# Native nanopore sequencing of multiple tumor sites reveals genetic and epigenetic intra-tumor heterogeneity in canine osteosarcoma

Ayse Keskus[1], Sergey Aganezov[2], Tanveer Ahmad[1], Xiaoguang Dai[2], Robert Walker[1], John Beaulaurier[2], Scott Hickey[2], Sissel Juul[2], Amy K. LeBlanc[1], Paul S. Meltzer[1], Mikhail Kolmogorov[1]

[1] NIH, National Cancer Institue, Bethesda, MD.
[2] Oxford Nanopore Technologies Inc, New York City, NY.

Contact: sergey.aganezov@nanoporetech.com, mikhail.kolmogorov@nih.gov

## Background

Osteosarcoma (OS) is an aggressive bone malignancy affecting children and young adults. Neoadjuvant chemotherapy has led to significant improvements in outcome, but patients with recurrence or metastases at diagnosis continue to have poor prognoses.

Whole genome sequencing (WGS) has revealed that a hallmark of OS is a highly rearranged genome with complex copy number aberrations rather than point mutations. A better understanding of the key genomic and epigenetic factors driving clonal evolution may provide insights into tumor progression and therapy response[1].

Canine OS recapitulates most of the hallmark genetic and molecular factors of human OS and is recognized as an informative model. A significant difference is that initial surgery in canine OS allows us to explore tumor evolution unperturbed by treatment, providing an invaluable resource for studying the genetic and molecular landscape of tumor evolution.

Here, we present a study in which we performed bulk WGS of multiple tumor sites and the matching normal genome in two dogs using nanopore long-read sequencing to analyze the genetic and epigenetic evolution of OS (Figure 1).

Briefly, we obtained 4 and 6 tumor samples from each dog, where each tumor sample derived from a distinct site within the primary tumor growth. We performed DNA extraction with short fragment elimination and shearing by Megaruptor, followed by Q20+ library preparation (SQK-LSK114) and native sequencing on PromethION with R10.4.1 flow cells. Reads were basecalled using `dorado` with the SUP model and with both 5mC and 5hmC methylation calling enabled. We obtained 45-60x average read depth for tumor samples and >30x for matching normals with read N50 values ranging between 20-26Kbp.

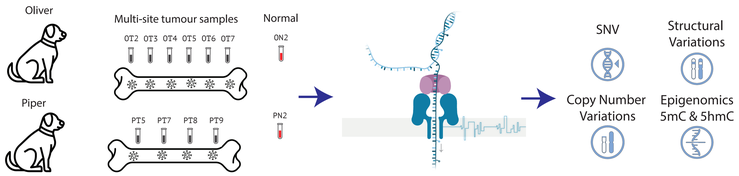


**Figure 1**. Multi-site primary tumor and matching normal native nanopore WGS bulk sequencing in 2 dogs followed by genome-wide genomic and epigenomic haplotype-resolved analysis.

## Germline analysis

We aligned nanopore reads using `minimap2` to the latest reference canine assembly `canFam6`. We then called germline SNPs in the normal samples with `Clair3`. We observed comparable summary counts of both homozygous and heterozygous variants in both dogs with an even distribution across the chromosomes (Figure 2a). We then used `HAPCUT2` to phase heterozygous germline SNPs using long nanopore reads from the normal samples. We observed on average ~50% of chromosomal lengths being phased in each dog with the rest being occupied by long runs of homozygosity. Roughly half of these regions were shared between both dogs, while the other half were specific to a single dog (Figure 2b,c). Following phasing, we then haplotagged all reads from both the normal and tumor samples for each dog using `whatshap`.

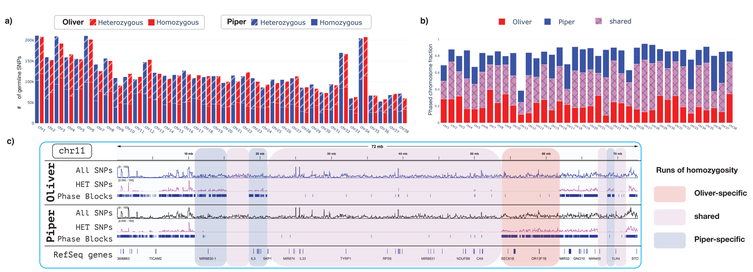


**Figure 2**. Germline SNPs and phasing characterization. a) Homozygous and heterozygous SNPs counts w/ per-chromosome distribution. b) Fractions of chromosomes phased with heterozygous SNPs in each dog. c) chr11 example with shared and unique long runs of homozygosity in each dog.

## Somatic structural variations

In both subjects, somatic structural variation detection using `Severus`[2] revealed chromothripsis-like rearrangements in chr11, leading to the deletion of CDKN2A, a known cell cycle regulator and tumor suppressor (Figure 3a,d). Interestingly, in Oliver, most of the SVs (n = 131) shared across all the samples were localized to chr5 (Figure 3c), coinciding with perturbed copy number profiles and bi-allelic frequency of the SNVs. In Piper, the majority of SVs shared across all the samples derived from the chr11 chromothripsis event (Figure 3f). The multi-site sample analysis revealed somatic SVs shared across only a subset of analyzed samples, identifying major subclonal clusters (e.g., OT2 w/ OT6 for Oliver and PT7 w/ PT9 for Piper) (Figure 3b,e). This indicates divergence in somatic SV accumulation across tumor locations.

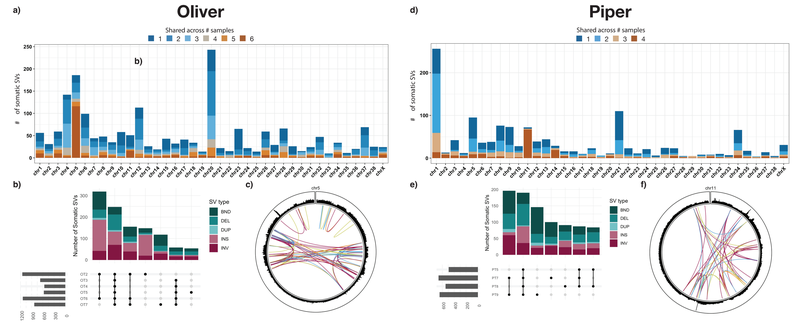


**Figure 3**. Somatic structural variations. Per-chromosome load of most to least sample-shared somatic SVs for Oliver (a) and Piper (d). Upset plots of sample-specific somatic SVs and their cross-sample presence in Oliver (b) and Piper (e). Chromothripsis-like somatic SV diagram on chr5 in Oliver (c) and chr11 in Piper (f).

## Somatic single nucleotide variants

We used `ClairS` to detect somatic SNVs in each individual tumor sample, then merged and genotyped somatic SNVs across all tumor samples in each dog using `bcftools` and pileup variant evidence.

We observed a large number of sample-private somatic SNVs evenly distributed across all chromosomes (Figure 4a,d), but with an uneven total SNVs count distribution across samples (higher counts in OT2, OT6 in Oliver and PT7 in Piper) (Figure 4b,e). We observed a similar grouping of samples (OT2 w/ OT6 in Oliver and PT7 w/ PT9 in Piper) in the case of somatic SVs. Somatic SNVs allele frequency distribution across tumor samples was consistent for each dog, with a notable exception of OT6 in Oliver (Figure 4c,f). These results point to a presence of subclonal somatic SNVs accumulation, while aligning with the cross-sample somatic SV results.

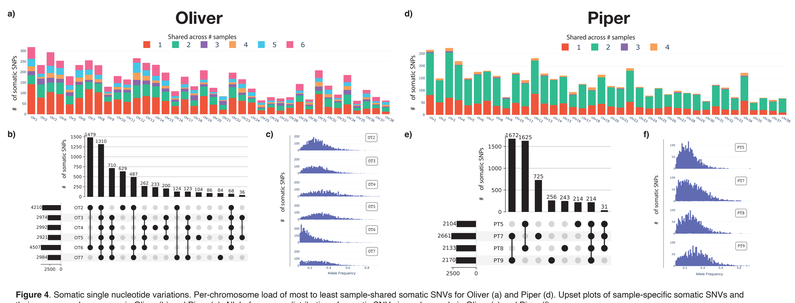


**Figure 4**. Somatic single nucleotide variations. Per-chromosome load of most to least sample-shared somatic SNVs for Oliver (a) and Piper (d). Upset plots of sample-specific somatic SNVs and their cross-sample presence in Oliver (b) and Piper (e). Allele-frequency distribution of somatic SNVs in each sample in Oliver (c) and Piper (f).

## Somatic copy number variations

To explore somatic CNVs we used `HATCHet`[3] with global clustering and joint analysis of multiple tumor samples. Briefly, for `HATCHet` input we fixed binned the referenced genome, computed read-depth-ratio (RDR) between tumor and normal samples and cross-sample-fixed B-allele frequencies. `HATCHet` inferred multi-clonal haplotype-specific CNV profiles, revealing the locations of copy-number heterogeneity (e.g., chr6 in Oliver and chr5 in Piper) and cases of mirrored haplotype CNVs (e.g., chr6, chr12, chr17 in Oliver and chr10, chr12, chr17 in Piper) (Figure 5a,b). The clonal composition in samples was consistent with the sample groupings observed in both the somatic SVs and SNVs (Figure 5c,d).

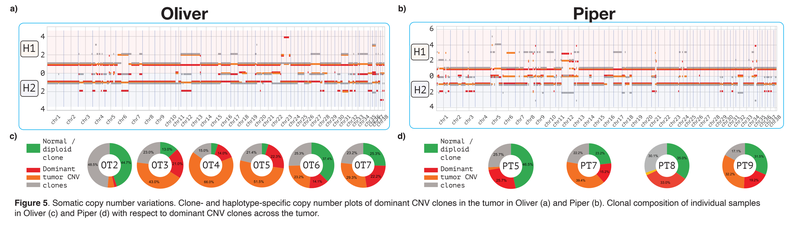


**Figure 5**. Somatic copy number variations. Clone- and haplotype-specific copy number plots of dominant CNV clones in the tumor in Oliver (a) and Piper (b). Clonal composition of individual samples in Oliver (c) and Piper (d) with respect to dominant CNV clones across the tumor.

## Somatic methylome landscape

We utilized `modkit` to generate genome-wide bedmethyl tables for both 5mC and 5hmC methylation modifications across all CpG loci. Tumor samples exhibited a reduced level of 5mC methylation compared to the normal genome, while 5hmC levels were either elevated or reduced, depending on the specific tumor sample (Figure 6a). We were able to assess the haplotype-specific, base-pair resolved methylation patterns across tumors samples within phased regions, allowing us to delineate haplotype-resolved instances of both localized and extended differentially methylated regions (Figure 6b).

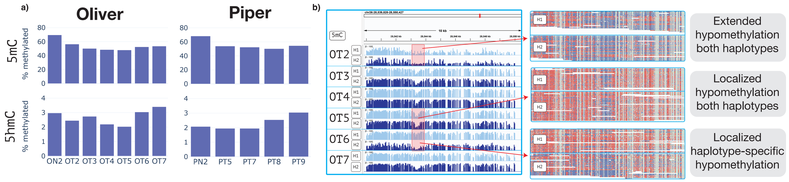


**Figure 6**. 5mC and 5hmC prevalence. a) Sample-specific 5mC and 5hmC methylation frequencies across CpG locations. b) Example of haplotype-resolved sample-unique 5mC DMR region in Oliver.

## Conclusion

A better understanding of the genomic and epigenetic factors driving somatic evolution of OS may provide insights into tumor progression and therapy response. Here we showcase how a single nanopore sequencing assay enables a simultaneous analysis of genetic and epigenetic tumor alterations with joint analysis of multiple tumor sites, shedding light on the clonal composition of OS. Somatic alterations of all types and sizes can be inferred in a haplotype-resolved fashion, providing a richer understanding of the drivers of somatic evolution and their association with the molecular landscape in OS.

Download poster

## References

[1] Meltzer, Paul S., and Lee J. Helman. "New horizons in the treatment of osteosarcoma." New England Journal of Medicine 385.22 (2021): 2066-2076

[2] Keskus, Ayse, et al. "Severus: accurate detection and characterization of somatic structural variation in tumor genomes using long reads." medRxiv (2024): 2024-03.

[3] Zaccaria, Simone, and Benjamin J. Raphael. "Accurate quantification of copy-number aberrations and whole-genome duplications in multi-sample tumor sequencing data." Nature communications 11.1 (2020): 4301.