medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Low-coverage genome sequencing for the detection of clinically relevant copy-number and mtDNA variants

Sander Pajusalu[1,2], Mikk Tooming[1,2], Kaisa Teele Oja[1,2], Ustina Šamarina[2], Tiina Kahre[1,2], Katrin Õunap[1,2]

[1] Department of Clinical Genetics, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia

[2] Genetics and Personalized Medicine Clinic, Tartu University Hospital, Tartu, Estonia

Grant numbers

MOBTP175, PRG471, PUT355, UM1 HG008900, R01 HG009141

Abstract and keywords

Background: Compared to exome sequencing, genome sequencing is widely appreciated for its superior ability to detect a wide range of genetic variations including copy-number variants (CNVs) and mitochondrial (mtDNA) variants. We assessed whether low-coverage genome sequencing, a considerably cheaper approach, would detect clinically relevant CNVs and mtDNA variants and would thus be a cost-efficient supplement to exome sequencing in rare disease diagnostics.

Methods: To assess the level of sequencing depth needed for variant detection, first, 30x mean coverage genome sequencing data were subsampled to 0.5x, 1x, 2x, and 4x coverage files *in silico* followed by CNV and mtDNA detection. Based on the analysis, 2x short-read sequencing was selected to be performed in 16 patients with putatively pathogenic CNVs or mtDNA variants to assess the empirical sensitivity.

Results: For CNV calling, 2x coverage was sufficient to detect all heterozygous CNVs greater than 10kb in size from *in silico* subsampled data. In experimental data, the results were similar, although a 16kb heterozygous deletion was once not detected. Regarding mtDNA variants, 2x coverage sufficed for variant confident variant calling and heteroplasmy assessment for all samples.

Conclusions: Low-coverage genome sequencing may be used to complement exome sequencing for simultaneous mtDNA variant and CNV detection.

Keywords:

Low-coverage genome sequencing, mtDNA sequencing, copy-number variants, genetic diagnosis

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Introduction

Since the first human exome sequencing study focused on finding a genetic cause for rare human diseases (Ng et al., 2009), novel sequencing technologies facilitating genome-wide simultaneous variant detections have revolutionized research and diagnostics of genetic disorders (Wright et al., 2018).

Currently, many diagnostic labs are performing exome sequencing and chromosomal microarray (CMA) in parallel to discover most of the clinically interpretable findings, as exome sequencing can detect single nucleotide variants and short deletions and insertions, while CMA is developed for detecting copy-number variants (CNVs), which are causative factors for ~10% of rare genetic disorders (Žilina et al., 2014). In case of suspicion of mitochondrial disorders, also mtDNA sequencing is requested.

High-depth genome sequencing can outperform CMAs in sensitivity for CNV detection, as it is not limited to the size resolution and can effectively detect both copy-number variants and balanced structural variants (e.g., translocations and inversions) (Collins et al., 2020). However, high computational and reagent costs challenge the usage in clinical settings. Generally, high-depth genome sequencing at standard 30x coverage is 3-5 times more expensive than exome sequencing, while CMA is cheaper than exome. Low coverage genome sequencing has been proposed as an alternative for CMA, as the lower coverage will reduce the costs proportionally while analytical sensitivity still outperforms CMA, especially for smaller deletions and duplications even at only 1x coverage (Dong et al., 2016; Zhou et al., 2018). Low-coverage genome sequencing has also been tested in prenatal settings (Wang et al., 2020). Several read-depth-based computational tools have been used for CNV detection from low-coverage genome sequencing data. Control-FREEC (Boeva et al., 2012) has shown the best performance with optimal computational resource usage (Smolander et al., 2021).

Another way to increase diagnostic yield of next-generation sequencing is to simultaneously detect mitochondrial DNA (mtDNA) variants, as exome and genome sequencing also cover mtDNA as a byproduct (Duan et al., 2018, 2019). An average PCR-free clinical genome sequencing has a mean read depth of around 30x, which results in above 2000x mtDNA coverage due to a large copy number of mtDNA in cells compared to autosomes (Laricchia et al., 2022). By reducing genome depth, sequencing costs decrease proportionally, but mtDNA coverage, although lower, could still be sufficient for variant detection. A previous study demonstrated that average autosome coverage of 1.6x resulted in average mtDNA coverage of 124x on genome sequencing (Rustagi et al., 2017). The sequencing depth of 100x or more is sufficient for detecting variants with heteroplasmy (proportion of mtDNA molecules having the non-reference allele) levels over 10% covering all clinically relevant variants if DNA from disease-relevant tissue is sequenced.

This study aims to assess whether low-coverage genome sequencing could be a reasonably cost-efficient solution for detecting CNVs and mtDNA variants and thus supplementing exome sequencing.

Methods

First, to assess the level of sequencing depth needed for variant detection, 30x mean coverage genome sequencing data from selected samples were subsampled to 0.5x, 1x, 2x, and 4x coverage files *in silico*. CNVs were detected using Control-FREEC (Boeva et al., 2012) and

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

annotated with AnnotSV (Geoffroy et al., 2018). The mtDNA variants were detected using the GATK4 mitochondrial pipeline (Laricchia et al., 2022) and annotated with HmtNote (Preste et al., 2019). The GATK4 mitochondrial pipeline also outputs theoretical sensitivity assessment for different heteroplasmy levels, which was used for selecting genome sequencing depth for the second part of the study. Five disease-causing deletions (sized 3.4kb, 4.6kb, 7.2kb, 16kb, and 90kb) and one possibly pathogenic heteroplasmic (heteroplasmy level 14.7%) mtDNA variant were used to assess sensitivity for clinically relevant variants.

Second, 16 samples from ten different families with known variants were selected for 2x genome sequencing. The chosen samples carried the following variants: eight (five unique) deletions, one duplication, three (one unique) inversions, and four (three unique) mtDNA variants (Tables 1 and 2). The sequencing run was carried out on Illumina NextSeq 500 in a single run using high output sequencing kit with 2x150bp paired-end reads. Fastq files were mapped to the hg38 reference genome using BWA MEM algorithm version 0.7.17-r1188 (Li & Durbin, 2009), duplicates were marked, and base quality scores recalibrated using GATK version 4.1.4.0 (van der Auwera et al., 2020). For quality control, genome sequencing metrics, including sequencing depth, were assessed with the Picards CollectWgsMetrics tool (http://broadinstitute.github.io/picard/). A read-depth assessment-based Control-FREEC software (Boeva et al., 2012) with 1kb and 10kb non-overlapping calling windows and Manta, combining paired and split-read evidence (Chen et al., 2016) was used to call structural variants from bam files. GATK4 mitochondrial pipeline (Laricchia et al., 2022) was used to detect mtDNA variants and assess heteroplasmy. Heteroplasmy was calculated as a ratio of alternate variant reads to the total sequencing depth for the same genome locus. Again, AnnotSV (Geoffroy et al., 2018) and HmtNote (Preste et al., 2019) were used to annotate structural and mtDNA variants, respectively.

The scripts used for the analysis are available at https://github.com/SanderEST/lcwgs.

This study was approved by the Research Ethics Committee of the University of Tartu (approval date 11/18/2018 and number 287M-15, and 19/10/2020 327T-3). Informed consent was obtained from patients or their legal guardians.

Results

First, in-silico subsampled data was assessed. For CNV calling, 2x coverage was sufficient to detect all heterozygous CNVs greater than 10kb. For smaller CNVs, even 4x coverage data did not suffice for CNV detection. An example of an estimated copy number using 10 kb windows on chromosome 13 around 16 kb deletion in both heterozygous and homozygous states is shown for different *in silico* subsampled depths in Figure 1. Regarding mtDNA variants, 2x coverage resulted in >99% theoretical sensitivity for heteroplasmy levels >10%. The possibly pathogenic heteroplasmic variant was detected with similar heteroplasmy levels in all depths (0.5x to 4x). We selected 2x as an aimed depth for the separate low-coverage genome sequencing experiment based on these results.

The sequencing depth for 16 samples selected for the experiment ranged from 1.62 to 2.12, following the aimed sequencing depth of 2x. Regarding the assessed variants, FREEC confidently detected 90kb heterozygous deletion and 16kb homozygous deletion (Table 1, Figure 1). The detection was inconsistent for heterozygous 16kb deletion, and the software

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

failed to detect smaller than 10kb deletions. Manta, using different algorithms, detected the variants with incomplete sensitivity which was not in direct concordance with the CNV size. For example, it was able to detect 7.2 kb deletion, and 14 kb duplication, which both were not detected by FREEC software, but was not able to detect even 90kb deletion and 16 kb homozygous deletion. We also assessed Manta's ability to detect large inversion on chromosome 9, which was detected in the homozygous state, and in one of the two heterozygous carriers.

The known putatively pathogenic mtDNA variants were all detected from 2x genome sequencing. Moreover, the heteroplasmy levels were concordant with the 30x genome data (Table 2).

Discussion

Although the field of rare disease diagnostics and research is shifting toward using high-depth genome sequencing as a first-tier test, the high cost for sequencing and computational demands make exome sequencing the most widely used test. Although possible, CNV detection from exome sequencing is challenged by the fragmented nature of the data (Pfundt et al., 2016). Thus, chromosomal microarrays are often used to supplement exome sequencing to detect clinically relevant CNVs. The detection limit of chromosomal microarrays depends on the array used, commonly ranging from 10kb to 100kb.

Also, mtDNA variants are often assessed separately. Although possible to detect from exome sequencing, the coverage is often poor and insufficient for heteroplasmy level assessment (Puusepp et al., 2018). Thus, a patient with a suspected genetic disorder but without a specific diagnostic hypothesis commonly receives three separate genetic tests, exome, mtDNA sequencing, and chromosomal microarray, making comprehensive testing expensive.

This study demonstrates that low-coverage genome sequencing can replace chromosomal microarray and mtDNA sequencing. However, some limitations have to be noted. For chromosomal microarrays, the resolution, i.e., the smallest size of a CNV that can be reliably detected, is provided by the manufacturer after sensitivity assessments. Similarly, low-coverage genome sequencing has its resolution, which may depend on the sequencing protocol and the bioinformatics pipeline. Each lab should assess the sensitivity and specificity of its protocol. Importantly, natural variation of sequencing depth should be considered as the depth for samples in the same run is never equal (Table 1).

Regarding mtDNA variant detection, the main limitation lies in the studied tissue. Exome sequencing is usually performed from the DNA extracted from the blood or saliva. In contrast, muscle or fibroblasts may be the preferred tissue for mtDNA variant detection due to differences in heteroplasmy levels between tissues. This should be noted, as it is tempting to use the already available DNA from exome sequencing for further studies. While the low heteroplasmy levels may not be detected from the low coverage genome sequencing, the sensitivity for heteroplasmy levels above 10% remains adequate. Thus this method is suitable for screening clinically relevant mtDNA variants.

As high-coverage genome sequencing is becoming cheaper, low-coverage genome sequencing may not be efficient in the future, where standard genome sequencing replaces exome sequencing, and other variant classes may be assessed from the same data. High coverage genome sequencing is probably more sensitive for other variant classes like repeat

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

expansion variant detection (Ibañez et al., 2022). However, for the next few years, low coverage genome sequencing can serve as a cost-effective complementing analysis for exome sequencing, widening the scope of variant detection.

Conclusions

Low-coverage genome sequencing may be used to complement exome sequencing for simultaneous mtDNA variant and structural variant detection. However, for smaller CNVs, higher coverage genome sequencing is needed for comprehensive variant detection.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Acknowledgments

This study was supported by European Regional Development Fund and the program Mobilitas Pluss grant MOBTP175. KÕ, TK, KTO, and MT received support from the Estonian Research Council grants PRG471 and PUT355. The high coverage genome sequencing used for methods development was provided by the Broad Institute of MIT and Harvard Center for Mendelian Genomics (Broad CMG) and was funded by the National Human Genome Research Institute, the National Eye Institute, and the National Heart, Lung and Blood Institute grant UM1 HG008900 and in part by National Human Genome Research Institute grant R01 HG009141.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Figures

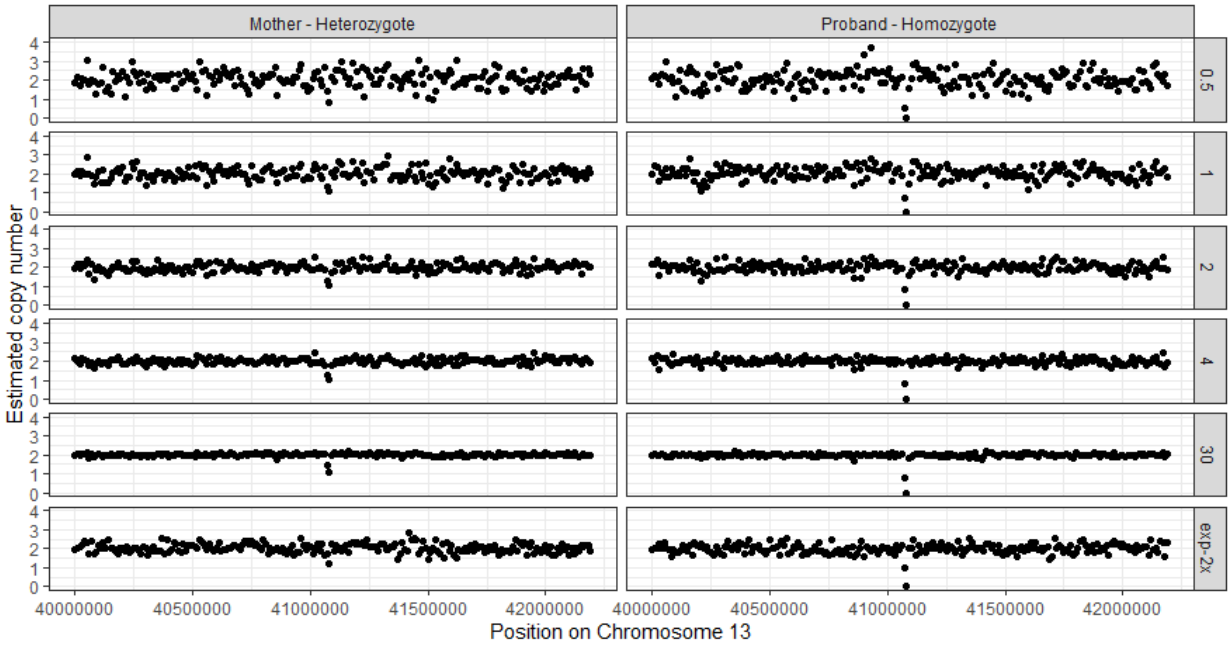

Figure 1. An example of an estimated copy number using 10 kb windows on chromosome 13 around 16 kb deletion in both heterozygous and homozygous states is shown for different *in silico* subsampled depths (0.5x, 1x, 2x, 4x) and the original 30x genome sequencing data as well as from the experimental 2x genome sequencing run (exp-2x).

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Tables

Table 1 – Cohort of patients with known structural variants selected for 2x genome sequencing (GS) and the results with 2 different variant callers. Control-FREEC software was used in two modes using either one kilobase or ten kilobase windows. Fam – family, DEL – deletion, DUP – duplication, INV – inversion.

Fam	**Relation**	**Known variant from 30x GS**	**Mean depth**	**Control-FREEC 10kb window**	**Control-FREEC 1kb window**	**Manta**
1	Proband	HOM 16kb DEL in *WBP4*	2.12	Detected	Detected	Not detected
1	Mother	HET 16kb DEL in *WBP4*	2.05	Not detected	Not detected	Not detected
1	Father	HET 16kb DEL in *WBP4*	1.82	Detected	Not detected	Not detected
2	Proband	HET 3.4kb DEL in *CTCF*	1.94	Not detected	Not detected	Not detected
3	Proband	HET 7.19kb DEL in *TRAPPC9*	1.86	Not detected	Not detected	Detected
3	Father	HET 7.19kb DEL in *TRAPPC9*	1.86	Not detected	Not detected	Not detected
4	Proband	HET 90.1kb DEL in *RBFOX1*	1.82	Detected	Detected	Not detected
5	Proband	HET 4.63kb DEL in *NIPBL*	1.98	Not detected	Not detected	Not detected
6	Proband	HET 14.2kb DUP in *SQOR*	1.95	Not detected	Detected	Detected
7	Proband	HOM 9Mb INV on chr 9	1.62	Not applicable	Not applicable	Detected
7	Mother	HET 9Mb INV on chr 9	1.78	Not applicable	Not applicable	Not detected
7	Father	HET 9Mb INV on chr 9	1.74	Not applicable	Not applicable	Detected

Table 2 - Cohort of patients with mitochondrial DNA variants selected for 2x genome sequencing (GS) and the results for the known variants.

Fam	Relation	Known mtDNA variant	Mean genome depth	Mean mtDNA depth	Heteroplasmy from 2x GS
8	Proband	m.9176T>C 95% heteroplasmy	2.16	166	93.8% (183/195)
9	Proband	m.3243A>G 9.5% heteroplasmy	1.79	202	7% (15/213)
10	Proband	m.15866A>G 14.7% heteroplasmy	2.07	299	17% (60/352)
10	Mother	m.15866A>G 5.6% heteroplasmy	2	225	4.5% (11/246)

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

References

Boeva, V., Popova, T., Bleakley, K., Chiche, P., Cappo, J., Schleiermacher, G., Janoueix-Lerosey, I., Delattre, O., & Barillot, E. (2012). Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics*, *28*(3), 423–425. https://doi.org/10.1093/BIOINFORMATICS/BTR670

Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Källberg, M., Cox, A. J., Kruglyak, S., & Saunders, C. T. (2016). Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*, *32*(8), 1220–1222. https://doi.org/10.1093/BIOINFORMATICS/BTV710

Collins, R. L., Brand, H., Karczewski, K. J., Zhao, X., Alföldi, J., Francioli, L. C., Khera, A. v., Lowther, C., Gauthier, L. D., Wang, H., Watts, N. A., Solomonson, M., O'Donnell-Luria, A., Baumann, A., Munshi, R., Walker, M., Whelan, C. W., Huang, Y., Brookings, T., … Talkowski, M. E. (2020). A structural variation reference for medical and population genetics. *Nature 2020 581:7809*, *581*(7809), 444–451. https://doi.org/10.1038/s41586-020-2287-8

Dong, Z., Zhang, J., Hu, P., Chen, H., Xu, J., Tian, Q., Meng, L., Ye, Y., Wang, J., Zhang, M., Li, Y., Wang, H., Yu, S., Chen, F., Xie, J., Jiang, H., Wang, W., Choy, K. W., & Xu, Z. (2016). Low-pass whole-genome sequencing in clinical cytogenetics: a validated approach. *Genetics in Medicine*, *18*(9), 940–948. https://doi.org/10.1038/gim.2015.199

Duan, M., Chen, L., Ge, Q., Lu, N., Li, J., Pan, X., Qiao, Y., Tu, J., & Lu, Z. (2019). Evaluating heteroplasmic variations of the mitochondrial genome from whole genome sequencing data. *Gene*, *699*, 145–154. https://doi.org/10.1016/j.gene.2019.03.016

Duan, M., Tu, J., & Lu, Z. (2018). Recent advances in detecting mitochondrial DNA heteroplasmic variations. In *Molecules* (Vol. 23, Issue 2). MDPI AG. https://doi.org/10.3390/molecules23020323

Geoffroy, V., Herenger, Y., Kress, A., Stoetzel, C., Piton, A., Dollfus, H., & Muller, J. (2018). AnnotSV: an integrated tool for structural variations annotation. *Bioinformatics*, *34*(20), 3572–3574. https://doi.org/10.1093/BIOINFORMATICS/BTY304

Ibañez, K., Polke, J., Hagelstrom, R. T., Dolzhenko, E., Pasko, D., Thomas, E. R. A., Daugherty, L. C., Kasperaviciute, D., Smith, K. R., Deans, Z. C., Hill, S., Fowler, T., Scott, R. H., Hardy, J., Chinnery, P. F., Houlden, H., Rendon, A., Caulfield, M. J., Eberle, M. A., … Zarowiecki, M. (2022). Whole genome sequencing for the diagnosis of neurological repeat expansion disorders in the UK: a retrospective diagnostic accuracy and prospective clinical validation study. *The Lancet Neurology*, *21*(3), 234–245. https://doi.org/10.1016/S1474-4422(21)00462-2/ATTACHMENT/B961BC1C-0911-45C2-B327-EA0431A8E1DB/MMC1.PDF

Laricchia, K. M., Lake, N. J., Watts, N. A., Shand, M., Haessly, A., Gauthier, L., Benjamin, D., Banks, E., Soto, J., Garimella, K., Emery, J., Rehm, H. L., MacArthur, D. G., Tiao, G., Lek, M., Mootha, V. K., & Calvo, S. E. (2022). Mitochondrial DNA variation across 56,434 individuals in gnomAD. *Genome Research*, *32*(3), 569–582. https://doi.org/10.1101/GR.276013.121/-/DC1

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E. E., Bamshad, M., Nickerson, D. A., & Shendure, J. (2009). Targeted

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280155; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

capture and massively parallel sequencing of 12 human exomes. *Nature*, *461*(7261), 272–276. https://doi.org/10.1038/nature08250

Pfundt, R., del Rosario, M., Vissers, L. E. L. M., Kwint, M. P., Janssen, I. M., de Leeuw, N., Yntema, H. G., Nelen, M. R., Lugtenberg, D., Kamsteeg, E.-J., Wieskamp, N., Stegmann, A. P. A., Stevens, S. J. C., Rodenburg, R. J. T., Simons, A., Mensenkamp, A. R., Rinne, T., Gilissen, C., Scheffer, H., … Hehir-Kwa, J. Y. (2016). Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. *Genetics in Medicine*, *August*, 1–9. https://doi.org/10.1038/gim.2016.163

Preste, R., Clima, R., & Attimonelli, M. (2019). Human mitochondrial variant annotation with HmtNote. *BioRxiv*, 600619. https://doi.org/10.1101/600619

Puusepp, S., Reinson, K., Pajusalu, S., Murumets, Ü., Õiglane-Shlik, E., Rein, R., Talvik, I., Rodenburg, R. J., & Õunap, K. (2018). Effectiveness of whole exome sequencing in unsolved patients with a clinical suspicion of a mitochondrial disorder in Estonia. *Molecular Genetics and Metabolism Reports*, *15*, 80–89. https://doi.org/10.1016/j.ymgmr.2018.03.004

Rustagi, N., Zhou, A., Watkins, W. S., Gedvilaite, E., Wang, S., Ramesh, N., Muzny, D., Gibbs, R. A., Jorde, L. B., Yu, F., & Xing, J. (2017). Extremely low-coverage whole genome sequencing in South Asians captures population genomics information. *BMC Genomics*, *18*(1), 396. https://doi.org/10.1186/s12864-017-3767-6

Smolander, J., Khan, S., Singaravelu, K., Kauko, L., Lund, R. J., Laiho, A., & Elo, L. L. (2021). Evaluation of tools for identifying large copy number variations from ultra-low-coverage whole-genome sequencing data. *BMC Genomics*, *22*(1), 1–15. https://doi.org/10.1186/S12864-021-07686-Z/TABLES/2

van der Auwera, G., O'Connor, B., & Safari, an O. M. Company. (2020). Using Docker, GATK, and WDL in Terra. *Genomics in the Cloud*, 300. https://www.oreilly.com/library/view/genomics-in-the/9781491975183/

Wang, H., Dong, Z., Zhang, R., Chau, M. H. K., Yang, Z., Tsang, K. Y. C., Wong, H. K., Gui, B., Meng, Z., Xiao, K., Zhu, X., Wang, Y., Chen, S., Leung, T. Y., Cheung, S. W., Kwok, Y. K., Morton, C. C., Zhu, Y., & Choy, K. W. (2020). Low-pass genome sequencing versus chromosomal microarray analysis: implementation in prenatal diagnosis. *Genetics in Medicine*, *22*(3), 500–510. https://doi.org/10.1038/S41436-019-0634-7

Wright, C. F., FitzPatrick, D. R., & Firth, H. v. (2018). Paediatric genomics: Diagnosing rare disease in children. In *Nature Reviews Genetics* (Vol. 19, Issue 5, pp. 253–268). Nature Publishing Group. https://doi.org/10.1038/nrg.2017.116

Zhou, B., Ho, S. S., Zhang, X., Pattni, R., Haraksingh, R. R., & Urban, A. E. (2018). Whole-genome sequencing analysis of CNV using low-coverage and paired-end strategies is efficient and outperforms array-based CNV analysis. *Journal of Medical Genetics*, *55*(11), 735–743. https://doi.org/10.1136/jmedgenet-2018-105272

Žilina, O., Teek, R., Tammur, P., Kuuse, K., Yakoreva, M., Vaidla, E., Mölter-Väär, T., Reimand, T., Kurg, A., & Õunap, K. (2014). Chromosomal microarray analysis as a first-tier clinical diagnostic test: Estonian experience. *Molecular Genetics & Genomic Medicine*, *2*(2), 166–175. https://doi.org/10.1002/mgg3.57