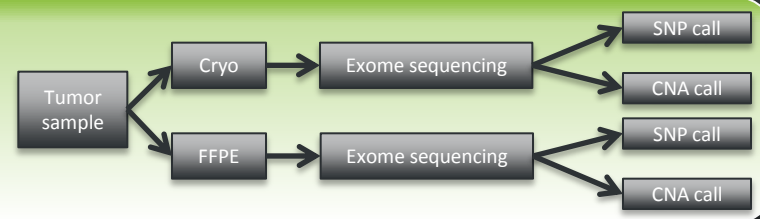
Exome sequencing of tumors: relevance in copy-number alteration (CNA) analysis and fixed tissue samples.

Stéphane Wenric [1] and Claire Josse [2]; Corinne Fasquelle [1]; Christophe Poulet [1]; Tiberio Sticca [1]; Meriem Boukerroucha [1]; Guy Jerusalem [2]; Vincent Bours [1]

[1] - University of Liege, GIGA-Research, Human Genetics Unit

[2] - University of Liege Hospital (ULg CHU), Medical Oncology Laboratory

Genomic DNA has been extracted from both **cryopreserved** and **formalin-fixed paraffin-embedded** forms of 2 different **tumor samples** (triple negative, and Her2+).
Exome sequencing has been performed on all 4 forms, as well as **SNP** and **CNA** detection. A comparison of the various metrics and results related to the sequencing, mapping, and variants detection has been done, outlining what can, and can't be done with exome data sequenced from cryopreserved and FFPE tissue.

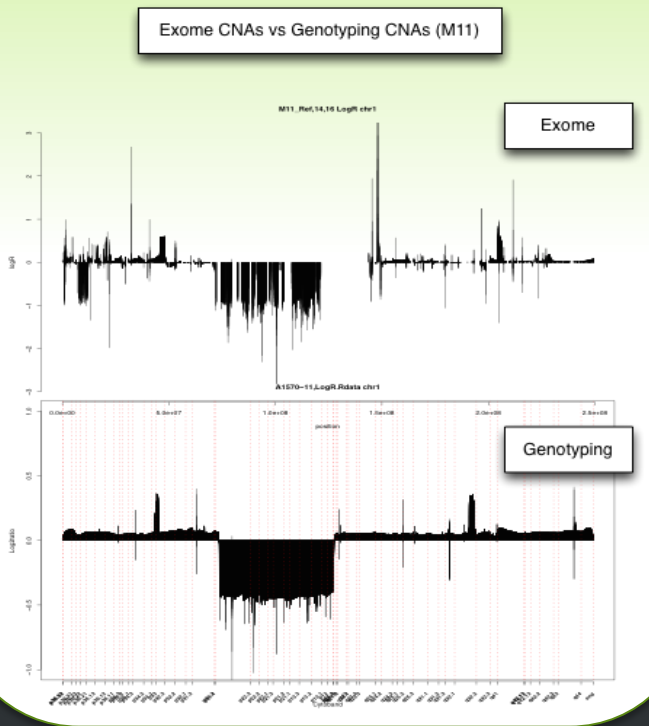


SAMPLE	TN Cryo	Her2+ Cryo	TN FFPE	Her2+ FFPE
GENOME_SIZE	3 101 804 741	3 101 804 741	3 101 804 741	3 101 804 741
TARGET_TERRITORY	62 286 318	62 286 318	62 286 318	62 286 318
TOTAL_READS	76 963 934	70 975 202	37 298 532	22 505 872
PF_UNIQUE_READS	72 257 335	66 407 261	34 041 479	21 028 247
PCT_PF_UQ_READS	93.88%	93.56%	91.27%	93.43%
MEAN_TARGET_COVERAGE	49.30	45.12	25.31	14.20
PCT_PF_UQ_READS_ALIGNED	87.97%	88.38%	90.13%	87.16%
ZERO_CVG_TARGETS_PCT	0.57%	0.73%	2.31%	3.53%

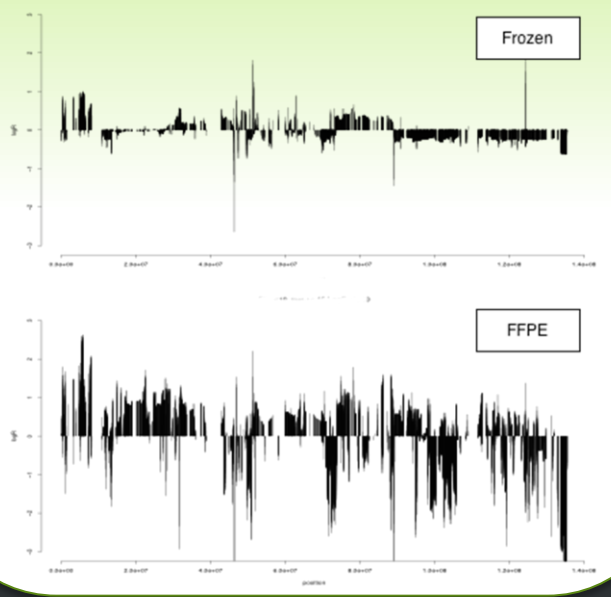
All 4 samples were pooled on the same flow-cell lane and were sequenced on an Illumina HiSeq 2000.
The data shows strong different capture affinity (and thus a difference in the number of reads) between cryo and FFPE samples, but roughly the **same proportion of aligned reads**.

The mean target on coverage drops in the FFPE samples, as is expected, considering the number of sequenced reads for these libraries.

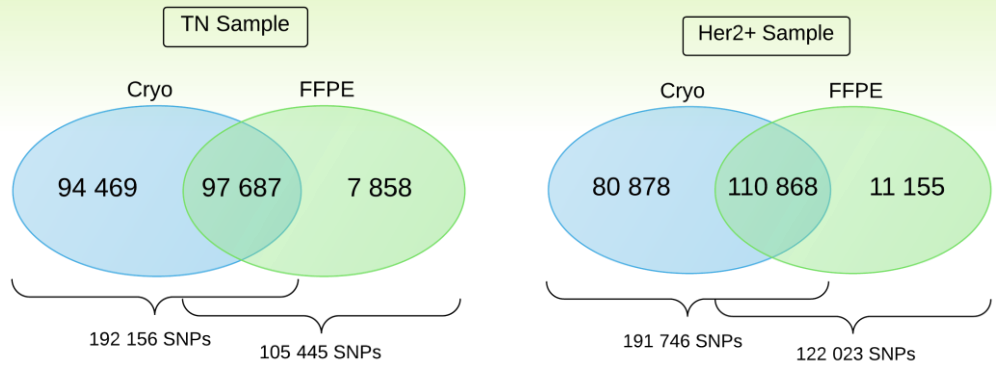
CNA calling with the cryopreserved exome data yields **good results**, which were compared with CGH data for validation.
All CNA regions which were previously detected with CGH analysis were correctly detected with the exome data.



CNA calling with the FFPE sample exome data yields **noisy results** (which may be caused by the use of frozen sample data as reference), but still allows for the detection of the Her2+ amplification.

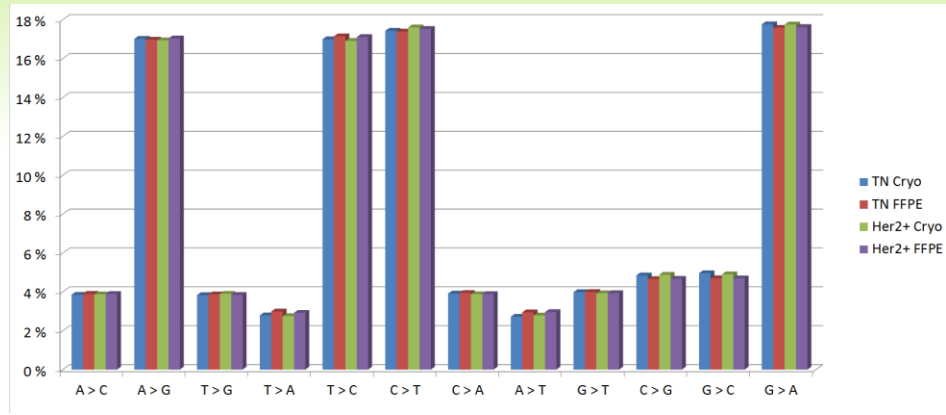


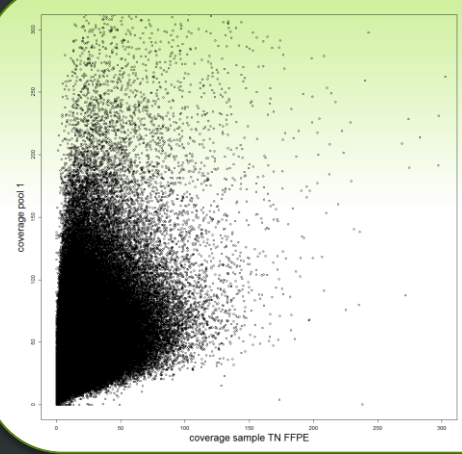
SNP calling was performed on all samples, **more than 90% of the SNPs found in FFPE exomes were also found in the corresponding cryo exomes**.
A lower target coverage in FFPE yields a lower number of detected SNPs, for both samples, but **FFPE-related degradations do not seem to cause false positives**, in either samples.



BRCA2 mutations were correctly detected in the FFPE exome data.

All samples show the **same substitution rates**, thus the FFPE-related degradation is either completely random or its non-randomness is corrected by the mapping algorithms.





Future work: we've conjectured that the poor performance in CNA detection with FFPE data was caused by a low read depth correlation between the FFPE samples and the references used.

We plan on testing the use of a pool of several FFPE samples as future reference to assess as to whether this method of CNA detection is also suited for FFPE data.

Moreover, we plan on increasing the read coverage of future FFPE exome libraries to allow for better comparisons with cryo libraries.

Contact : c.josse@ulg.ac.be ; s.wenric@ulg.ac.be