

CHROMOSOME-SCALE ASSEMBLY OF KAZAKH INDIVIDUALS USING MODERN GENOMICS APPROACHES

Aiziya Paizulla¹, Kuanysh Sarkytbayev², Anel Ordabayeva², Asset Daniyarov², Askhat Molkenov², Diana Samatkyzy³, Saule Rakhimova³, Ainur Akilzhanova³ and Ulykbek Kairov^{1,2}

¹*School of Sciences and Humanities, Nazarbayev University, Kazakhstan.* ²*Laboratory of Bioinformatics and Systems Biology, Center for Life Sciences, National Laboratory Astana, Nazarbayev University, Kazakhstan.*

³*Laboratory of Genomic and Personalized Medicine, Center for Life Sciences, National Laboratory Astana, Nazarbayev University, Kazakhstan.*

*Corresponding author (s): ulykbek.kairov@nu.edu.kz

Background: Population-specific reference genomes are crucial for accurate genomic analysis and precision medicine applications. Currently, there is a lack of high-quality chromosome-level reference genomes specifically representing the Kazakh population, which limits the effectiveness of biomedical investigations and genomic studies in this demographic. This project aims to address this gap by creating a comprehensive chromosomal-level assembly of Kazakh individuals' whole genomes using advanced genomic technologies, including next-generation sequencing (NGS, Illumina), third-generation sequencing (TGS, Oxford Nanopore), optical genome mapping, and Hi-C chromosomal conformation capture.

Materials and methods: We generated chromosome-scale assemblies of Kazakh individuals by integrating multiple modern genomic technologies. Whole-genome sequencing was performed using both long-read and short-read platforms. The assemblies were polished with short-read data, scaffolded with Bionano optical genome maps, and

organized at the chromosomal level using Hi-C chromosomal conformation data. Quality assessment was conducted at each stage, and comparative analyses will be performed against other global reference genomes.

Results: The project delivered high-quality, chromosome-level assemblies for individuals of the Kazakh population. These assemblies provide insights into unique genetic features, improve the accuracy of population-specific genomic analyses, and will be deposited in open-access repositories for use in biomedical and bioinformatics research worldwide.

Acknowledgement: This study is supported by a grant AP23490594 from the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan.

Keywords: Chromosome-scale assembly; Kazakh genome; long-read sequencing; Bionano optical mapping; Hi-C; de novo assembly; population genomics.