen et al., 1999a,b). A recent study found that downregulation of *TSPAN10* resulted in reduced melanoma migration (Seong et al., 2012). These results indicate that DEGs *HLA-B* and *TSPAN10* are related to metastatic melanoma.

SNVs only exist in tumors located in *MARVELD1*, *SLC16A3*, and *VAV3*. *MARVELD1* was widely expressed in normal human tissues and was downregulated in multiple primary tumors derived from the breast, ovary, kidney, and liver, among others (Wang et al., 2009a). Zeng et al. (2011) examined the function of *MARVELD1* in mice and demonstrated that overexpression of mouse *MARVELD1* inhibited cell proliferation, G1-phase arrest, and cell migration. In hepatocellular carcinoma, *MARVELD1* is a tumor suppressor that regulates tumor growth and may be a potential biomarker for hepatocellular carcinoma therapy (Yu et al., 2012).

SLC16A3, which belongs to the proton-linked transmembrane monocarboxylate transporter (MCT) family, is related to the transportation of metabolic monocarboxylates such as lactate, acetate, and pyruvate (Halestrap and Price, 1999). *SLC16A3* is overexpressed in breast cancers compared with normal breast tissue (Whitaker-Menezes et al., 2011) and is up-regulated in bladder tumors, which showed no response to methotrexate, vinblastine, cisplatin vinblastine, doxorubicin, and cisplatin regimens (Takata et al., 2005).

VAV3 is a member of the Vav family and is encoded by *VAV3* (Faccio et al., 2005). In prostate cancer, the *VAV3* oncogene is overexpressed and regulates cell growth and androgen receptor function (Dong et al., 2006). Lee et al. (2008) also observed overexpression of *VAV3* in breast cancer specimens. It also has been reported that activation of *VAV3* in tumors resulted in the development of anaplastic large cell lymphomas (Colomba et al., 2008).

In summary, DEGs and SNVs were identified in metastatic melanoma. *YBX1*, *HLA-B*, and *TSPAN10* were identified in our analysis, which is consistent with previous studies in melanoma. *MARVELD1*, *SLC16A3*, and *VAV3* were newly reported for melanoma, and they may be potential biomarkers for metastatic melanoma therapy. However, further experiments are required to confirm our results.

[Supplementary material](#)

REFERENCES

- Atkins MB, Lotze MT, Dutcher JP, Fisher RI, et al. (1999). High-dose recombinant interleukin 2 therapy for patients with metastatic melanoma: analysis of 270 patients treated between 1985 and 1993. *J. Clin. Oncol.* 17: 2105-2116.
- Bedikian AY, Millward M, Pehamberger H, Conry R, et al. (2006). Bcl-2 antisense (oblimersen sodium) plus dacarbazine in patients with advanced melanoma: the Oblimersen Melanoma Study Group. *J. Clin. Oncol.* 24: 4738-4745.
- Bittner M, Meltzer P, Chen Y, Jiang Y, et al. (2000). Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406: 536-540.
- Chepelev I, Wei G, Tang Q and Zhao K (2009). Detection of single nucleotide variations in expressed exons of the human genome using RNA-Seq. *Nucleic Acids Res.* 37: e106.
- Clark EA, Golub TR, Lander ES and Hynes RO (2000). Genomic analysis of metastasis reveals an essential role for RhoC. *Nature* 406: 532-535.
- Colomba A, Courilleau D, Ramel D, Billadeau DD, et al. (2008). Activation of Rac1 and the exchange factor Vav3 are involved in NPM-ALK signaling in anaplastic large cell lymphomas. *Oncogene* 27: 2728-2736.

- Comis RL (1976). DTIC (NSC-45388) in malignant melanoma: a perspective. *Cancer Treat. Rep.* 60: 165-176.
- Dong Z, Liu Y, Lu S, Wang A, et al. (2006). Vav3 oncogene is overexpressed and regulates cell growth and androgen receptor activity in human prostate cancer. *Mol. Endocrinol.* 20: 2315-2325.
- Faccio R, Teitelbaum SL, Fujikawa K, Chappel J, et al. (2005). Vav3 regulates osteoclast function and bone mass. *Nat. Med.* 11: 284-290.
- Gorden A, Osman I, Gai W, He D, et al. (2003). Analysis of BRAF and N-RAS mutations in metastatic melanoma tissues. *Cancer Res.* 63: 3955-3957.
- Griffioen M, Ouwerkerk IJ, Harten V and Schrier PI (1999a). HLA-B downregulation in human melanoma is mediated by sequences located downstream of the transcription-initiation site. *Int. J. Cancer* 80: 573-580.
- Griffioen M, Ouwerkerk IJ, Harten V, Gobin SJ, et al. (1999b). Role of the interferon-stimulated response element in HLA-B downregulation in human melanoma cell lines. *Immunogenetics* 49: 287-294.
- Halestrap AP and Price NT (1999). The proton-linked monocarboxylate transporter (MCT) family: structure, function and regulation. *Biochem. J.* 343 Pt 2: 281-299.
- Hodi FS, O'Day SJ, McDermott DF, Weber RW, et al. (2010). Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* 363: 711-723.
- Kennedy GC, Matsuzaki H, Dong S, Liu WM, et al. (2003). Large-scale genotyping of complex DNA. *Nat. Biotechnol.* 21: 1233-1237.
- Kim M, Gans JD, Nogueira C, Wang A, et al. (2006). Comparative oncogenomics identifies NEDD9 as a melanoma metastasis gene. *Cell* 125: 1269-1281.
- Lee K, Liu Y, Mo JQ, Zhang J, et al. (2008). Vav3 oncogene activates estrogen receptor and its overexpression may be involved in human breast cancer. *BMC Cancer* 8: 158.
- Lens MB and Dawes M (2004). Global perspectives of contemporary epidemiological trends of cutaneous malignant melanoma. *Br. J. Dermatol.* 150: 179-185.
- Ley TJ, Mardis ER, Ding L, Fulton B, et al. (2008). DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* 456: 66-72.
- Li H and Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754-1760.
- Li H, Handsaker B, Wysoker A, Fennell T, et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078-2079.
- Lindblad-Toh K, Tanenbaum DM, Daly MJ, Winchester E, et al. (2000). Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. *Nat. Biotechnol.* 18: 1001-1005.
- Middleton MR, Grob JJ, Aaronson N, Fierbeck G, et al. (2000). Randomized phase III study of temozolomide versus dacarbazine in the treatment of patients with advanced metastatic malignant melanoma. *J. Clin. Oncol.* 18: 158-166.
- Nagalakshmi U, Waern K and Snyder M (2010). RNA-Seq: a method for comprehensive transcriptome analysis. *Curr. Protoc. Mol. Biol.* Chapter 4: Unit-13.
- Schitteck B, Psenner K, Sauer B, Meier F, et al. (2007). The increased expression of Y box-binding protein 1 in melanoma stimulates proliferation and tumor invasion, antagonizes apoptosis and enhances chemoresistance. *Int. J. Cancer* 120: 2110-2118.
- Seong I, Min HJ, Lee JH, Yeo CY, et al. (2012). Sox10 controls migration of B16F10 melanoma cells through multiple regulatory target genes. *PLoS One* 7: e31477.
- Smalley KS, Haass NK, Brafford PA, Lioni M, et al. (2006). Multiple signaling pathways must be targeted to overcome drug resistance in cell lines derived from melanoma metastases. *Mol. Cancer Ther.* 5: 1136-1144.
- Takata R, Katagiri T, Kanehira M, Tsunoda T, et al. (2005). Predicting response to methotrexate, vinblastine, doxorubicin, and cisplatin neoadjuvant chemotherapy for bladder cancers through genome-wide gene expression profiling. *Clin. Cancer Res.* 11: 2625-2636.
- Trapnell C, Williams BA, Pertea G, Mortazavi A, et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28: 511-515.
- Valsesia A, Rimoldi D, Martinet D, Ibberson M, et al. (2011). Network-guided analysis of genes with altered somatic copy number and gene expression reveals pathways commonly perturbed in metastatic melanoma. *PLoS One* 6: e18369.
- Wang K, Li M and Hakonarson H (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38: e164.
- Wang S, Li Y, Han F, Hu J, et al. (2009a). Identification and characterization of MARVELD1, a novel nuclear protein that is downregulated in multiple cancers and silenced by DNA methylation. *Cancer Lett.* 282: 77-86.
- Wang Z, Gerstein M and Snyder M (2009b). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10: 57-63.

- Weeraratna AT, Jiang Y, Hostetter G, Rosenblatt K, et al. (2002). Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma. *Cancer Cell* 1: 279-288.
- Whitaker-Menezes D, Martinez-Outschoorn UE, Lin Z, Ertel A, et al. (2011). Evidence for a stromal-epithelial “lactate shuttle” in human tumors: MCT4 is a marker of oxidative stress in cancer-associated fibroblasts. *Cell Cycle* 10: 1772-1783.
- Yu Y, Zhang Y, Hu J, Zhang H, et al. (2012). MARVELD1 inhibited cell proliferation and enhance chemosensitivity via increasing expression of p53 and p16 in hepatocellular carcinoma. *Cancer Sci.* 103: 716-722.
- Zeng F, Tian Y, Shi S, Wu Q, et al. (2011). Identification of mouse MARVELD1 as a microtubule associated protein that inhibits cell cycle progression and migration. *Mol. Cells* 31: 267-274.