BIOF 520: Tracking Clonal Dynamics in Cancer

Diana Lin[1,2]

[1]Canada's Michael Smith Genome Sciences Centre, BC Cancer, Vancouver, BC, Canada

[2]Bioinformatics Graduate Program, University of British Columbia, Vancouver, BC, Canada

To determine the genetic underpinnings of treatment resistance and identify new targets for novel therapies in treatment resistant tumours, a computational and experimental workflow has been designed and outlined below.

To profile the genomes (and the genomic variants) of these patients, tissue samples are collected from multiple regions of the tumour and normal tissue and then separately homogenized to get maximal representation of all the clones of each tissue[1]. The samples are clonally profiled using bulk DNA sequencing and PyClone[2]. Next, the samples are directly xenografted[3] into humanized NSG mice[1] using established protocols[4], yielding patient derived xenografts (PDXs)[5]. Afterwards, a validation step of clonal profiling is done to ensure that the PDXs are representative of the primary tumour samples (and normal samples) before serial passaging[4]. PDXs are derived from as many patients as funding budget permits and done in replicates, as very few PDXs will take and pass validation[5]. Patients selected for PDXs depend on the amount of tissue cells that can be sampled from the patient; those with large enough samples are chosen[6]. For each patient's PDXs, some are treated with drugs, while others remain untreated to serve as controls for monitoring drug resistance evolutionary patterns[4]. The matched patient tumour and normal PDXs allows for comparison of somatic variation[4].

To make use of the two cohorts of patients (treatment-sensitive/-resistant), genomic variants common to each cohort are analyzed to see if there are variants that confer resistance. Bulk tissue genome and RNA sequencing are used to characterize each patient's genomic variants[7] at each passage/time point. Resources permitting, single cell sequencing can also be done[8] and used as a reference for clonal profiling or bulk tissue cell type deconvolution[9]. Bulk tissue genome sequencing (WGS), is chosen as it is a relatively easy and cheap protocol when compared to single cell sequencing, which is substantially more expensive and requires more skill to do, with higher risk to reward ratio. Bulk RNA-seq (WTS) does not add a lot cost-wise, but can significantly enrich the analysis, as some tools are able to use gene expression profiles to deconvolute bulk tissue data[10].

In each cohort, genomic variants such as single nucleotide variants (SNVs) and copy number variations (CNVs) will be called and annotated to characterize the profiles of each patient and to characterize the cancer mutational heterogeneity and clonal dynamics[4]. Given the resultant clonal and variant profiles of each cohort, pathway analysis and variant prioritization will reveal mutations under positive selection and driver mutations with oncogenic potential as treatment targets, while drug prioritization and drug sensitivity correlations will determine optimal treatments for those target variants. All tools used in this workflow are presented in **red**, in **Figure 1**.

Figure 1. The computational and experimental workflow to track clonal dynamics in treatment-resistant cancers.

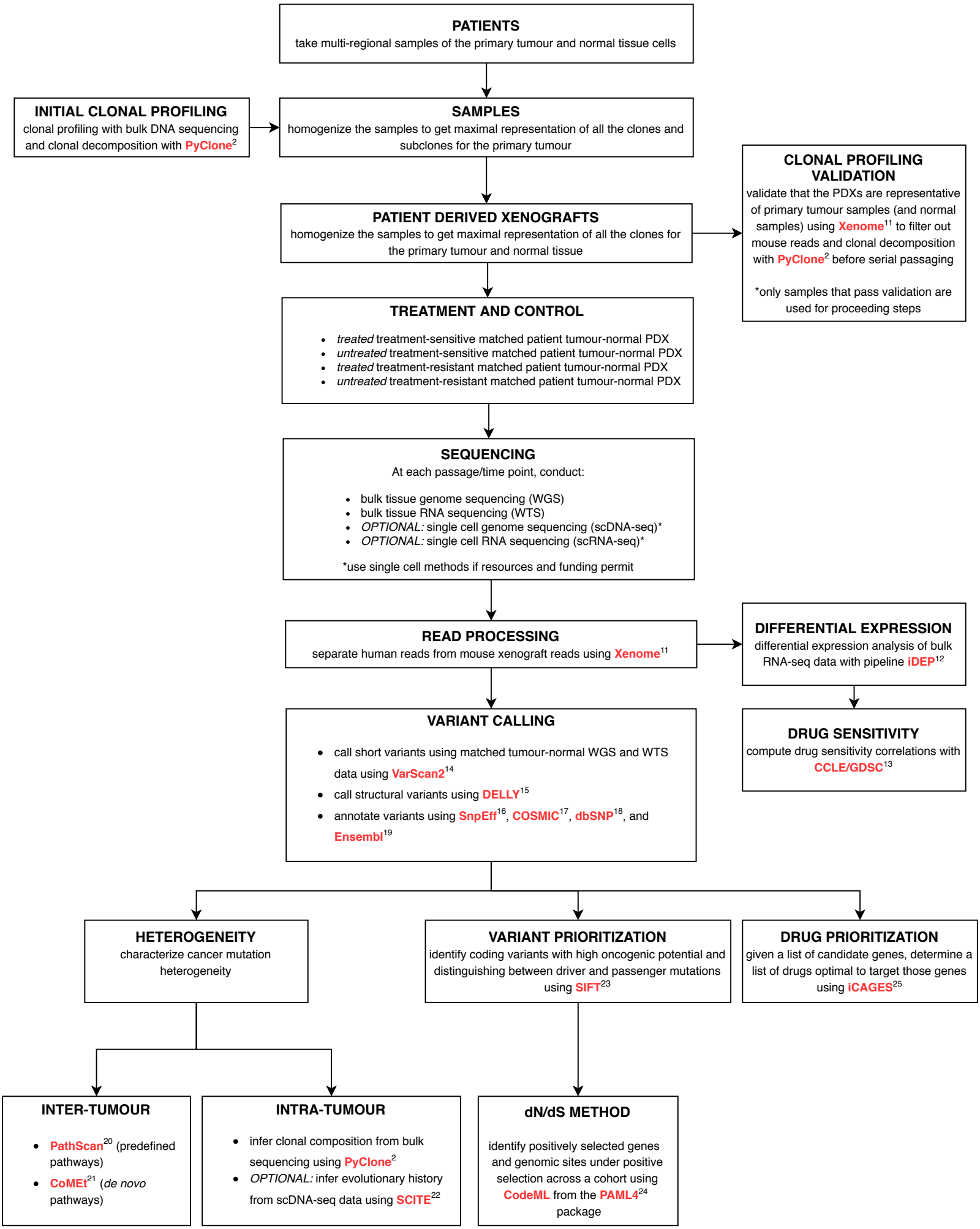
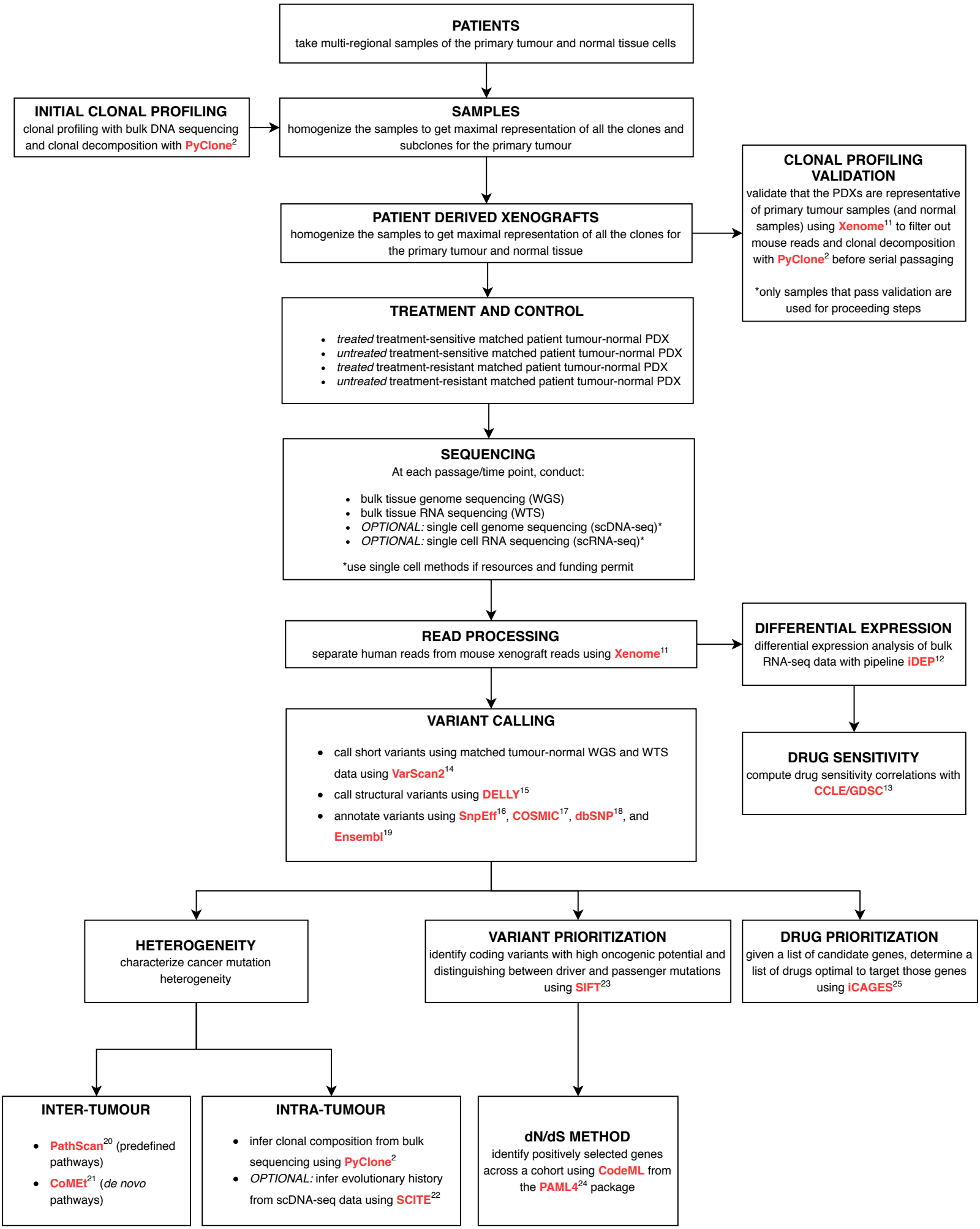


Figure 1. The computational and experimental workflow to track clonal dynamics in treatment-resistant cancers.



REFERENCES

1. Jung, J., Seol, H. S. & Chang, S. The Generation and Application of Patient-Derived Xenograft Model for Cancer Research. *Cancer Res. Treat.* **50**, 1–10 (2018).
2. Roth, A. *et al.* PyClone: statistical inference of clonal population structure in cancer. *Nat. Methods* **11**, 396–398 (2014).
3. Ng, W. Mouse xenograft models for elucidating drug resistance mechanisms. *PeerJ Preprints* (2017) doi:10.7287/peerj.preprints.1049v1.
4. Eirew, P. *et al.* Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature* **518**, 422–426 (2015).
5. Murayama, T. & Gotoh, N. Patient-Derived Xenograft Models of Breast Cancer and Their Application. *Cells* **8**, (2019).
6. Hidalgo, M. *et al.* Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer Discov.* **4**, 998–1013 (2014).
7. Woo, X. Y. *et al.* Genomic data analysis workflows for tumors from patient-derived xenografts (PDXs): challenges and guidelines. *BMC Med. Genomics* **12**, 92 (2019).
8. Kyrochristos, I. D., Ziogas, D. E., Goussia, A., Glantzounis, G. K. & Roukos, D. H. Bulk and Single-Cell Next-Generation Sequencing: Individualizing Treatment for Colorectal Cancer. *Cancers* **11**, (2019).
9. Sun, W., Jin, C., Gelfond, J. A., Chen, M.-H. & Ibrahim, J. G. Joint analysis of single-cell and bulk tissue sequencing data to infer intratumor heterogeneity. *Biometrics* (2019) doi:10.1111/biom.13198.
10. Wang, X., Park, J., Susztak, K., Zhang, N. R. & Li, M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nat. Commun.* **10**, 380 (2019).
11. Conway, T. *et al.* Xenome--a tool for classifying reads from xenograft samples. *Bioinformatics* **28**, i172–8 (2012).
12. Ge, S. X., Son, E. W. & Yao, R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **19**, 534 (2018).
13. Qin, Y., Conley, A. P., Grimm, E. A. & Roszik, J. A tool for discovering drug sensitivity and gene expression associations in cancer cells. *PLoS One* **12**, e0176763 (2017).
14. Koboldt, D. C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
15. Rausch, T. *et al.* DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012).
16. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012).
17. Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).
18. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research* **29**, 308–311 (2001).
19. Hunt, S. E. *et al.* Ensembl variation resources. *Database* **2018**, (2018).
20. Wendl, M. C. *et al.* PathScan: a tool for discerning mutational significance in groups of putative cancer genes. *Bioinformatics* **27**, 1595–1602 (2011).
21. Leiserson, M. D. M., Wu, H.-T., Vandin, F. & Raphael, B. J. CoMEt: a statistical approach to identify combinations of mutually exclusive alterations in cancer. *Genome Biology* vol. 16 (2015).
22. Jahn, K., Kuipers, J. & Beerenwinkel, N. Tree inference for single-cell data. *Genome Biology* vol. 17 (2016).
23. Ng, P. C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).
24. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
25. Dong, C. *et al.* iCAGES: integrated CAncer GEnome Score for comprehensively prioritizing driver genes in personal cancer genomes. *Genome Med.* **8**, 135 (2016).