Background: We provide a solution of pressing needs in preparation of FFPE samples for genomic analysis: small sample size, unwanted admixture of normal cells, analysis of tumor-rare subpopulations present at low percentages in the tumor fraction.

Methods: We disaggregated into cell suspensions archival FFPE samples from 1 ovarian, 2 pancreatic, 9 lung, 3 adenocarcinoma rectosigmoid, 7 adenocarcinome rectum, 1 adenocarcinergic sigmoid cancer patients. Cells were stained for Vimentin, Keratin and nuclear stain. We sorted by DEPAarray™ precise number (mean=140, median 94, range=5-600) of pure homogenous cells from the major population of tumor cells, the contaminant diploid stromal cells, and other minority tumor cell types indicative of epithelial-to-hemenchymal transition (EMT). Using IonTorrent AmpliSeq™ CHP2, we generated NGS sequencing libraries, after direct lysis of the sorted pure cells recovered by DEPAarray™ (n=94), or of the unsorted samples (either QIAmp DNA columns or disaggregated cells). Libraries were sequenced with IonTorrent™ PGM (mean depth=2,000x), and analyzed using Torrent Suite.

Results: On several loci, we detected somatic mutations with 100% variant frequency, only observable as heterozygous in the unsorted and as wild-type in stromal cells of same patient, confirming 100% purity of sorted cells. Moreover, in the EMT-phenotype subpopulations we identified clear somatic mutations, different from tumor cells majority and undetectable in unsorted samples. Frequently, for loci harboring germ-line heterozygous SNPs with variant frequency around 50% for pure stromal cells, we readily detected loss-of-heterozygosity in tumor cells subpopulations as binary (0%/100%) variants. Quantitative traits such as copy number gains and losses were also reproducibly identified in tumor cell replicates as deviations from the 50% variant frequency of germline SNPs of pure stromal cells. Furthermore, we observed an excellent coverage uniformity (mean = 96%) for recoveries (n=56) in the range of 81-600 cells, even higher than the uniformly obtained with (n=3) QIAmp-purified DNA (90%). Mean uniformly gradually decreased to 87% for cell recoveries (n=20) in the range 21-80, and further decreased to 71% for lower cell numbers (n=16). Moreover, decreasing the cell number (under 60 cells), we observed an higher number of false positive variants, also with high frequency values, that make it difficult to filter them out. These artifacts are probably due to formalin-dependent DNA damage boosted by the low amount of DNA material.

Conclusion: We solved two pressing problems in preparation of FFPE samples for genomic analysis: small sample size and unwanted admixture of normal cells. Regardless of the tumor fraction in input, the procedure leads to robust and reproducible results because i)stromal contaminants are eliminated, ii) the number of input cells is controlled precisely by the instrument. The approach makes possible to reliably analyze also quantitative traits such as CNVs and LoH, which cannot be evaluated precisely in an unsorted sample.

Highlights: Sorting tumor rare-cell subpopulations reveals their genetic characteristics, undetectable in unsorted samples. Analyzing homogenous cell subpopulations as boosts of detection-to-noise ratio working around inherent sensitivity/specificity trade-offs of rare-variant calls. The proposed method enables reliable detection of quantitative traits such as CNVs. Sorting pure stromal cells yields internal controls for archival samples.