

This is the final peer-reviewed accepted manuscript of:

Giannuzzi D, Marconato L, Elgendy R, et al. Longitudinal transcriptomic and genetic land-scape of radiotherapy response in canine melanoma. *Vet Comp Oncol.* 2019;17:308–316.

The final published version is available online at: <https://doi.org/10.1111/vco.12473>

#### Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

*This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)*

***When citing, please refer to the published version.***

# **Longitudinal transcriptomic and genetic landscape of radiotherapy response in canine melanoma**

## **ABSTRACT**

Canine malignant melanoma is a highly aggressive tumor with a low survival rate, and represents an ideal spontaneous model for the human counterpart. Considerable progress has been recently obtained, but the therapeutic success in canine melanoma is still challenging. Little is known about the mechanisms beyond pathogenesis and melanoma development, and the molecular response to radiotherapy has never been explored before. A faster and deeper understanding of cancer mutational processes and developing mechanisms are now possible through next generation sequencing technologies, and they can now be applied in canine species. In this study, we matched whole exome sequencing and transcriptome sequencing in four dogs affected by malignant melanoma at diagnosis and at disease progression to identify possible genetic mechanisms associated with therapy failure. According to previous studies, a genetic similarity between canine malignant melanoma and its human counterpart was observed. Several somatic mutations were functionally related to MAPK, PI3K/AKT and p53 signaling pathways, but located in genes other than BRAF, RAS and KIT. At disease progression, several mutations were related to therapy-induced genomic effects. Natural killer cell-mediated cytotoxicity and several immune-system-related pathways resulted activated opening a new scenario on the role of T-cell lymphocytes in this tumor. In conclusion, this preliminary study suggests a potential role of the immune system associated to radiotherapy in canine melanoma, but a larger sample size associated with functional studies are needed.

## **KEY WORDS**

Dog, Malignant Melanoma, Radiation therapy, Exome sequencing, RNA-seq

## INTRODUCTION

Malignant melanoma (MM) is among the deadliest forms of cancer, and it is characterized by a high degree of local invasiveness and high metastatic propensity (Rohli). In dogs, MM frequently involves the oral cavity, followed by mucosal junctions, feet (including digits) and haired skin (Spangler). Regardless of the anatomical location, complete surgical resection associated with regional lymphadenectomy and/or radiotherapy is the preferred treatment modality. Systemic chemotherapy has yielded controversial and quite disappointing results. Recently, attention has been paid to the role of the immune system in the pathogenesis and progression of canine melanoma, resulting in the development of two immunotherapeutic approaches, namely a xenogeneic DNA vaccine against tyrosinase (Oncept, Merial), and a CSPG4-immunotargeting vaccine.<sup>1</sup>

Despite a significant improvement in treatment during the last years, canine MM continues to be a major clinical challenge. Consequently, there is a need for further research on disease etiology and pathogenesis, possibly leading to the identification of biomarkers and drug targets, ultimately translating into an improvement in early diagnosis and personalized treatments. Although incompletely outlined, the genetic profile of canine MM has revealed similarities with the human counterpart. The mitogen-activated protein kinase (MAPK) and phosphoinositide 3-kinase (PI3K) pathways were found frequently activated in dogs,<sup>2-5</sup> conversely activating mutations in *BRAF*, *NRAS* and *KIT* were scarcely identified.<sup>4,5,6</sup> In addition, Poorman et al.<sup>7</sup> described amplification of *MYC* and *KIT* loci, deletion of *CDKN2A* locus and reported distinctive copy number abnormalities in canine chromosome 30 encoding MAPK pathway genes, such as *SPRED1* and *TRPM7*.

Whole-Exome Sequencing (WES) has been extensively applied in human medicine for several diseases (reviewed in Retterer et al.<sup>8</sup>) and recent studies in human oral MM have identified potential therapeutic targetable mutations.<sup>9</sup> In dogs, studies to identify global somatic mutations by next generation sequencing (NGS) were conducted only in lymphoma<sup>10</sup> and in rare congenital<sup>11</sup>, retinal<sup>12</sup> and neurodegenerative<sup>13</sup> disorders.

In the current study, WES was performed in dogs with treatment-naïve MM at diagnosis and at disease progression, to highlight the genetic mechanisms beyond progression. In addition, whole-transcriptome sequencing (RNA-seq) was performed on the same samples. The main objectives were (i) to characterize the mutational signature via the identification of somatic short and structural variants of canine MM both at diagnosis and at disease progression, (ii) to combine gene mutations with gene

expression alterations by functional analysis and finally (iii) to compare gene expression profiles at diagnosis and disease progression.

## **MATERIALS AND METHODS**

### ***Animal recruitment and sample acquisition***

Client-owned dogs affected by MM, undergoing a complete staging work-up and receiving treatment were eligible for recruitment. At first presentation (T0), all dogs had to undergo a surgical biopsy to confirm the diagnosis (Table S1). Furthermore, all dogs had to be followed-up and undergo to a second surgical biopsy at disease progression (T1). Signed owner consent was always obtained before inclusion.

DNA and RNA extraction, WES and RNA-seq library preparation and sequencing are described in Supplementary Material and Methods.

### ***RNA-Seq data analysis***

The raw 50 bp single-end sequences (Sanger/Illumina 1.9 encoding) were quality-controlled by FastQC v.0.11.4 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), and the low-quality bases (quality scores < 20) and adaptor contamination (if present) were removed by Trimmomatic<sup>14</sup> v.0.36 using the parameters 'ILLUMINACLIP:TruSeq3-SE:2:30:10 LEADING:3 SLIDINGWINDOW:4:15 MINLEN:25'. Quality assessment by FastQC was performed again on the clean reads to ensure contaminant-free data. The high-quality reads were mapped to the CanFam3.1 genome assembly from Ensembl ([ftp://ftp.ensembl.org/pub/release-90/fasta/canis\\_familiaris/dna/](ftp://ftp.ensembl.org/pub/release-90/fasta/canis_familiaris/dna/)) by HISAT2<sup>15</sup> v.2.1.5. The uniquely-mapped reads aligned to exons were counted with HTSeq<sup>16</sup> v.0.6.1, then tested - by the DESeq2<sup>17</sup> R package v.1.14.1 - for the presence of differentially expressed genes (DEGs). Genes with a false discovery rate (FDR)  $\leq 0.1$  and at least 2-fold change ( $\text{Log}_2\text{FC} > 1$  or  $\text{Log}_2\text{FC} < -1$ ) in their expression between T0 and T1 were considered DEGs. The sequencing data (FASTQ files) associated with this study were deposited in the GenBank's Sequence Read Archive (SRA) under the accession ID SRP150063.

When comparing T0 versus T1 samples, the enriched gene ontology (GO) terms and signaling pathways associated with DEGs were identified by the Gene Set Enrichment Analysis (GSEA; <http://www.broadinstitute.org/gsea/index.jsp>), using the GSEA pre-ranked option. Briefly, all Ensembl

gene IDs were collapsed to their corresponding HUGO gene symbols, then the entire normalized transcriptome dataset was ranked by the logarithm transformed (base 2) FC, where the up- and down-regulated genes were assigned positive and negative values, respectively. The pre-ranked dataset was analyzed (1000 permutations) against the curated canonical KEGG pathways (c2.cp.kegg.v6.1), GO terms (c5.all.v6.1) and Hallmarks (h.all.v6.1) catalogs, from the Molecular Signatures Database<sup>18</sup> (MsigDB). In the present study, all gene sets with a FDR  $\leq$  25% were considered, but only those with a FDR  $\leq$  10% are reported here, as they represent the most negatively or positively enriched (based on the normalized enrichment score; NES) gene sets in the T1 samples.

### ***Somatic mutations discovery***

For WES data processing, the paired-end reads sequence quality was first verified using the reports generated by FASTQC toolkit (Babraham Institute, Cambridge, UK). Trimmomatic<sup>14</sup> v.0.36 was used to clip adapters and trim low-quality regions accordingly (parameter ILLUMINACLIP:TruSeq3-PE-fa-2:10:10:8 SLIDINGWINDOW:5:20 HEADCROP:2 MINLEN:70). The reads were aligned to the CanFam3.1 reference genome (Broad Institute, Cambridge, MA; released Sep. 2011; downloaded from Ensembl release 87) using BWA<sup>19</sup> mem v.0.7.7. SAMtools<sup>20</sup> v.1.8 flagstat was used to evaluate the number and percentage of reads that mapped on the genome. Sorting, flagging of PCR duplicates and read group assignment were performed by means of PICARD toolkit (<http://broadinstitute.github.io/picard/>). A local realignment, using the Genome Analysis Toolkit<sup>21</sup> (GATK) v.3.7 tools RealignerTargetCreator and IndelRealigner, were performed to correct any misalignment around small insertions and deletions (indels). Then, the GATK tools BaseRecalibrator and PrintReads were used for base quality score recalibration. On the resulting BAM files, single nucleotide variants (SNVs) and indels were identified with Mutect2<sup>22</sup> and VarScan<sup>23</sup> v.2.4.1, described as the two outperforming variant callers in previous studies.<sup>24,25</sup> Additionally, loss of heterozygosity (LOH) was further investigated by VarScan. The dbSNP database (version 146) was used as a source of known variants to localize any region needing realignment (IndelRealigner), to ignore known mismatch on the aligned reads (BaseRecalibrator) and to distinguish true somatic variants from false positives by establishing stricter logarithm of odds (LOD) score threshold for the normal samples (Mutect2). Then, variants identified by Mutect2 and VarScan to be somatic rather than germline were selected. To obtain high-confidence calls, default hard filters made by Mutect2 were applied and only variants flagged as “pass” were retained. Conversely, for VarScan the option --tumor-purity equal to 0.90 for tumor samples was added, and SNVs, indels and somatic loss of heterozygosity (LOH) with a p-value  $\leq$  0.05 were

selected. Finally, filtered SNVs and indels were annotated with ANNOVAR<sup>26</sup> and Variant Effect Predictor<sup>27</sup> (VEP). LOHs were annotated with Homer.<sup>28</sup> For each sample, SNVs, indels and LOH loci were considered as the union of Mutect2 and VarScan filtered somatic mutations. For further analysis, only those SNVs and indels that were shared by a minimum of two dogs were employed. In order to assess the impact of tumorigenic mutations on gene expression, mRNA levels of genes harboring somatic SNVs were retrieved. A customized perl script was then applied to extract the read coverage of each allelic variant (i.e. reference and alternate allele), only genes with a minimum total coverage of 15 were investigated. Raw Illumina reads were deposited in the GenBank's SRA repository under the accession ID SRP151732.

### ***Copy Number Aberration analysis***

For each sample, WES data were analyzed by NEXUS Copy Number v9.0 software (Biodiscovery Inc., CA, USA) to investigate genomic copy number variation. Copy number aberrations (CNAs) were identified starting from deduplicated BAM files with a matched approach ("BAM ngCGH-matched"). Then, a FASST2 segmentation algorithm with a significance threshold of 5.5e-06 was applied. Aberrations were defined as a minimum of 3 consecutive segments with log2 tumor/reference ratio value of >1.0 (high gain), 1.0 to 0.2 (gain), -0.25 to -1.5 (loss), <-1.5 (big loss). Recurrent CNAs across multiple samples were determined within NEXUS using an involvement threshold of 50%. Statistical significance was then determined using the GISTIC algorithm with a G-score cut off of  $G > 1.0$  and a significance of  $Q < 0.05$ . Copy number aberration frequency comparisons amongst T0 and T1 were performed in NEXUS using Fisher's exact test with differential threshold of > 50 % and significance p-value < 0.05.

### ***Functional analysis***

To call variants that differ between T0 and T1, VarScan was used with the *somatic* functionality. Variants were annotated with ANNOVAR and for intergenic locations only variants within 5 kb distance were retained. Then, to extract biological information by the identification of significant gene ontologies and KEGG pathways, variants were collapsed to genes and mutated genes (MGs) were analyzed by the Cytoscape plugin ClueGO.<sup>29</sup> In this study, a right-side hypergeometric test was used for the p-value calculation, followed by the Benjamini-Hochberg adjustment for multiple test correction.<sup>30</sup> A pathway/ontology with a FDR <0.05 was considered significant. Based on the *kappa* score threshold ( $\geq 0.4$ ), the significant pathways/ontologies were functionally clustered into groups and displayed in a

network, in which a node represented a biological function, and an edge between the nodes indicated that genes are shared between the pathways.

## RESULTS

### *Clinico-Pathological features*

Four dogs were enrolled: 1 mixed breed, 1 Golden retriever, 1 Labrador retriever and 1 English setter. There were 3 intact males and 1 spayed female. Age ranged between 8 and 13 years, and weight ranged between 20 and 38 kg. Three dogs had an oral MM, and 1 dog had a nail bed MM. Three MMs were staged T2N0M0, and 1 T3N1M0. All dogs were irradiated with 5 bi-weekly fractions of 6 Gy each, up to a total dose of 30 Gy; 3 dogs received adjuvant chemotherapy and 1 dog was treated with immunotherapy. At the time of disease progression, the tumor was biopsied again in all dogs, after 139, 175, 204, and 231 days, respectively, from diagnosis. At the end of the study, all dogs had died because of local progression (n=2) after 139 and 270 days, respectively, or distant metastasis to skin (n=1) after 188 days, or lungs and liver (n=1) after 221 days from diagnosis.

### *Gene expression profiling reveals immune signatures differentiating MMs at first presentation versus disease progression*

Four pairs of matched melanomas obtained from dogs at T0 and T1 were analyzed by RNA-Seq. A summary of the sequencing output, quality scores, and mapping percentage is reported in Table S2. Comparing the transcriptomes of T1 and T0 samples (T1 vs T0), a total of 25 and 29 protein-coding genes were found differently regulated in the biological and technical replicates, respectively (Table S3). To define a list of non-biased DEGs, only those genes that overlapped (n=18) between the two independent analyses (biological and technical replicates) were considered as the final DEGs list (Table S3). Of the 18 DEGs, 15 genes were downregulated at T1, with *ASPRVI*, *MSTN* and *SPINK7* being the most downregulated genes, while the other 3 genes were upregulated (Table S3). Interestingly, by GSEA, immune system-related pathways were mainly enriched in T1, including genes associated with natural killer cell mediated cytotoxicity, chemokines and T-cell receptors. A complete list of all enriched GO terms, KEGG pathways, Cancer Hallmarks and the relative enrichment plots are reported in Table S4, Figure S1, Figure S2 and Figure S3.

### ***CNA analysis identifies a similar pattern of chromosomal aberrations in MMs at first presentation and disease progression***

The CNA analysis showed both focal and broad chromosomal aberrations in all MMs, totaling from 10 Mb to 100 Mb and affecting hundreds of genes (Figure 1). Noteworthy, amplifications prevailed over deletions. The extension and the number of CNAs was increased in T1 compared to T0. However, the comparison of CNA frequency by GISTIC revealed no significant regions distinguishing firstly-occurring and progressed tumors. Analyzing T0 and T1 separately, no regions reached the threshold for recurrent CNAs calling. Conversely, exploring all the MMs together, recurrent gains were found on *Canis Familiaris* autosome (CFA) 10, 27, 30 and 38 ( $G > 4.8$  and  $Q < 1.945e-02$ ), whereas CFA 2, 11, 30, 31 and 38 presented recurrent losses ( $G > 2.4$  and  $Q < 3.877e-02$ ) (Table S5). The most frequent copy number gains were located on CFA 6, 8, 9, 17 and 25 (Table S6) and occurred in all tumors. The most recurrent copy number loss (50% samples at T0, and 75% at T1) involved CFA 11 and was a ~600kb region located at 41,055,923-41,652,538 (Table S6) encompassing *CDKN2A* locus. A focal amplification located on CFA 10 and corresponding to *MDM2* locus (50% samples at T0 and 50% at T1) was also identified. Interestingly, 2 out of 4 dogs showed a characteristic aberration on CFA 30 displaying a sigmoidal pattern of copy number loss (spanning 0.5-12 Mb) followed by a gain (spanning 9.5-19 Mb). In this articulated region, the amplification contained 12 gene loci, including *SLC27A2*, *GABPB1*, *USP8*, *TRPM7*, *SPPL2A* and *CYP19A1*, whereas *SPRED1* and *FAM98B* were located within the region of copy number loss.

### ***WES shows an additional burden of mutations in MMs at disease progression***

At T0, a total of 6,807 somatic mutations were found: 4,036 SNVs, 971 indels and 1,800 LOH calls, respectively (Table S7). However, only 1 non-synonymous SNV, four non-coding SNVs and 1 frameshift substitution were in common in at least 3 tumors and none when considering all T0 samples (Figure 2A and 2C). The median tumor mutational burden (TMB) was 4.01 coding mutations/Mb (range 3.37-5.63). At T1, 8,246 somatic mutations were identified: 4,866 SNVs, 1,094 indels and 2,286 LOH calls, respectively. Two non-coding SNVs were shared by all tumors, while 5 non-coding SNVs and 1 LOH call were in common in 3 dogs (Figure 2B and 2D) The median TMB was 5.1 coding mutations/Mb (range 4.6-6.16).

A total of 8,902 SNVs were identified at T0 and T1, and the ratio of transitions (5,779) to transversions (3,123) was 1.85. Also, a significant higher rate of C>T/G>A transitions (Fisher's test,  $p < 0.01$ ) was calculated (Table S8 and Figure 3). Using signatures of mutational processes published by Alexandrov<sup>31</sup>



in 2011, MMs revealed a prevalence of signature 1B, 7 and 11, resembling human head and neck cancers and melanoma.

A total of 101 and 109 somatic mutations (SNVs and indels), which were shared by at least 2 dogs, were observed at T0 and at T1, respectively (Table S9). By comparing the somatic mutations in each dog at diagnosis and disease progression, on average the 15.5% (10.3%-23.6%) was found in common, the 20.7% (15%-27.6%) of which was exonic with predicted protein sequence effects (Table S10). Overall, most of the somatic mutations were identified in intronic and intergenic regions. A total of 17 (17%) and 18 (16.5%) somatic mutations were identified in the exons at T0 and T1, respectively. Distribution and consequences of coding mutations according to ANNOVAR and VEP are depicted in Figure 3. Interestingly, 15 somatic mutations were in common across all samples including genes such as *TMEM256-PLSCR3*, *DCAF10*, *LAMA3*, *GALC*, *SMARCAL1*, *TSPYL1*, *SOX17*, *MCM5*, and *CYP2A13*. The frame shift substitutions harbored on *TMEM256-PLSCR3* (chr5:32330460:A>-C) and on *SOX17* (chr29:5979449:G>-TT) and the non-silent SNV on *GALC* (chr8:59317742:A>C) were the sole coding mutations. A total of 94 mutations were found in at least 50% of T1 samples and never detected at T0. Exonic mutations were 7 non-silent SNVs, 1 non-frameshift substitution and 1 non-frameshift deletion. Notably, the non-frameshift deletion was located on *TRPM1* (chr3:37948404:CCC>-) and one of the non-silent SNV on gene *eEF1A1* (chr3:60899961:G>A).

When integrating somatic mutations with gene expression data, an average of 70% of mutations was expressed. Only 11 and 9 somatic mutations were supported by RNA-seq reads in at least 2 dogs at T0 and T1, respectively. When comparing expressed mutations between T0 and T1, only a small number was in common (26.5%). Genes belonging to keratin family and keratin associated proteins (KRTAP) were considered to be related to the sampling site, rather than being associated to the disease itself.<sup>32</sup> The coding mutation located on *GALC* gene was covered by RNA-seq reads in 6 MMs.

To investigate the presence of RNA-seq reads that supported genes with somatic mutations, SNVs and indels were collapsed to genes. Particularly attractive were genes including *AHNAK*, being mutated in 6 out of 8 samples, *LINS1* and *UBAI* at T1, and *PRKCDBP* at T0. A functional analysis was also performed using Varscan and a total of 2,471 variants were identified differentiating MMs at T0 and T1. When variants were collapsed to genes, a total of 12 genes were shared among all dogs. Using ClueGO with the MGs in common in at least 2 samples (n=171), pathways related to drug and hormone metabolism were significantly overrepresented (Table S11 and Figure S4). However, considering the low number of DEGs identified in the differential expression analysis, no overlapping emerged comparing DEGs and MGs.

## DISCUSSION

Understanding the relationship between therapeutic treatments and the dynamics of mutations in naïve and progressed tumors has dramatically changed the therapeutic perspectives in many human cancers. Overall and event free survival in dogs with MM are relatively short, and the molecular mechanisms beyond tumor progression have never been investigated. In this study, we used WES, incorporating a comprehensive analysis of copy number, SNVs, and mutational signatures, from prospectively collected canine MM to compare primary tumors with matched samples at disease progression after radiotherapy. Furthermore, we performed RNA-seq on the same samples to evaluate the phenotypical impact of variants and to identify mutated alleles that are then RNA transcribed.

When considering mutational profiles, treatment-naïve melanomas showed similar signatures to human melanoma and TMB also reflected the high complexity;<sup>33,34</sup> furthermore, a significant higher rate of C>T/G>A transitions was found. At progression, the signatures remained unchanged, whereas TMB was increased. By definition, TMB measures the number of mutations carried by tumor cells. Even if increased, TMB may be associated in primary instance to tumor clonal evolution after treatment due to selective pressure. Recent studies have reported that tumors with a higher number of mutations have a higher likelihood to respond to immunotherapy, but this is still unexplored in dogs.

Due to the low number of cases in our series, GISTIC algorithm didn't identify chromosomal aberrations being able to differentiate MM at first presentation and at progression. However, interesting data were obtained when comparing specific aberrations between matched tumors. In two dogs, a peculiar sigmoidal pattern aberration on CFA 30, that was previously described by Poorman et al.<sup>7</sup> in canine oral MM, was observed in treatment-naïve melanomas and retained at progression. The amplification comprises important genes, such as *SLC27A2*, *GABPB1*, *USP8*, *TRPM7* and *SPPL2A*. In the same genomic region, loss of *SPRED1* was also found. The deletion of *SPRED1* and the amplification of *TRPM7* have been demonstrated to cause activation of MAPK signaling pathway.<sup>4,7</sup> Examining RNA-seq data, *TRPM7* expression was increased up to 20 times in tumors presenting the amplification than in MM showing no aberration. Similarly to other canine tumors, loss in *CDKN2A* locus was observed in two dogs both at T0 and T1, and one dog acquired the deletion at disease progression.<sup>7,35</sup> In human mucosal melanoma, the *CDKN2A* locus is a commonly deleted region, and inactivation leads to unregulated cell proliferation<sup>36</sup>. A recurrent focal amplification on *MDM2* locus was identified,

underlining an additional interaction with p53 signaling pathway. Moreover, tumors with amplification showed up to 19 times higher than MM with no aberrations. *MDM2* is a p53 negative regulator promoting tumor formation by targeting several tumor suppressor proteins.<sup>37</sup> This gene itself is transcriptionally-regulated by p53, but overexpression due to amplification was reported in a variety of different human cancers, including melanomas, but never reported in dogs before.<sup>37</sup>

Unexpectedly, when comparing mutations at T0 and T1, only a small fraction was in common; furthermore only 20% of these mutations were exonic with predicted protein sequence effects. However, this reflects a global result that was obtained from WES analysis. Indeed, the majority of mutations were identified in the intronic and intergenic regions, revealing a significant percentage of off-target sequences in the canine exome enrichment protocol.<sup>38</sup> This result also explains the observed transitions/transversions ratio (2.0) that is generally associated to SNVs located in regions other than exome.<sup>39</sup> Finally, a more pronounced selective pressure in the exonic regions of the tumor might also explain a higher variation in noncoding regions.<sup>40</sup>

Among the genes affected by exonic mutations both at T0 and T1, *GALC* showed a non-silent SNV that was confirmed by the absence of reads in RNA-seq. *GALC* encodes a lysosomal protein, which hydrolyzes glycosphingolipids leading to the consequent production of bioactive lipids that negatively affect cancer progression.<sup>41,42</sup> Also, *Sox17* harbored at the genomic level a frameshift mutation that was not supported by RNA-seq reads. This transcription factor is considered an antagonist to canonical Wnt/ $\beta$  catenin signaling in several malignant tumors, and its down-regulation is associated with progression in both primary and metastatic human melanomas.<sup>43</sup> Conversely, target melanoma oncogenes, such as *BRAF*, *RAS* and *KIT*, did not harbor variants, thereby confirming previous findings.<sup>4-6</sup>

When exploring genes harboring mutations only in T1, a non-frameshift deletion located on *TRPM1* was identified. The number of reads mapping *TRPM1* was reduced compared to dogs not harboring this mutation. *TRPM1* gene encodes for a calcium permeable cation channel involved in regulation of melanocyte physiology and melanoma development.<sup>44</sup> Loss of *TRPM1* expression has been correlated to melanoma aggressiveness<sup>45</sup> and is now considered an independent prognostic marker of disease free and overall survival in human MM.<sup>46,47</sup> Our results showed a relationship between *TRPM1* non-frameshift mutation and gene expression but no association with event free or overall survival was obtained, possibly due to the low number of samples. However, no radiotherapy effects are documented for *TRPM1* molecular changes. A non-silent SNV was located on *eEF1A1* and further characterized by absence of reads on variant allele. *eEF1A1* has been proposed as a premature senescence marker<sup>48</sup> and its expression was found to be decreased after radiotherapy both *in vitro* and *in vivo*, suggesting that

premature senescence represents a significant pathway in radiation-induced tumor regression.<sup>49</sup> Mutations on *eEF1A1* might be one of the causative mechanisms by which radiotherapy induces cellular senescence contributing to tumor response.

We next investigated RNA-seq data to compare transcriptome of naïve melanomas with matched samples at progression. A positive enrichment of several immune-system-related pathways, such as natural killer cell-mediated cytotoxicity, was found in progressed MM. Melanoma is an immunogenic tumor and its relationship with the host immune system is currently under investigation<sup>50</sup>. Different studies have shown the capacity of melanoma to activate or modulate the adaptive immune response, resulting in a balance between survival and proliferation of tumor cells.<sup>50,51</sup> Interestingly, tumors that progressed after radiotherapy were able to selectively modulate the immune system in their favor. Our data opens a new perspective on the role of this lymphocytes subset in canine MM after radiotherapy.<sup>50,52</sup> In MMs at progression, a negative enrichment of several fatty acid metabolism pathways was also identified. Even if never described in dogs, this data is not surprising considering the ability of cancer cells to reprogram their metabolic pathways and microenvironment to satisfy both proliferation and survival demands.<sup>53-55</sup> This ability should be considered in future targeted therapies.

In conclusion, the preliminary results from our study confirm the genetic similarity of canine MM with its human counterpart.<sup>2,56,57</sup> Genetic molecular mechanisms also seem to be more related to copy number aberrations rather than point mutations. Notably, somatic mutations are located in genes other than *BRAF*, *RAS* and *KIT*, nevertheless modulating the same signaling pathways, such as MAPK, PI3K/AKT and p53 signaling pathways. We identified several allelic mutations related to therapy-induced genomic effects. A larger number of cases and functional studies will help to identify their biological functions in future studies. Finally, on one side, transcriptomic pathway analysis has revealed the importance of adaptive immune system regulation and on the other side the increased of TMB in disease progression showed a higher likelihood to respond to immunotherapy, opening the opportunity to further investigate the potential of immunotherapy in canine melanoma.

## **CONFLICT OF INTEREST**

The authors declare no conflict of interests.

## REFERENCES

1. Rolih V, Barutello G, Iussich S, et al. CSPG4: a prototype oncoantigen for translational immunotherapy studies. *J Transl Med.* 2017;**15**: 151.
2. Gillard M, Cadieu E, De Brito C, et al. Naturally occurring melanomas in dogs as models for non-UV pathways of human melanomas. *Pigment Cell Melanoma Res.* 2014;**27**: 90-102.
3. Simpson RM, Bastian BC, Michael HT, et al. Sporadic naturally occurring melanoma in dogs as a preclinical model for human melanoma. *Pigment Cell Melanoma Res.* 2014;**27**: 37-47.
4. Fowles JS, Denton CL, Gustafson DL. Comparative analysis of MAPK and PI3K/AKT pathway activation and inhibition in human and canine melanoma. *Vet Comp Oncol.* 2015;**13**: 288-304.
5. Shelly S, Chien MB, Yip B, et al. Exon BRAF mutations are uncommon in canine oral malignant melanomas. *Mamm Genome.* 2005;**16**: 211-7.
6. Murakami A, Mori T, Sakai H, et al. Analysis of KIT expression and KIT exon 11 mutations in canine oral malignant melanomas. *Vet Comp Oncol.* 2011;**9**: 219-24.
7. Poorman K, Borst L, Moroff S, et al. Comparative cytogenetic characterization of primary canine melanocytic lesion using array CGH and fluorescence in situ hybridization. *Chromosome Res.* 2015;**23**: 171-86.
8. Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med.* 2016;**18**: 696-704.
9. Lyu GY, Yeg YH, Yeh YC, Wang YC. Mutation load estimation model as a predictor of the response to cancer immunotherapy. *NPJ Genom Med.* 2018;**3**: 12.
10. Elvers I, Turner-Maier J, Swofford R, et al. Exome sequencing of lymphomas from three dog breeds reveals somatic mutation patterns reflecting genetic background. *Genome Res.* 2015;**25**: 1634-45.
11. Evans JM, Cox ML, Huska J, et al. Exome sequencing reveals a nebulin nonsense mutation in a dog model of nemaline myopathy. *Mamm genome.* 2016;**27**: 495-502.
12. Kropatsch R, Akkad DA, Frank M, et al. A large deletion in RPGR causes XLPPRA in Weimaraner dogs. *Canine Genet Epidemiol.* 2016;**3**: 7.
13. Tsuboi M, Watanabe M, Nibe K, et al. Identification of the PLA2G6 c.1579G>A Missense Mutation in Papillon Dog Neuroaxonal Dystrophy Using Whole Exome Sequencing Analysis. *PloS One.* 2017;**12**: e0169002.
14. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.

- Bioinformatics*. 2014;**30**: 2114-20.
15. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015; **12**: 357-60.
  16. Anders S, Pyl PT, Huber W. HTseq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015; **31**: 166-9.
  17. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;**15**: 550.
  18. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;**102**: 15545-50.
  19. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; **25**: 1754-60.
  20. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; **25**: 2078-9.
  21. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;**20**: 1297-303.
  22. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013; **31**: 213-9.
  23. Koboldt DC, Larson DE, Wilson RK. Using VarScan 2 for Germline Variant Calling and Somatic Mutation Detection. *Curr Protoc Bioinformatics*. 2013; **44**: 15.4.1-17.
  24. Alioto TS, Buchalter I, Derdak S, et al. A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing, *Nat Commun*. 2015;**6**: 10001.
  25. Cai L, Yuan W, Zhang Z, et al. In-depth comparison of somatic point mutation callers based on different tumor next-generation sequencing depth data, *Sci Rep*. 2016;**6**: 36540.
  26. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;**38**: e164.
  27. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;**17**: 122.
  28. Heinz S, Benner C, Spann N, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell*. 2010;**38**: 576-589.

29. Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. 2009;**25**: 1091- 3.
30. Thissen D, Steinberg L, Kuang D. Quick and easy implementation of the Benjamini-Hockberg procedure for controlling the false positive rate in multiple comparisons. *J Educ Behav Stat*. 2002;**27**: 77-83.
31. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;**500**: 415-21.
32. Laquer VT, Hevezi A, Albrecht H, Chen TS, Zlotnik A, Kelly KM. Microarrays analysis of port wine stains before and after pulsed dye laser treatment. *Lasers Surg Med*. 2013; **45**: 67-75.
33. Furney SJ, Turajlic S, Stamp G, et al. The mutational burden of acral melanoma revealed by whole-genome sequencing and comparative analysis. *Pigment Cell Melanoma Res*. 2014;**27**: 835-8.
34. Berger MF, Hodis E, Heffernan TP, et al. Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature*. 2012;**485**: 502–6.
35. Shearin AL, Hedan B, Cadieu E, et al. The MTAP-CDKN2A locus confers susceptibility to a naturally occurring canine cancer. *Cancer Epidemiol Biomarkers Prev*. 2012; **21**: 1019-27.
36. Kabbarah O, Chin L. Revealing the genomic heterogeneity of melanoma. *Cancer Cell*. 2005;**8**: 439-441.
37. Forslund A, Zeng Z, Qin LX, et al. MDM2 gene amplification is correlated to tumor progression but not to the presence of SNP309 or TP53 mutational status in primary colorectal cancers. *Mol Cancer Res*. 2008; **6**: 205-11.
38. Broeckx BJG, Coopman F, Verhoeven GE, et al. Development and performance of a targeted whole exome sequencing enrichment kit for the dog (*Canis Familiaris* Build 3.1) *Sci Rep*. 2014;**4**: 5597.
39. Bainbridge MN, Wang M, Wu Y, et al. Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. *Genome Biol*. 2011;**12**: R68.
40. Piskol R, Ramaswami G, Li J. B. Reliable identification of genomic variants from RNA-seq data. *Am J Hum Genet*. 2013;**93**: 641–51.
41. Hakamori S. Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. *Cancer Res*. 1996;**56**: 5309-18.

42. Peng J, Chen B, Shen Z, et al. DNA promoter hypermethylation contributes to down-regulation of galactocerebrosidase gene in lung and head and neck cancers. *Int J Clin Exp Pathol.* 2015;**8**: 11042-50.
43. Lu J, Zhang G, Cheng Y, et al. Reduced expression of SRY-box containing gene 17 correlates with an unfavorable melanoma patient survival. *Oncol Rep.* 2014;**32**: 2571-9.
44. Devi S, Kedlaya R, Maddodi N, et al. Calcium homeostasis in human melanocytes: role of transient receptor potential melastatin 1 (TRPM1) and its regulation by ultraviolet light. *Am J Physiol Cell Physiol.* 2009;**297**: C679-87.
45. Miller AJ, Du J, Rowan S, Hershey CL, Widlund HR, Fisher DE. Transcriptional regulation of the melanoma prognostic marker melastatin (TRPM1) by MITF in melanocytes and melanoma. *Cancer Res.* 2004;**64**: 509-16.
46. Duncan LM, Deeds J, Hunter J, et al. Down-regulation of the novel gene melastatin correlates with potential for melanoma metastasis. *Cancer Res.* 1998;**58**: 1515-20.
47. Brozyna AA; Guo H, Yang SE, et al. TRPM1 (melastatin) expression is an independent predictor of overall survival in clinical AJCC stage I and II melanoma patients. *J Cutan Pathol.* 2017;**44**: 328-337.
48. Byun HO, Han NK, Lee HJ, et al. Cathepsin D and eukaryotic translation elongation factor 1 as promising markers of cellular senescence. *Cancer Res.* 2009;**69**: 4638-47.
49. Kim BC, Yoo HJ, Lee HC, et al. Evaluation of premature senescence and senescence biomarkers in carcinoma cells and xenograft mice exposed to single or fractionated irradiation. *Oncol Rep.* 2014;**31**: 2229-35.
50. Passarelli A, Mannavola F, Stucci LS, Tucci M, Silvestris F. 2017. Immune system and melanoma biology: a balance between immunosurveillance and immune escape. *Oncotarget.* 2017;**8**: 106132-106142.
51. Shankaran V, Ikeda H, Bruce AT, et al. IFN $\gamma$  and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature.* 2011;**410**: 1107-11.
52. Langman RE, Cohn M. Self-nonsel self discrimination revisited. *Semin Immun.* 2000;**12**: 159–62.
53. Zhang F, Du G. Dysregulated lipid metabolism in cancer. *World J Biol Chem.* 2012;**3**: 167-74.
54. Tennant DA, Durán RV and Gottlieb E. Targeting metabolic transformation for cancer therapy. *Nat Rev Cancer.* 2010;**10**: 267-77.
55. Maciel E, Felgueiras J, Silva EM, et al. Lipid remodelling in human melanoma cells in response to UVA exposure. *Photochem Photobiol Sci.* 2017;**16**: 744-752.



56. Hayward NK, Wilmott JS, Wadell N, et al. Whole-genome landscapes of major melanoma subtypes. *Nature*. 2017; **545**:175-180.
57. Hernandez B, Adissu HA, Wei BR, Michael HT, Merlino G, Simpson RM. Naturally Occurring Canine Melanoma as a Predictive Comparative Oncology Model for Human Mucosal and Other Triple Wild-Type Melanomas. *Int J Mol Sci*. 2018; **19**. pii: E394.

## FIGURE LEGENDS

**Figure 1. Circos plot depicting the Copy Number Aberrations (CNAs) in the eight tumor samples.** Chromosome numbers are identified in the outer-most ring. From outer to inner circles, second and third circles represent CNAs in disease progression samples (T1), as heat map and scatter diagram, respectively. Fourth and fifth circles represent CNAs in naive samples (T0), as heat map and scatter diagram, respectively. Blue dots correspond to amplifications ( $>0.25$ ) while red dots correspond to deletions ( $<-0.25$ ). The dimension of the dots increase with the size of the amplification/deletion.

**Figure 2. Venn's diagrams of somatic variants across samples.** Illustrative representation of single nucleotide variants (SNVs) and indels at T0 (**A**) and at T1 (**B**) melanomas and LOH calls across the T0 (**C**) and T1 (**D**) samples.

**Figure 3. Illustration of ratio frequencies of transitions and transversions at presentation and disease progression.** Naïve and progressed melanomas showed similar pattern of mutations, with a significant prevalence of C>T/G>A transitions.

**Figure 4. Coding somatic mutations (SNVs and indels) distribution and consequences across the eight tumor samples.** Genomic distribution of somatic mutations (**A**). Number and consequences of coding somatic mutations (**B**). Number and consequences of coding somatic SNVs (**C**).

## SUPPORTING INFORMATION DESCRIPTION

**Supplementary Materials and Methods.** Detailed description of DNA/RNA isolation and WES/RNA-seq library preparation and sequencing.

**Table S1.** Clinical data of dogs included in the study.

**Table S2.** RNA-seq sequencing yields, quality scores, and reads mapping summary.

**Table S3. Differentially expressed genes (DEGs) at disease progression- compared with naïve-canine melanomas (T1 vs T0), in both the biological and technical data sets.** Table S3 contains three worksheets organized as follow: S3.1. T1 vs T0 DEGs in biological replicates; S3.2. T1 vs T0 DEGs in technical replicates; S3.3. List of overlapping DEGs between biological and technical replicates. Column A: Dog ENSEMBL gene ID. Column B: mean count. Columns C, D: log<sub>2</sub> fold change (FC), and standard error. Columns E, F: nominal and adjusted (Benjamini-Hochberg) p-value. Column G: dog gene symbol. Column H: gene type. Column I: gene description.

**Table S4. Gene Ontology (GO) terms, KEGG pathways and Cancer Hallmarks significantly enriched identified by Gene Set Enrichment Analysis in disease progression tumor samples (T1).** Column A: enrichment direction (POS=positively enriched; NEG= negatively enriched). Column B: GO/KEGG/Cancer Hallmarks terms ID. Columns C, D, E: gene set size, enrichment score (ES) and normalized enrichment score (NES). Columns F, G, H: nominal p-value, False Discovery Rate (FDR) q-value and familywise-error rate (FWER) p-value.

**Table S5.** Recurrent Copy Number Aberrations identified in the eight samples by GISTIC algorithm using NEXUS software.

**Table S6.** Most frequent copy number gains and losses in all eight samples.

**Table S7.** Distribution of SNVs, Indels and LOH across all the tumors.

**Table S8.** Ratio of transitions and transversions at presentation and disease progression.

**Table S9. List of common somatic mutations (SNVs and indels) across the T0 and T1 melanomas.** Table S9 contains two worksheets organized as follow: S9.1. List of common somatic mutations in at least 2 dogs across the T0 samples; S9.2. List of common somatic mutations in at least 2 dogs across the T1 tumors. Columns A, B: Dog ID samples presenting the mutations and number of mutations per group of samples. Columns C, D: variant location on CanFam genome. Columns E, F: reference allele and variant allele. Column G: Dog ENSEMBL gene ID, Columns H, I, J: ANNOVAR variant annotation, gene symbol and gene type, Columns K, L: Variant Effect Predictor (VEP) impact and Amino Acids substitutions, Columns M: dbSNP variant ID, Column N: SNVs Sorting Intolerant from Tolerant (SIFT) effect prediction.

**Table S10.** Distribution of somatic mutations (SNVs and indels): intra-individual comparison.

**Table S11.** List of overrepresented pathways and GO terms identified by ClueGO using the list of mutated genes (T1 vs T0).

**Figure S1. Gene Set Enrichment Analysis (GSEA) enrichment plot (score curves) of the most ‘negatively-enriched’ KEGG pathways in T1, compared with T0.** Gene Set Enrichment Analysis was performed with the canonical pathway gene sets in GSEA Molecular Signatures Database (MSigDB). The green curve corresponds to the enrichment score curve, which is the running sum of the weighted enrichment score obtained from GSEA software. a, b, c, and d denote the most negatively enriched (negative NES) pathways (i.e., gene sets): Retinol metabolism, sphingolipid metabolism, arachidonic acid metabolism, and steroid biosynthesis pathway, respectively. NES = Normalized Enrichment Score; FDR = False Discovery Rate (q value).

**Figure S2. Gene Set Enrichment Analysis (GSEA) enrichment plot (score curves) of the most ‘positively-enriched’ KEGG pathways in T1, compared with T0.** Gene Set Enrichment Analysis was performed with the canonical pathway gene sets in GSEA Molecular Signatures Database (MSigDB). The green curve corresponds to the enrichment score curve, which is the running sum of the weighted enrichment score obtained from GSEA software. a-h denote the most positively enriched (positive NES) pathways (i.e., gene sets): (a) natural killer cell-mediated cytotoxicity, (b) primary immunodeficiency, (c) complement and coagulation cascades, (d) prion diseases, (e) B cell receptor signaling pathway, (f) antigen processing and presentation, (g) chemokine signaling pathway and (h) T cell receptor signaling pathway, respectively. NES = Normalized Enrichment Score; FDR = False Discovery Rate (q value).

**Figure S3. Gene Set Enrichment Analysis (GSEA) enrichment plot (score curves) of the top 5 ‘negatively-enriched’ biological processes (GO terms) in T1, compared with T0.** Gene Set Enrichment Analysis was performed with the GO term gene sets in GSEA Molecular Signatures Database (MSigDB). The green curve corresponds to the enrichment score curve, which is the running sum of the weighted enrichment score obtained from GSEA software. a, b, c, d and e denote the most negatively enriched (negative NES) biological processes: keratinocyte differentiation, skin development, epidermal cell differentiation, keratinization, and epidermis development, respectively. NES = Normalized Enrichment Score; FDR = False Discovery Rate (q value).

**Figure S4. Functional map of the mutated genes (T1 vs T0) using ClueGO.** In the network, the Gene Ontology (GO) biological terms and pathways were identified as nodes and linked based on their *kappa* score level ( $>0.4$ ) and False Discovery Rate (FDR)  $< 0.05$ . The size node represents the term enrichment significance. Functionally related groups partially overlapped. The similar GO terms/KEGG pathways are labelled in the same color. The list of mutated genes in common in at least 2 samples was used (n=171).