

Auramatcher: Scalable Reanalysis of 17,000 Clinical Genomes to Boost Rare Disease Diagnostics within the French Genomic Medicine Initiative

Background: Genome sequencing (GS) has an overall diagnostic yield of 35% in rare diseases, with routine discovery of new pathogenic variants, including outside coding regions. Retrospective analyses improve diagnosis but raise organizational and computational challenges. Within the French Genomic Medicine Initiative 2025, GCS AURAGEN performs large-scale clinical GS, processing up to 360 germline genomes at 30× weekly. By the end of 2025, more than 50,000 genomes had been analyzed.

Material and Methods: The AURAGEN bioinformatics team developed Auramatcher, an in-house tool similar to Genematcher. Using proactive literature monitoring, it enables targeted interrogation of GS data and live extraction of rare SNVs/indels. Reported variants include rare ones (gnomAD <1%, allelic balance >25%) predicted to affect coding or splicing, and those prioritized in Aurapport, a phenotype-driven shortlist. Mitochondrial, copy number, structural variants and short tandem repeats are available on request, with future direct access planned.

Results: Over 12 months, 1,944 genes were queried (≈160/month). Among them, 33 novel disease genes were identified in 54 families and returned after initial negative interpretation. Auramatcher also supports gene-based cohorts for transversal analysis of fully penetrant monogenic disorders.

Conclusion: This rapid targeted request strategy applied to large-scale GS enables efficient reanalysis and is a practical, scalable way to improve diagnostic yield after negative reports. It complements existing reanalysis workflows and supports iterative reinterpretation over time in routine clinical practice settings internationally. Continuous monitoring of new pathogenic variants and genes will combine variant- and gene-driven reanalyses to improve diagnosis of suspected rare disorders in diverse populations.