Melan-OMICS: whole-exome and transcriptome sequencing to dissect molecular complexity of cutaneous malignant melanoma

Scan to download PDF poster

Cifola Ingrid[1], Pietrelli Alessandro[1], Severgnini Marco[1], Magi Alberto[2], Tattini Lorenzo[2], Mangano Eleonora[1], Santosh Anand[1,3], Russo Vincenzo[4], Battaglia Cristina[1,5], De Bellis Gianluca[1].

[1] Institute for Biomedical Technologies, National Research Council (ITB-CNR), Milan, Italy
[2] Dept. of Clinical and Experimental Medicine, University of Florence, Florence, Italy
[3] Dept. of Science and Technology, University of Sannio, Benevento, Italy.
[4] Cancer Gene Therapy Unit, San Raffaele Scientific Institute, Milan, Italy
[5] Dip.Biotecnologie Mediche e Medicina Traslazionale, Università di Milano, Italy

SUMMARY

Cutaneous melanoma is the most fatal skin cancer and, although some effective molecular therapies exist, novel targets and drugs are still needed. To provide new insights for novel targets discovery, we performed an extensive characterization by NGS of a collection of melanoma cell lines derived from metastatic cases. Samples were profiled by whole-exome sequencing (WES) and RNA-sequencing using Illumina technology.

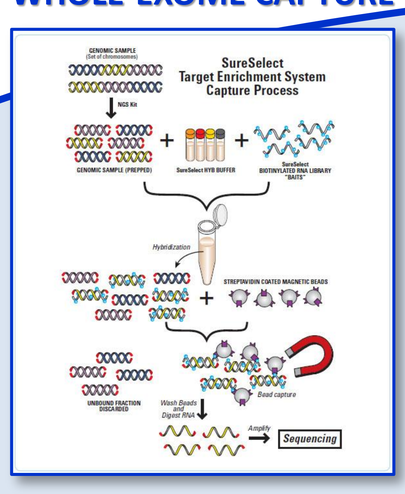
Starting from **WES data**, we developed a **bioinformatics pipeline** to catalogue **mutations** affecting melanoma key **biological pathways** already targeted by current therapies, as well as genes never described for melanoma [*Cifola, Plos One 2013*]. Moreover, WES data were used to perform **copy number alteration (CNA) analysis** using a novel software developed by us, called **Excavator**, which is very sensitive and precise in DNA copies estimation even in situations of great sample heterogeneity [*Magi, Genome Biol 2013*]. CNA results were used to explore CN state of mutated genes. To collect and share these results, we created a free and public **Melanoma Exome Database**.

On the same samples, we carried out **RNA-sequencing** and performed both a traditional **gene expression analysis** and more sophisticated **structural evaluations**. Focusing on **fusion transcripts**, we identified 72 putative events generated by either inter-chromosomal translocations or intra-chromosomal rearrangements, recently defined "conjoined genes" and representing an additional gene regulation mechanism.

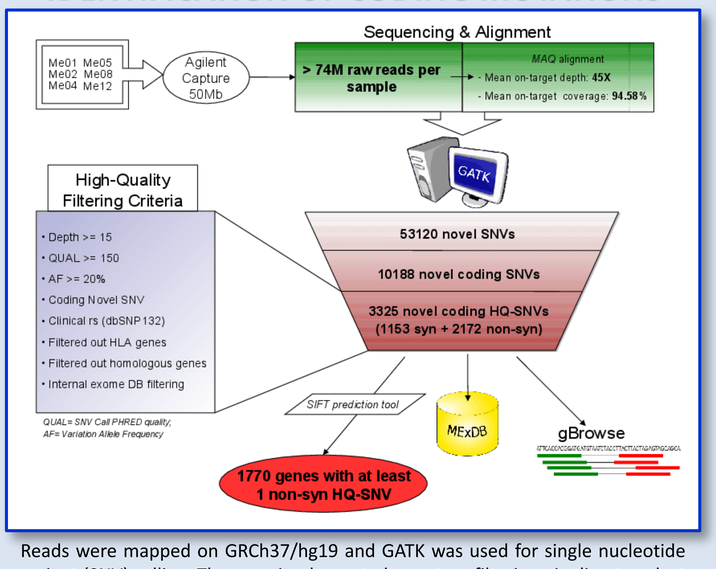
Globally, NGS proved to be extremely powerful to dissect cancer complexity at both genomic and transcriptomic levels, and to identify novel potential targets for personalized treatment of cutaneous melanoma.

WHOLE-EXOME CAPTURE

From 3ug DNA, we used the Agilent Sure Select Human All Exon 50 Mb kit to capture coding regions. Sequencing was done on Illumina GAIIx, in 76-cycle paired-end runs.



BIOINFORMATICS CUSTOM PIPELINE FOR IDENTIFICATION OF CODING MUTATIONS



Reads were mapped on GRCh37/hg19 and GATK was used for single nucleotide variant (SNV) calling. Then, we implemented a custom filtering pipeline to select novel coding highly confident variants (HQ-SNVs).

CREATION OF A COMPREHENSIVE AND UPDATED LIST OF MELANOMA-RELATED GENES

Canonical KEGG Melanoma Pathway (map05218) *30 gene families, 195 genes* + 10 NGS recent papers (2008-2012) *26 gene families, 352 genes* = **547 potentially melanoma-related genes**

We manually compiled an updated list of melanoma-related genes by combining already existing knowledge (KEGG Melanoma Pathway) and recent literature.
We used this list (547 genes) to screen our samples for known and novel mutations.

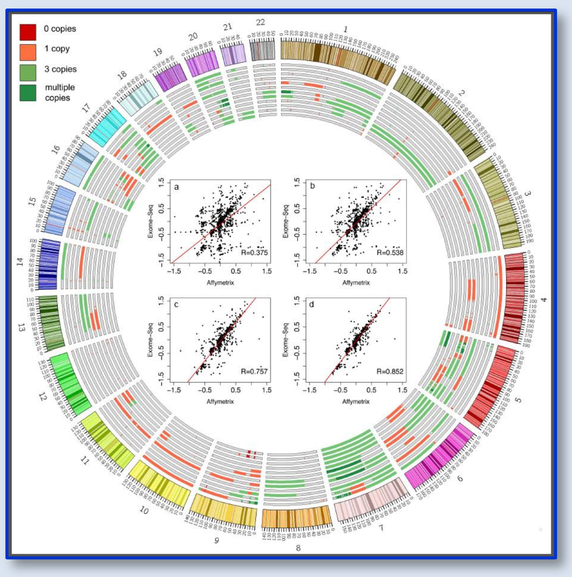
AFFECTED GENES AND PATHWAYS RELEVANT FOR MELANOMA AND TARGETED DRUGS

	Me01	Me02	Me04	Me05	Me08	Me12	NS/S
Signalling cascades							
PI3K/Akt signalling	*PIK3R4*	*PIK3C2G*[*§]	*PIK3C2G, PREX2*[§]	-	*PIK3CG, PREX2*[§]	-	**10/1 (10)**
MAPK signalling	*MAP3K4, BRAF*	-	*BRAF*[Δ]	*MAPK6*	*MAP2K3, BRAF*	*MAP3K5, BRAF*	**8/3 (2.7)**
Glutamate signalling	*GRIN2B, GRIN3A, GRM5, PLCB1, PLCB4, PLCE1, PLCZ1*	*GRIN3A*	*GRM1, PLCXD2*	-	-	-	10/13 (0.8)
Molecular function classes							
RAS-RAF Ser/Thr kinases	*BRAF*	*NRAS*[Δ]	*BRAF*[Δ]	-	*BRAF*	*BRAF*	**5/0**
Protein tyrosine kinases	*FGFR1*	*MET*	*PTK2B*	*PTK7*	-	-	4/2 (2)
Metalloproteinases	*ADAM22*[&], *ADAMTS18, MMP24, MMP25*	*ADAMTS9*	*ADAMTS12**	-	*ADAM23, ADAMTS6, ADAMTS9, MMP19*	-	10/7 (1.4)
Protein tyrosine phosphatases	*PTPN1, PTPRK*	*PTPN13*	*PTPN13, PTPRF*[Δ]	*PTPRD*[&]	-	*PTPLA*[Δ]	7/3 (2.3)
G protein-coupled receptors	*GPR64, GPR101*[Δ], *GPR112*[Δ], *GPR158, GRM5*	*GPR113*[Δ]	*GPR151*[*Δ], *GRM1*	*GPR112*[Δ], *GPR113, GPR133*	*GPR158*	-	12/5 (2.4)

Legend: *stop codon; [§] double-mutated in the same sample; [Δ] homozygous mutation; [&] deletion coupled with mutation of the remaining allele. Genes mutated in at least two samples are underlined. NS/S, ratio of non-synonymous to synonymous mutations, it indicates potential driver role.

Mutated melanoma-related genes were grouped according to pathway or molecular function in order to highlight particularly impacted processes. Some of these pathways, such as RAS-RAF, MAPK and PI3K/Akt, are already addressed by current melanoma targeted drugs. Thus, our results might give indications for novel targets for personalized combined approaches.

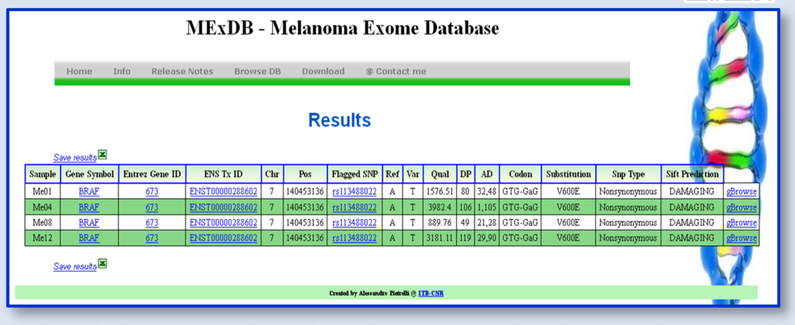
COPY NUMBER ANALYSIS FROM WES DATA USING EXCAVATOR



Excavator software identified gain and loss regions in each sample compared to a diploid reference, starting from WES data. By Circos plot, we visualized and compared copy number alterations (CNA) detected by WES and by SNP array.

CREATION OF A PUBLIC DATABASE FOR DATA STORAGE AND SHARING

(MExDB @ *https://155.253.6.64/MExDB/*)



Full details about HQ-SNVs detected in each sample were collected in our public MExDB and can be visualized by the linked gBrowse genome viewer tool, together with CNA regions.

Whole-exome seq

RNA-seq

ESTIMATION OF TRANSCRIPT EXPRESSION LEVELS (FPKM)

RNA-seq reads were mapped with TopHat on hg18 and used to perform the quantitative estimation of transcript expression levels (FPKM), using Cufflinks and RefSeq gene model database (36,098 transcripts).
When compared to microarray, RNA-seq offers a wider dynamic range, greater sensitivity for low expressed genes, better discrimination of expression differences, and distinct values for any transcript isoform, both known and novel.

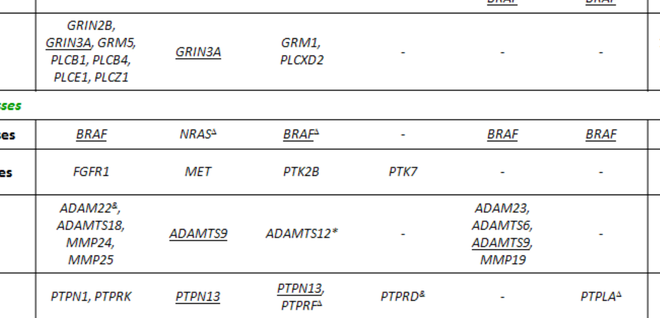
VARIANT CALLING ON TRANSCRIPTOME

By using CLC and GATK, we performed variant calling on RNA-seq reads. This strategy allows to narrow the search for sequence changes, focusing only on expressed genes.

When having also WES calls for the same sample, you can compare variants present on RNA to those on DNA, and elucidate a lot of interesting underlying mechanisms !!!

TRANSCRIPTOME	EXOME	Mechanism
het	het	Balanced expression
hom	het	Allele-specific expression
hom	hom (punctual or extended)	Single variant in homozygous state in genome Deletion with LOH or CNN-LOH

FUSION TRANSCRIPTS



By using FusionMap and RefSeq gene model database, we identified fusion transcripts, which can generate from either inter-chromosomal translocations, or intra-chromosomal rearrangements, such as deletions or non-traditional splicing between adjacent genes (the so-called "conjoined genes").

Events	Type	Origin
72 putative fusion transcripts	38 inter-chr	Translocations between two different chrs
	34 intra-chr	11 deletions
		23 conjoined genes

CONTACT
Mail: **ingrid.cifola@itb.cnr.it**
Phone: +39-02-26422738
Address: via F.lli Cervi 93, 20090 Segrate (Milan), Italy

ACKNOWLEDGEMENTS
This work was supported by Cariplo Foundation grant (Genomic, epigenetic and transcriptional analysis of cancer by next-generation sequencing, Rif. 2006.0771).