DOI: 10.1111/exd.12054
www.blackwellpublishing.com/EXD

High-throughput sequencing of the melanoma genome

Manfred Kunz[1], Michael Dannemann[2] and Janet Kelso[2]

[1]Department of Dermatology, Venereology and Allergology, University of Leipzig, Leipzig, Germany; [2]Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Correspondence: Manfred Kunz, MD, Department of Dermatology, Venereology and Allergology, University of Leipzig, Philipp-Rosenthal-Str. 23, 04103 Leipzig, Germany, Tel.: +49-341-9718610, Fax: +49-341-9718609, e-mail: manfred.kunz@medizin.uni-leipzig.de

Abstract: Next-generation sequencing technologies are now common for whole-genome, whole-exome and whole-transcriptome sequencing (RNA-seq) of tumors to identify point mutations, structural or copy number alterations and changes in gene expression. A substantial number of studies have already been performed for melanoma. One study analysed eight melanoma cell lines with RNA-Seq technology and identified 11 novel melanoma gene fusions. Whole-exome sequencing of seven melanoma cell lines identified overlapping gain of function mutations in *MAP2K1* (*MEK1*) and *MAP2K2* (*MEK2*) genes. Integrative sequencing of cutaneous melanoma metastases using different sequencing platforms revealed a new somatic point mutation in *HRAS* and a structural rearrangement affecting *CDKN2C* (a CDK4 inhibitor). These latter sequencing-based discoveries may be used to motivate the inclusion of the affected patients into clinical trials with specific signalling pathway inhibitors. Taken together, we are at the beginning of an era with new sequencing technologies providing a more comprehensive view of cancer mutational landscapes and hereby a better understanding of their pathogenesis. This will also open interesting perspectives for new treatment approaches and clinical trial designs.

Key words: genetics – molecular biology – mutations – tumor

Accepted for publication 19 October 2012

Introduction

Melanoma is a malignant tumor of rapidly increasing incidence and high metastatic potential (1). Much progress has been made in recent years in our understanding of the genetic basis of this tumor. Activating mutations in *BRAF* and *NRAS* oncogenes have been identified in the majority of melanomas, and both are mutually exclusive (2,3). There is substantial evidence that improved overall survival of melanoma patients may be achieved by treatment with recently developed specific BRAF inhibitors such as vemurafenib, at least in patients with tumors that carry activating *BRAF* V600E mutations (4,5). Although of high clinical relevance, many of the patients (~50%) with initial treatment response to specific BRAF inhibitors experience recurrences. Recurrences due to BRAF inhibitor resistance may be due to secondary *NRAS* mutations, overexpression of platelet-derived growth factor receptor (PDGFRβ1) or alternative-spliced BRAF(V600E) variants, but may also involve other mechanisms such as MAP3K8 (COT) and stromal cell secretion of hepatocyte growth factor (6–11).

Moreover, *c-KIT* mutations are found in 2% of skin melanomas, and a significant percentage of patients with metastatic lesions with *c-KIT* mutations respond to treatment with imatinib, a multikinase inhibitor targeting c-KIT (16% durable responses compared with 8–10% responses after classical chemotherapy) (12,13). Phosphatase and tensin homologue (PTEN) is a protein and lipid phosphatase that inhibits activation of Akt kinase by negative interference with phosphoinositide 3-kinase (PI3K). Loss of heterozygosity of the chromosomal locus of *PTEN* was demonstrated in up to 30% of melanomas, and approximately 10% of melanomas have *PTEN* mutations (14). Interestingly, expression of mutated *BRAF* together with *PTEN* silencing in mice induced melanomas with 100% penetrance (15). PI3K/Akt signalling pathway is also activated in non-melanoma skin cancer (16). In addition to targeting mutated pathway molecules, signalling assays may be used to support treatment decisions (17).

Together, although the pathogenic concept focussing on BRAF, c-KIT and PTEN pathways appears to be promising in subsets of melanomas, further search for new pathogenic mechanisms and genetic alterations is needed. In particular, the molecular mechanisms for treatment resistance and high recurrence rates after targeted therapy with BRAF inhibitors are still enigmatic. Genetic tumor heterogeneity may be one reason for this, which would have significant consequences for future therapeutic approaches.

Next-generation sequencing technologies

The first next-generation sequencing (NGS) platform was released in 2005, and a number of different methodologies and platforms have been established and are described in detail in a series of recent reviews (18–21). Commonly used NGS platforms are produced by Roche, Illumina, Life/Applied Biosystems, Helicos BioSciences and Pacific Biosciences. The major distinguishing feature of the NGS technologies is their ability to generate many thousands/millions sequence reads in parallel (20,22). NGS also differs from Sanger sequencing in library/template preparation, sequencing reaction and sequence detection, which will be described briefly here.

For library preparation in NGS, adaptors are ligated to fragmented DNA and then amplified by PCR before being sequenced. Two major methods are typically used to generate amplified templates, one is emulsion PCR (23) and the other solid-phase amplification (24). In emulsion PCR, a library of small DNA fragments is used to prepare templates which are bound to beads in a ratio of one DNA molecule per bead (Fig. 1). Amplified DNA

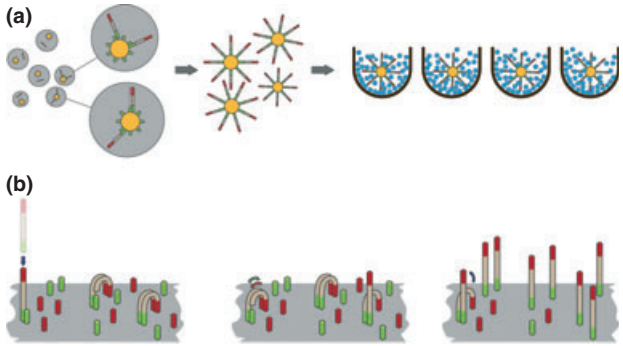

Figure 1. Examples of next-generation sequencing technologies. (a) Emulsion PCR of a library of small DNA fragments (with a ratio of one fragment per bead) is performed in small aqueous droplets of an oil–aqueous emulsion to generate bead-bound templates for sequencing. Each beat carries a large number of specific DNA templates which are sequenced by pyrosequencing (pyrosequencing enzymes indicated as blue beads) for generation of the fluorescence signal. (b) Sequencing templates may also be generated by primers covalently attached to a slide, which bind the DNA fragments from a sample and undergo the so-called bridge amplification. This generates clusters of identical sequences as templates for the sequencing reaction. Sequencing is performed by cyclic reversible termination. Symbols: Three colour bars represent sample DNA fragments with adaptors for universal priming sites; small single-colour bars attached to the surface of small beads (a) or a solid surface (b) represent complementary primers to adaptors.

fragments on each bead are then immobilized and sequenced. Amplified templates may also be generated on a solid support by covalently attached primers. Fragmented adapter-ligated DNA molecules are bound to these primers and amplified through a series of the so-called bridge amplifications to generate clusters of identical sequences (e.g. Illumina), which provide templates for the sequencing reaction (Fig. 1). Some technologies also read single DNA sequences, which are immobilized on solid supports and directly sequenced (25).

The sequencing reaction itself can be carried out by cyclic reversible termination, single-nucleotide addition and real-time sequencing, or by ligation (19). In cyclic reversible termination, as used in the Illumina and Helicos systems, chemically modified nucleotides are used as terminators of the sequencing reaction and are subsequently removed for addition of the next nucleotide in the next cycle. The HiSeq system from Illumina was launched in 2010 and provides up to 600 Gb of sequence per run (26). In the case of sequencing by ligation, as applied in the Life Technologies/Applied Biosystems SOLiD system, a fluorescently labelled nucleotide probe hybridizes to the DNA template adjacent to a primer (27). The Ion PM (Personal Genome Machine) from Ion Torrent (now Life technologies) uses semiconductor technology which measures the release of a proton after nucleotide incorporation as a sequencing signal (26, 28). In pyrosequencing reactions such as those used by Roche/454, a labelled nucleotide is detected when an inorganic pyrophosphate from the incorporated nucleotide emits a light signal after enzymatic transformation (29). Newer 'third-generation' technologies recently released include that from Pacific Biosciences (30). On this platform, DNA polymerase molecules are attached to the surface of specific detectors that analyse signals from dye-labelled nucleotides after nucleotide extension. Two principal features distinguish this approach: PCR amplification of the template is not required, and the signals are detected in real-time (26). Pacific Bioscience launched Pacbio RS as a commercial system in 2011. In nanopore sequencing (Oxford Nanopore Technologies), another third-generation sequencing method, sequencing is not based on extension of a template. Instead, DNA passes through a protein nanopore where different bases produce different electrical signals which are measured by an electrophysiological technique (31). So far no large-scale studies have been published with the latter two technologies.

These NGS technologies have facilitated whole-genome, whole-exome and whole-transcriptome analyses at an unprecedented scale. The comprehensive sequencing and analysis of cancer genomes (32,33) including large-scale projects such as The Cancer Genome Atlas (34), the International Cancer Genome Consortium (35) as well as the 1000 Genomes Project have been made possible with these technologies.

Next-generation sequencing bioinformatics and statistics

Computational processing and statistical analysis are critical components in making sense of the vast amounts of sequence data generated by these high-throughput platforms. A number of challenges are posed by the next-generation sequencing technologies, and these must be overcome through the production of both the appropriate type and the amount of sequence data, as well as through the use of appropriate algorithms and statistical tests. There is active, ongoing development of software for the analysis of high-throughput sequencing data for both cancer and non-cancer applications. Most researchers choose to customize widely applied tools to deal with the complexities of tumor samples, which are distinguished in many ways from normal samples, including by sample quality, sample volume, the complexity of their genome sequences and the inclusion of non-tumor material (36,37).

Processing pipelines differ depending on the main aims of the analysis. In Figure S1, we present one typical analysis which seeks to identify mutations and structural rearrangements and to quantify alterations in the expression levels of genes using transcriptome sequence data. Usually, a comparison of tumor sample/s with matched normal sample/s is used.

Beginning with raw data produced by the sequencing instrument, the samples proceed through (i) base calling to identify as accurately as possible the base at each position in each read and to assign an informative base-quality score which reflects well the confidence in the basecall (38), (ii) quality filtering to remove adaptor sequences and low-quality reads and (iii) alignment to a reference genome to identify the origin of the sequence read (Fig. 2). Following this basic processing, the mutations and genomic rearrangements can be identified by comparison to the normal sample.

A major challenge for all applications is the development of the appropriate software that can deal with the very high number of short sequence reads – millions to billions of reads per instrument run – generated by these platforms. Read lengths are critical for accurate mapping to a reference genome. Although read lengths have increased considerably from the 30–50 nt reads initially possible with these NGS platforms, reads are still substantially shorter (30–500 bp) than the kilobase sequences typically generated on the Sanger platform. Further, the error rates for these technologies are higher (10^{-3}–10^{-2}), which must be taken into account in analysis and which affects the identification of polymorphisms, particularly those present at low frequencies (39).

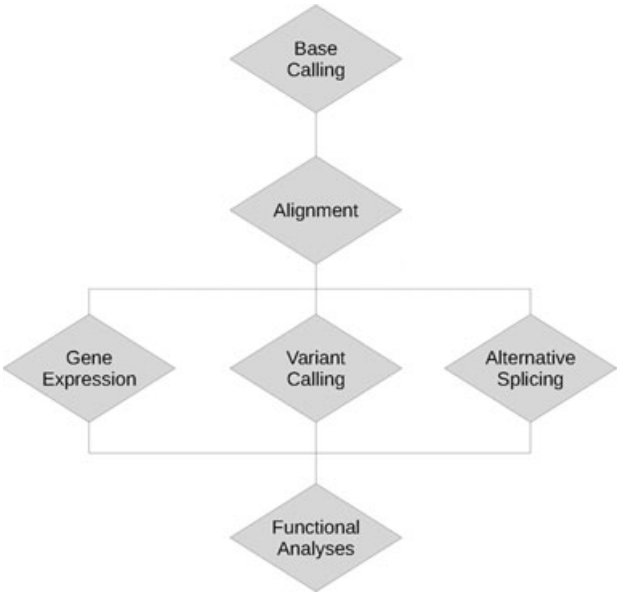

Figure 2. Schematic representation of next generation sequencing data analysis. After base calling and alignment of sequences, next-generation sequencing allows the analysis of differences in gene expression, identification of mutations and structural alterations (variant calling) and alternative splicing of genes.

Alignment

Most analysis pipelines begin by aligning the sequence reads to the reference human genome. Accurate and efficient mapping of millions/billions of short reads to the human reference is crucial for the correct identification of the genomic origin for a sequence read as this forms the basis for all downstream analyses including variant calling (40).

The complex combination of mutation and structural rearrangement present in cancer genomes means that alignment strategies must ensure that reads which are divergent from the reference genome are aligned with maximal sensitivity and specificity (41). Uniqueness of a read mapping is a useful measure of map reliability. Most short-read mappers provide a map-quality score which reflects the probability that a given read is misplaced in the reference genome. This map-quality score allows reads which align to multiple alternative genomic loci to be assigned a lower map-quality score which reflects the uncertainty about its placement and which can be used to filter reads in order to ensure high-quality single nucleotide variant calls. Read length and error rate are important factors in obtaining accurate mapping, that is, longer, error-free reads can be aligned unambiguously more easily than short reads with a high number of mismatches to the reference. Longer reads also carry more information about structural variation than short reads. A number of aligners have been developed that are able, to different extents, to deal with the short read lengths and high error rates typical of high-throughput sequencing platforms (reviewed in 41). A few of those commonly used for the mapping of genomic reads are MAQ (42), BWA (43), Bowtie (44), SOAP3 (45) and SHRiMP (46). Other mappers such as SpliceMap (47), TopHat (48) and GSNAP (49) allow for the spliced alignment of reads from whole-transcriptome sequencing (RNA-Seq). Appropriate mapping software should be chosen with the data type, read length and error rate in mind.

Identifying mutations

Both genome and transcriptome sequencing provide data suitable for the identification of mutations within tumor samples. Confident variant identification requires multiple obseravtions for each nucleotide, with current recommendations suggesting a coverage of at least 20-fold (50). Approaches to variant calling range from simple allele counting (51) to more advanced probabilistic approaches such as those implemented in GATK's Unified Genotyper (52,53) and SAMtools bcf (54). The advantage of using approaches that implement likelihood ratio tests or Bayesian methods is that these can be used with lower coverage data and provide a probabilistic measure of the uncertainty associated with each single nucleotide variant call that takes multiple samples as well as sources of error into account.

Identifying structural variation

Paired-end sequence reads from molecules of defined size are particularly useful for the identification of structural variations (SVs) that are often seen in tumor samples. In the simplest approaches, reads from either end of a single molecule are mapped to the reference genome, and reads that map to different chromosomes or in incorrect orientations are posited to be due to an interchromosomal rearrangement (55). A number of approaches such as PEMer (56), BreakDancer (57), PinDel (57) and GASV (58) allow intrachromosomal rearrangement to be identified. These approaches infer only the approximate genomic locations of a SV but do not allow the more precise identification of breakpoints as is possible with CREST (59). As short reads are less likely to be mapped uniquely in genomic regions with a high repeat content or recent duplications, SVs involving low-complexity sequence may be under-represented in precisely the parts of the genome most likely to carry structural variation (60).

Gene fusions, the creation of a hybrid gene due to a genomic rearrangement, are prevalent in some cancers. Approaches to the identification of gene fusion events in RNASeq data that make use of the discordant alignment of paired-end reads have been successfully applied to the identification of gene fusions in a chronic myelogenous leukemia cell line and in prostate cancer samples (61).

Expression profiling

High-throughput sequencing approaches to transcript quantification provide a number of advantages over array-based methods, including the ability to assay all transcripts without prior knowledge of transcript sequences, reduced background due to elimination of cross-hybridization and probe-specific effects and an increased dynamic range of detection allowing both low and high expression to be more accurately quantified (62,63). A typical approach to expression quantification begins with spliced alignment of sequence reads to the reference genome, quantification of transcript abundance and finally detection of differential expression between tumor and matched normal samples. A widely used software suite including the Bowtie, TopHat and Cufflinks packages (64) provides one approach to this workflow. Bowtie and TopHat carry out the alignment of reads and splice-site discovery, while Cufflinks assembles and quantifies transcripts and Cuffdiff identifies differential expression between samples. For the comparison of expression levels across genes within and between experiments, the abundance of each gene per sample is estimated (63).

Insights into the melanoma genome

SPDEF/PTEN

In the first whole-cancer-genome sequencing study, the genomes of a melanoma cell line (COLO-829) and a lymphoblastoid cell line (COLO-829BL) from the same patient were compared (52). Of 292 somatic base substitutions identified in protein-coding sequences, 187 were non-synonymous (changing the protein code). Interestingly, the ratio of non-synonymous to synonymous substitutions did not differ from that expected by chance and thus provided no evidence for selection at these sites.

Regarding candidates for new cancer genes for melanoma pathogenesis, authors described two heterozygous missense mutations in *SPDEF*, a member of the ETS transcription factor family. In line with this, enhanced SPDEF expression has been described earlier in tumor progression of prostate, breast and ovarian cancer (65).Moreover, a missense mutation was found in MMP28, a member of the matrix metalloproteinase gene family.

An eight to 12-fold copy number increase was observed on chromosome 3p which harbours four genes: *RARB*, *TOP2B*, *NGLY1* and *KS (OXSM)*. Copy number changes of these genes have not been described in other cancers, but *RARA*, another member of the retinoic receptor family, is known to be rearranged in acute promyelocytic leukaemia (66).The majority of somatic mutations in COLO-829 were C>T/G>A transitions, 360 of 510 dinucleotide substitutions were CC>TT/GG>AA changes. This mutational spectrum is highly suggestive of an induction by UV light (67). This study showed that a cancer genome may harbour a large number of somatic mutations. However, generation of more genome-wide catalogues of melanoma mutations is necessary to identify common genetic variants. Table 1 gives an overview of currently available major melanoma-sequencing studies.

RB1-ITM2B, PREX2

RNA-seq was performed for eight patient-derived melanoma short-term cultures and two known melanoma cell lines (MeWo and 501 Mel) (68). Overall, 11 novel heterozygous gene fusions were identified. The tumor suppressor gene *RB1* was part of one of these gene fusions. Although its precise role in melanoma has not been clarified up to now, RB1 is located in a region with frequent copy number losses (2). Other fusions identified in this study such as *RECK-ALX3* also involved cancer-related genes. RECK is a known inhibitor of tumor invasion and metastasis (69). Interestingly, none of the fusion transcripts were present in more than one cell line, nor in additional 90 short-term cultures from metastatic lesions and must thus be regarded as private (sample-specific) mutations. Overall, 12% of all transcribed genes were sequenced in this study. On average, 130 genetic variants were identified per sample, of which 30% might be true somatic mutations, because 70% were also found in matched germline DNA and may thus be regarded as single nucleotide polymorphisms (SNPs). Several of these new mutations occur in genes previously shown to be mutated in cancer, namely *A2M, CAST, CENTD3, FUS, NUP133, SF3B1, TNFRSF14* and *TRIB3*. The majority (86%) of mutations were CG > TA transitions, again indicative of UV induction. Of note, mutations in *BRAF* and *NRAS*, the two most commonly mutated melanoma oncogenes, were not identified, likely due to the low abundance of the transcripts.

In a subsequent study, the same group performed whole-genome sequencing of 25 metastatic melanomas and matched germline DNA (70). Overall, 9653 missense, nonsense or splice-site mutations were detected in 5712 genes and on average 97 structural rearrangements per melanoma genome. Eleven genes were found to be significantly mutated across the 25 samples, among which were *BRAF* and *NRAS*, mutated in 16 and nine samples, respectively. The top significantly mutated genes with recurrent mutations included *PREX2, MUC4, PRG4* and *MST1*. In an extension cohort of 107 samples, a 14% mutation frequency was found for *PREX2*. *PREX2* (phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 2) interacts with PTEN tumor suppressor. However, its precise function in melanoma is not yet understood.

MMP8

Automated Sanger sequencing was used in a focused mutational analysis of the matrix metalloproteinase (MMP) protein family (71). The human MMP family consists of 23 genes and was analysed by sequencing of PCR-amplified MMP exons in 32 individuals with metastatic melanoma lesions. An additional set of 47 metastatic melanomas was analysed for genes which showed one or more non-synonymous mutations. Overall, 28 non-synonymous somatic mutations were found in eight genes present in 23% of all

Table 1. Genetic aberrations found in melanoma by recent studies using automated Sanger sequencing (PCR sequencing) and next-generation sequencing (NGS)

No.	Tissue	Material	Technology	Genes with genetic aberrations	Ref.
1.	Metastatic melanoma lesions	DNA	PCR sequencing	*MMP8, MMP24, MMP27, MMP28*	71
2.	Metastatic melanoma lesions	DNA	PCR sequencing	*ERBB4, FLT1, EPHA10, PDGFRA, PTK2B*	73
3.	Melanoma cell line COLO-829	DNA	Whole-genome sequencing (NGS)	*BRAF, SPDEF, MMP28, RARB, TOPB2, PTEN*	51
4.	Melanoma short-term cultures, cell lines	RNA	RNA-seq (NGS)	*RB1-ITM2B, RECK-ALX3, A2M, CAST, CENTD3, FUS*	68
5.	Metastatic melanoma lesions, cell lines	DNA	Whole-exome sequencing (NGS)	*BRAF, GRIN2A, TRRAP, DCC, ZNF831, TMEM132B*	75
6.	Metastatic melanoma lesions	DNA	PCR sequencing	*GRM3, GPR98, CHRM3, GRM8, LPHN2*	79
7.	Melanoma cell lines	DNA	Whole-exome sequencing (NGS)	*BRAF, NRAS, MAP2K1/2, PTEN, FAT4, DSC1, LRP1B*	81
8.	Cutaneous metastases	DNA, RNA	Whole-genome and whole-exome sequencing, RNA-seq (NGS)	*HRAS, ELK1, CDKN2C*	83
9.	Acral melanoma and lymph node metastasis	DNA	Whole-genome and whole-exome sequencing (NGS)	*SCAF1, WNT1, ASB9, FAT2, PTRF, RHOB, CNDP2, DROSHA, ERCC5, LRRK1*, and *LRRFIP1*	85
10.	Metastatic melanoma lesions	DNA	Whole-exome sequencing (NGS)	*BRAF* amplification	91
11.	Metastatic melanoma lesions	DNA	Whole-exome sequencing (NGS)	*BRAF, NRAS, PREX2, MUC4, PRG4, MST1*	70
12.	Metastatic melanoma lesions, primary melanomas, melanoma cell lines	DNA	Whole-exome sequencing (NGS)	*BRAF, NRAS, PPP6C, RAC1, SNX31, TACC1, STK19*	92
13.	Metastatic melanoma lesions, primary melanomas,	DNA	Whole-exome sequencing (NGS)	*BRAF, NRAS, p PPP6C, DCC, PTPRK, GRM3*	93

melanoma samples, with a high non-synonymous to synonymous ratio (28:5). MMP members with most frequent mutations were *MMP8* and *MMP27*. Authors focussed on MMP8, because of its protective role in skin tumor development (72).

In a series of biologic experiments, tumor-derived MMP8 mutants showed less proteolytic activity for collagen I than wild-type MMP8. Moreover, cells expressing wild-type *MMP8* had a lower migration capacity than cells expressing mutant *MMP8*. In a mouse metastasis model, numerous lung metastases were observed in mice injected with *MMP8* mutant cell clones, but none in mice injected with wild-type clones. These data suggest that wild-type *MMP8* may inhibit tumor progression in melanoma while the mutant variant promotes it.

ERBB4

Prickett et al. (73) performed a focused mutational analysis by automated Sanger sequencing of protein tyrosine kinases, a gene family frequently mutated in cancers. The coding exons of the kinase domains of all 86 members of this gene superfamily were sequenced in 29 metastatic melanomas (72). Nineteen genes were identified containing a total of 30 somatic mutations. In accordance with UV patterns described previously, the number of C>T mutations was significantly higher than the number of other nucleotide substitutions.

ERBB4 was the most highly mutated. It harboured 24 somatic mutations in 19% of all samples. Mutated *ERBB4* variants transfected into HEK 293T cells had higher kinase activity than the wild-type variant. Moreover, *ERBB4* mutants had a stronger transforming capacity for NIH 3T3 cells than wild-type *ERBB4* and were similar in their potency to oncogenic K-RAS (G12V).

Knockdown of ERBB4 with short hairpin (shRNA) showed that targeting of ERBB4 in wild-type cells had only slight effects on proliferation, but targeting of ERBB4 in cells expressing mutant *ERBB4* significantly reduced their growth. These findings suggest that mutant *ERBB4* seems to be indispensible for the growth of melanomas harbouring these mutations. Treatment of melanoma cells with lapatinib (GW2016), a pan-ERBB pharmacological inhibitor (74), significantly reduced cell proliferation of ERBB4 mutant cells but not of wild-type cells. Thus, ERBB4 appears to be an interesting therapeutic target for specific small-molecule inhibitors, at least in cases with *ERBB4* mutations.

GRIN2A

In a more recent study, whole-exome sequencing was performed in 14 metastatic melanoma samples (72, 75). Overall, 4226 putative somatic alterations after quality filtering were identified. Of these alterations, 2813 were non-synonymous. Again, a majority (59 of 116) of dinucleotide substitutions were CC>TT/GG>AA changes, suggestive for UV-induced alterations.

Besides the well-known *BRAF* V600G alteration found in seven of 14 samples, nine additional genes had recurrent mutations. Among these was TRRAP, which encodes a transformation/transcription domain-associated protein. *GRIN2A* was mutated in six of the 14 melanomas in the discovery screen and showed an additional 11 somatic mutations in a prevalence screen. In total, 34 distinct *GRIN2A* mutations were identified in 135 samples (25.2%). Thus, *GRIN2A* was the highest mutated gene in this screen. *GRIN2A* encodes for glutamate (*N*-methyl-(D)-aspartic acid (NMDA)) receptor subunit of an ionotropic glutamate receptor.

Gene amplifications of *GRIN2B* and *GRIN2C* have been shown in a study on seven melanoma cell lines using single nucleotide polymorphism analysis, comparative genomic hybridization and RNA-seq (76). Moreover, the metabotropic glutamate receptor GRM1 has been described as an oncogene for melanoma cells (77). Thus, glutamate receptor signalling appears to be a pathogenic pathway at least in a subset of melanomas. These findings might be of therapeutic relevance, because glutamate receptor antagonists have been shown in a recent report to make melanoma cells sensitive to ionizing radiation (78).

GRM3

Targeted exome sequencing was used to analyse the mutational status of 734 G protein–coupled receptors in 11 metastatic melanoma samples (79). Overall, 755 non-synonymous mutations were identified. Eleven genes of a total of 106 mutated genes harboured at least two somatic mutations. These were further analysed in 80 additional melanoma samples. *GRM3* and *GPR98* were shown to be most frequently mutated.

GRM3, a group 2 glutamate receptor, showed non-synonymous mutations in 13 of 80 tumors (16.3%), and *GPR98* showed non-synonymous mutations in 22 of 80 tumors (27.5%). In an additional set of 57 melanomas, a mutational hotspot for GRM3 was verified (p.Glu870Lys). Stably *GRM3* wild-type transfected melanoma cell clones were growing at significant lower rate than mutant clones. Activation of GRM3 by DCG-IV (a carboxylic derivative of non-selective glutamate receptor agonists) led to enhanced activity of MEK1/2 in mutant melanoma cell clones compared with wild-type cells. Moreover, introduction of mutant *GRM3* enhanced migration of melanoma cell-line A375 and transformed melanocytic cell-line Mel-STR. NOD/SCID mice injected with melanoma cells expressing vector alone developed a significant lower number of gross lung metastases than mice injected with cells expressing mutant *GRM3* variants, although the total number of metastases was not different. Moreover, GRM3 knockdown reduced tumor growth in subcutaneously injected mice with melanoma cell clones harbouring mutant *GRM3*.

Finally, mutant cells were shown to be more sensitive than wild-type cells to growth inhibition by AZD-6244 (selumitinib), a small molecule inhibitor against MEK1/2. Thus, targeting of MEK1/2 signalling in the presence of *GRM3* mutations might be a promising strategy for the treatment of metastatic disease (80).

MAP2K1/2

Whole-exome sequencing of a set of seven melanoma cell lines from metastatic lesions identified 3611 somatic variants in protein-coding regions of 2586 genes (81). Again, the majority of mutations were C>T/G>A transitions, indicative of UV damage–induced lesions. The exomes of two cell-line patients were derived from a regional lymph node and a distant metastatic site, respectively, of the same patient. The first one was excised 3 years and the second one 15 years after diagnosis. More than 70% of the mutations overlapped between both samples, suggestive of a common tumor origin of both metastases and the occurrence of mutations early in tumor progression.

All samples analysed harboured mutations in the RAS–RAF–MAPK pathway. As determined by single cell analysis, *BRAF* mutations were present in the same cells as *MAP2K1* or *MAP2K2* mutations. MAP2K1-, MAP2K2-, BRAF- and NRAS-coding sequences were analysed in 127 additional melanoma samples.

Overall, 10 of 127 melanomas (8%) had mutations in either *MAP2K1* or *MAP2K2*.

MAP2K1 mutants transfected into HEK293T cells showed significant ERK1/3 phosphorylation, while no phosphorylation was observed with the wild-type kinase, indicating that the mutants are constitutively active. Moreover, cells with *MAP2K2* E207K mutation were more resistant to MAP2K1/2 (MEK1/2) inhibitor AZD6244 than cells carrying only the *BRAF* V600G mutation. Together, *MAP2K1* and *MAP2K2* mutations might have treatment implications for melanoma regarding combined treatment with *BRAF* V600E inhibitors (82).

HRAS

A set of different methodologies (termed integrative sequencing) was used in a recent patient study including one patient with metastatic melanoma (83). This study was performed to explore the applicability of high-throughput sequencing in clinical oncology for rational treatment decisions. Whole-genome sequencing of the tumor, targeted whole-exome sequencing of tumor and normal DNA and transcriptome sequencing (RNA-Seq) of the tumor were performed. A multidisciplinary Sequencing Tumor Board combined clinical and sequencing information to decide about treatment options.

Sequencing of punch biopsies from four melanoma skin metastases identified an activating mutation of *HRAS* Q61L, a point mutation in the ETS transcription factor family member *ELK1* R74C and an inactivating rearrangement of cyclin-dependent kinase inhibitor 2c (*CDKN2C* or p18INK4C), which were chosen to decide about treatment options. Interestingly, mutations were not observed in *BRAF*, *c-KIT* or *NRAS*. HRAS was regarded as the most promising target, because RAS signalling leads to RAF-MAPK and PI3K-mTOR pathway activation, and there are a number of ongoing clinical trials for MEK (MAP2K1/2), PI3K and mTOR (84). This patient was regarded as a good candidate for an upcoming clinical trial with combined treatment of PI3K and MEK inhibitors (www.clinicaltrials.gov; NCT01363232).

WNT1

Whole-genome sequencing was used to characterize somatic mutations in a primary acral melanoma and a concurrent lymph node metastasis after microdissection of tumor cells (85). A non-synonymous mutation in the *SCAF1* gene was found in the primary tumor but not the metastasis, and a non-synonymous mutation in *WNT1* and a point mutation in a splice site of SUPT5H were present in the metastasis but not the primary tumor. While WNT1 signalling is known to play a role in melanoma pathogenesis (86), SUPT5H, a regulator of transcriptional elongation, has not yet been described in this tumor.

Some of the genes found to be mutated in both tumor samples were regarded as potentially relevant for melanoma, such as *ASB9*, *FAT2, PTRF, RHOB, CNDP2, DROSHA, ERCC5, LRRK1* and *LRRFIP1*, because they had already been linked to other cancers. Some, such as *KSR1* (kinase suppressor of Ras 1), *PREX2* (exchange factor in PI3K signalling) and *MFI2* (melanotransferrin), had been linked to melanoma in earlier reports (87–90). Again, 60% of mutations were C>T transitions and thus reminiscent of UV damage–induced changes.

BRAF

The mechanisms of resistance of metastatic melanoma to BRAF inhibitor vemurafenib were addressed in a recent study performing exome sequencing on two matched pairs of baseline and disease progression samples (91). Interestingly, these pairs showed *BRAF* V600E copy number gains. Subsequent focussed analysis in a set of 20 samples identified further two samples with *BRAF* V600E copy number gains.

Generation of seven vemurafenib-resistant melanoma cell lines *in vitro* provided one cell line with an increased copy number of the *BRAF* V600E gene. This cell line showed elevated levels of phospho-ERK and was highly resistant to BRAF(V600E) inhibition. Similarly, overexpression of BRAF(V600E) using viral gene constructs conferred vemurafenib resistance to another cell line at concentrations up to 1 μM, but resistance could partly be overcome at a concentration of 10 μM. Addition of MEK inhibitor AZD6244 (selumetinib) could restore sensitivity. A combination of BRAF and MEK inhibitor treatment worked synergistically. Together, these findings might provide a rationale for a dose escalation of vemurafenib in patients with acquired *BRAF* amplifications or for combination therapies with MEK1/2 inhibitors.

RAC1, PPPC6C, STK19

In a very recent study, whole-exome sequencing for 121 melanoma samples was performed (15 primary tumors, 30 metastatic samples and 76 short-term cultures from metastatic lesions) (92). Authors used a statistical approach that supported the identification of positively selected gene mutations based on the distribution of exon/intron mutations, together with an analysis of the functional impact of these mutations. Among eleven genes found by this approach were the known melanoma genes *BRAF*, *NRAS*, *PTEN*, *TP53*, *p16INK4a* and *MAP2K1* and five new candidates (*PPP6C*, *RAC1*, *SNX31*, *TACC1* and *STK19*). Three of the new candidates (*RAC1*, *PPP6C* and *STK19*) harboured recurrent mutations. *PPP6C* encodes for a subunit of a protein phosphatase complex and has been described as a tumor suppressor. STK19 is a predicted kinase with an as yet unknown function. *RAC1*, a RAS-related member of the Rho subfamily of GTPases, showed the most frequent recurrent mutation after *BRAF* and *NRAS* (*RAC1* P29S 5% of all cases). RAC1 is known to be involved in cytoskeleton rearrangement. The *RAC1* P29S mutation was also identified in another recent study on 147 melanoma samples, again as the most frequent recurrent mutation after BRAF and *NRAS* (9.2% of sun-exposed melanomas and 4.7% of all tumors) (93). Functional experiments showed that this mutation enhanced binding activity of RAC1 to downstream effectors such as PAK1 and MLK3 and melanocyte proliferation and migration.

Intratumor heterogeneity

Intratumor heterogeneity was analysed in spatially separated samples obtained from four primary renal carcinomas and associated metastases (94). For two of the four patients, whole-exome sequencing was performed.

In the first patient, a total of 128 mutations were found in nine different regions of the primary tumor and in three metastatic lesions. These were subdivided into 40 ubiquitous mutations (shared by all specimens), 59 mutations shared by several but not all regions of the primary tumor and 29 mutations that were unique to specific regions (private mutations). Authors constructed a phylogenetic tree of the tumor regions and metastases (95). Interestingly, this analysis did not reveal a linear tumor evolution but rather showed branching with one branch of the tree that led into clones present in metastatic sites and the other that led into primary tumor regions.

Multiregion exome sequencing was also performed for the primary tumor and a metastasis from patient 2. Samples from patient 2 showed a total of 119 mutations, 36 of which were shared between all samples (ubiquitous mutations). Only four of the total of 30 samples from the four patients of this study had identical allelic imbalance profiles.

Taken together, this study demonstrated significant genetic intratumor heterogeneity. The mutation pattern given by a single tumor biopsy may thus be misleading regarding treatment decisions for targeted therapy. Genetic hetereogeneity may also support tumor recurrences.

Conclusions

There are a number of pros and cons for each sequencing platform, and with the emergence of technologies such as PacBio RS and Oxford Nanopore, which have not yet been used in larger studies, it is likely that the selection of the ideal platform will continue to change for some time. At present, cost and throughput considerations favour the use of the Illumina HiSeq and Life/SOLiD technologies for larger-scale projects. Automated Sanger sequencing is highly accurate (99.999%) and may have its place in small-scale analyses in the range of kilobases to megabases. A critical issue is the amount of DNA/RNA material needed, which is normally 2–5 μg for most NGS technologies. This may drop to 100 ng, when amplification protocols are used (96). However, with the disadvantage that PCR amplification has a small but significant error rate.

There is great variability in data output per run in NGS with a maximum of 600 Gb per run for HiSeq, although other systems such as SOLiD5500xl are improving this aspect. A major issue are costs per base, currently 20-fold higher for the 454 GS FLX than for HiSeq and SOLiD5500xl. However, the 454 system provides significantly longer reads (1000 bp) compared with HiSeq and SOLiD5500 systems (150–250 bp). Short reads can make the final assembly or the mapping process more difficult. Sequencing accuracy varies between 98% (HiSeq), 99.9% (454 GS FLX), and 99.999% for Sanger sequencing.

Recently developed desktop-sequencing systems MiSeq (Illumina), Ion Torrent PGM (Life Technologies) and 454 GS Junior (Roche) are small in size, have high throughput but limited amounts of data per run (28). These will be used for small genomes (such as bacterial genomes), clone or amplicon checking, but may also be of interest for cancer mutation analysis, for example, for sequencing of mutational hotspots (97,98).

In contrast to exon and whole-genome or exome sequencing, RNA-Seq provides information about gene expression levels and splicing variants. But genetic aberrations such as mutations in low-abundance genes may be missed (64). On the other hand, exon and whole-genome sequencing do not provide the information whether a mutated gene is indeed expressed. The choice of an appropriate system must be based on the individual scientific or clinical question.

The described sequencing approaches have made possible many advances regarding the genetic basis of melanoma, some of which will surely translate into new treatment approaches or new designs of clinical trials. For example, ERBB4, GRIN2A and MAP2K1/2 identified in these studies are promising targets for new treatment approaches. However, some important issues have to be addressed. As many studies relied on a limited number of cell lines, short-term cultures or metastatic tumors from different locations, larger studies and studies with more homogeneous tumor material are needed. Moreover, the increased throughput made possible by NGS approaches revealed an unexpected heterogeneity, often with little overlap of genetic variations between samples. This is complicated by the fact that a single tumor may even have different mutational landscapes in different areas (94).

A major challenge will also be the identification of genetic differences between treatment responsiveness and treatment resistance of tumors after recurrence.

Acknowledgements

M.D. and J.K. are supported by funding from the Max Planck Society; M.K. is supported by funding from the Deutsche Krebshilfe, grant no. 109716. All authors contributed equally in drafting and writing of the manuscript. M.D. and J.K. analysed the melanoma sequence mentioned in the Figure S1. All authors declare that there are no conflicts of interest. Sequencing of the primary melanoma (Figure S1) was performed after informed, written consent of the patient and according to the Declaration of Helsinki. It has been approved by the local ethics committee of the University of Leipzig (July 26, 2011; number 224-11-11072011).

Conflict of interests

The authors have declared no conflicting interests.

References

1 Miller A J, Mihm M C Jr. N Engl J Med 2006: **355**: 51–65.
2 Curtin J A, Fridlyand J, Kageshita T *et al.* N Engl J Med 2005: **353**: 2135–2147.
3 Davies M A, Samuels Y. Oncogene 2010: **29**: 5545–5555.
4 Chapman P B, Hauschild A, Robert C *et al.* N Engl J Med 2011: **364**: 2507–2516.
5 Sosman J A, Kim K B, Schuchter L *et al.* N Engl J Med 2012: **366**: 707–714.
6 Nazarian R, Shi H, Wang Q *et al.* Nature 2010: **468**: 973–977.
7 Poulikakos P I, Persaud Y, Janakiraman M *et al.* Nature 2011: **480**: 387–390.
8 Kudchadkar R, Paraiso KH, Smalley KS. Cancer J 2012: **18**: 124–131.
9 Straussman R, Morikawa T, Shee K *et al.* Nature 2012: **487**: 500–504.
10 Johannessen C M, Boehm JS, Kim SY *et al.* Nature 2010: **468**: 968–972.
11 Yadav V, Zhang X, Liu J *et al.* J Biol Chem 2012: **287**: 28087–28098.
12 Waller C F. Recent Results Cancer Res 2010: **184**: 3–20.
13 Carvajal R D, Antonescu C R, Wolchok J D *et al.* JAMA 2011: **305**: 2327–2334.
14 Chudnovsky Y, Khavari P A, Adams AE. J Clin Invest 2005: **115**: 813–824.
15 Dankort D, Curley D P, Cartlidge R A *et al.* Nat Genet 2009: **41**: 544–552.
16 Hafner C, Landthaler M, Vogt T. Exp Dermatol 2010: **19**: e222–e227.
17 Passeron T, Lacour J P, Allegra M *et al.* Exp Dermatol 2011: **20**: 1030–1032.
18 Mardis E R. Annu Rev Genomics Hum Genet 2008: **9**: 387–402.
19 Metzker M L. Nat Rev Genet 2010: **11**: 31–46.
20 Haimovich A D. Yale J Biol Med 2011: **84**: 439–446.
21 Lee H, Tang H. Methods Mol Biol 2012: **855**: 155–174.
22 Kircher M, Kelso J. BioEssays 2010: **32**: 524–536.
23 Dressman D, Yan H, Traverso G *et al.* Proc Natl Acad Sci USA 2003: **100**: 8817–8822.
24 Fedurco M, Romieu A, Williams S *et al.* Nucleic Acids Res 2006: **34**: e22.
25 Harris T D, Buzby P R, Babcock H *et al.* Science 2008: **320**: 106–109.
26 Liu L, Li Y, Li S *et al.* J Biomed Biotechnol 2012: **2012**: 251364.
27 Valouev A, Ichikawa J, Tonthat T *et al.* Genome Res 2008: **18**: 1051–1063.
28 Loman N J, Misra RV, Dallman TJ *et al.* Nat Biotechnol 2012: **30**: 434–439.
29 Elahi E, Ronaghi M. Methods Mol Biol 2004: **255**: 211–219.
30 Eid J, Fehr A, Gray J *et al.* Science 2009: **323**: 133–138.
31 Eisenstein M. Nat Biotechnol 2012: **30**: 295–296.
32 Ding L, Wendl M C, Koboldt D C *et al.* Hum Mol Genet 2010: **19**: R188–R196.
33 Brosnan J A, Iacobuzio-Donahue C A. Semin Cell Dev Biol 2012: **23**: 237–242.
34 Collins F S, Barker A D. Sci Am 2007: **296**: 50–57.

35 International Cancer Genome Consortium, Hudson T J, Anderson W *et al.* Nature 2010: **464**: 993–998.
36 Meyerson M, Gabriel S, Getz G. Nature Rev Genet 2010: **11**: 685–696.
37 Robison K. Brief Bioinform 2010: **11**: 524–534.
38 Kircher M, Stenzel U, Kelso J. Genome Biol 2009: **10**: R83.
39 1000 Genomes Project Consortium. Nature 2010: **467**: 1061–1073.
40 Trapnell C, Salzberg S L. Nat Biotechnol 2009: **27**: 455–457.
41 Mardis E R, Wilson R K. Hum Mol Genet 2009: **18**: R163–R168.
42 Li H, Ruan J, Durbin R. Genome Res 2008: **18**: 1851–1858.
43 Li H, Durbin R. Bioinformatics 2009: **25**: 1754–1760.
44 Langmead B, Trapnell C, Pop M *et al.* Genome Biol 2009: **10**: R25.
45 Liu C M, Wong T, Wu E *et al.* Bioinformatics 2012: **28**: 878–879.
46 Rumble S M, Lacroute P, Dalca A V *et al.* PLoS Comput Biol 2009: **5**: e1000386.
47 Au K F, Jiang H, Lin L *et al.* Nucleic Acids Res 2010: **38**: 4570–4578.
48 Trapnell C, Pachter L, Salzberg S L. Bioinformatics 2009: **25**: 1105–1111.
49 Wu T D, Nacu S. Bioinformatics 2010: **26**: 873–881.
50 Nielsen R, Paul JS, Albrechtsen A *et al.* Nat Rev Genet 2011: **12**: 443–451.
51 Pleasance E D, Cheetham R K, Stephens PJ *et al.* Nature 2010: **463**: 191–196.
52 McKenna A, Hanna M, Banks E *et al.* Genome Res 2010: **20**: 1297–1303.
53 DePristo M A, Banks E, Poplin R *et al.* Nat Genet 2011: **43**: 491–498.
54 Li H, Handsaker B, Wysoker A *et al.* Bioinformatics 2009: **25**: 2078–2079.
55 Korbel J O, Urban A E, Affourtit J P *et al.* Science 2007: **318**: 420–426.
56 Chen K, Wallis J W, McLellan M D *et al.* Nat Methods 2009: **6**: 677–681.
57 Ye K, Schulz M H, Long Q *et al.* Bioinformatics 2009: **25**: 2865–2871.
58 Sindi S S, Onal S, Peng L *et al.* Genome Biol 2012: **13**: R22.
59 Wang J, Mullighan C G, Easton J *et al.* Nat Methods 2011: **8**: 652–654.
60 Onishi-Seebacher M, Korbel J O. BioEssays 2011: **33**: 840–850.
61 Maher C A, Kumar-Sinha C, Cao X *et al.* Nature 2009: **458**: 97–101.
62 Mortazavi A, Williams B A, McCue K *et al.* Nat Methods 2008: **5**: 621–628.
63 Wang Z, Gerstein M, Snyder M. Nat Rev Genet 2009: **10**: 57–63.
64 Trapnell C, Roberts A, Goff L *et al.* Nat Protoc 2012: **7**: 562–578.
65 Sood A K, Saxena R, Groth J *et al.* Hum Pathol 2007: **38**: 1628–1638.
66 Scaglioni PP, Pandolfi P P. Curr Top Microbiol Immunol 2007: **313**: 85–100.
67 Pfeifer G P, You Y H, Besaratinia A. Mutat Res 2005: **571**: 19–31.
68 Berger M F, Levin J Z, Vijayendran K *et al.* Genome Res 2010: **20**: 413–427.
69 Noda M, Oh J, Takahashi R *et al.* Cancer Metastasis Rev 2003: **22**: 167–175.
70 Berger M F, Hodis E, Heffernan T P *et al.* Nature 2012: **485**: 502–506.
71 Palavalli L H, Prickett TD, Wunderlich JR *et al.* Nat Genet 2009: **41**: 518–520.
72 Balbín M, Fueyo A, Tester AM *et al.* Nat Genet 2003: **35**: 252–257.
73 Prickett T D, Agrawal N S, Wei X *et al.* Nat Genet 2009: **41**: 1127–1132.
74 Qiu C, Tarrant M K, Choi S H *et al.* Structure 2008: **16**: 460–467.
75 Wei X, Walia V, Lin J C *et al.* Nat Genet 2011: **43**: 442–446.
76 Valsesia A, Rimoldi D, Martinet D *et al.* PLoS ONE 2011: **6**: e18369.
77 Pollock P M, Cohen-Solal K, Sood R *et al.* Nat Genet 2003: **34**: 108–112.
78 Khan A J, Wall B, Ahlawat S *et al.* Clin Cancer Res 2011: **17**: 1807–1814.
79 Prickett T D, Wei X, Cardenas-Navia I *et al.* Nat Genet 2011: **43**: 1119–1126.
80 Patel S P, Kim K B. Expert Opin Investig Drugs 2012: **21**: 531–539.
81 Nikolaev S I, Rimoldi D, Iseli C *et al.* Nat Genet 2011: **44**: 133–139.
82 Solit D, Sawyers C L. Nature 2010: **468**: 902–903.
83 Roychowdhury S, Iyer MK, Robinson DR *et al.* Sci Transl Med 2011: **3**: 111ra121.
84 Flaherty K T, Hodi FS, Fisher D E. Nat Rev Cancer 2012: **12**: 349–361.
85 Turajlic S, Furney S J, Lambros M B *et al.* Genome Res 2012: **22**: 196–207.
86 O'Connell MP, Weeraratna A T. Pigment Cell Melanoma Res 2009: **22**: 724–739.
87 Takata M, Goto Y, Ichii N *et al.* J Invest Dermatol 2005: **125**: 318–322.
88 Bertrand Y, Demeule M, Michaud-Levesque J *et al.* Biochem Biophys Res Commun 2007: **353**: 418–423.
89 Muchemwa F C, Ma D, Inoue Y *et al.* Br J Dermatol 2008: **158**: 411–413.
90 McKay M M, Ritt D A, Morrison D K. Curr Biol 2011: **21**: 563–568.
91 Shi H, Moriceau G, Kong X *et al.* Nat Commun 2012: **3**: 724.
92 Hodis E, Watson I R, Kryukov GV *et al.* Cell 2012: **150**: 251–263.
93 Krauthammer M, Kong Y, Ha BH *et al.* Nat Genet 2012: **44**: 1006–1014.
94 Gerlinger M, Rowan A J, Horswell S *et al.* N Engl J Med 2012: **366**: 883–892.
95 Merlo L M, Pepper J W, Reid B J *et al.* Nat Rev Cancer 2006: **6**: 924–935.
96 Linnarsson S. Exp Cell Res 2010: **316**: 1339–1343.
97 Harismendy O, Schwab RB, Bao L *et al.* Genome Biol 2011: **12**: R124.
98 Chan M, Ji S M, Yeo Z X *et al.* J Mol Diagn 2012: **14**: 602–612.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. RNA-seq of a primary melanoma.