

MPEG-G's formatted WES datasets have 76%-95% less size compared to legacy formats.

How does the Compression work?

MPEG-G's compression benefit is a result of a 2 stage compression process. In the 1st stage, MPEG-G **splits the data into specific Descriptors** other than legacy formats. In the 2nd stage, by leveraging **Optimized Entropy Coders**, **redundancies** amongst the data are **removed** efficiently and in a lossless way and thus resulting in smaller file sizes.

What is MPEG-G?

MPEG-G (ISO/IEC 23092) is an ISO international standard for the representation of genome sequencing data and associated metadata.

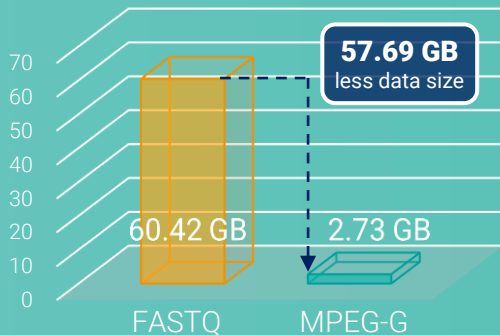
The standard MPEG-G aims to provide a framework for developing interoperable applications towards genuinely efficient and economical handling of genomic information.

Use-Case: Compression Benefit

A genomic laboratory is analyzing *8'000 samples per year with Whole-Exome Sequencing (WES) with 100x mean coverage*. Their average FASTQ file amounts to *60.42 GB* while converted into MPEG-G the same file is *2.73 GB*. The genomic data is stored in the cloud, using *AWS S3 standard storage*, costing *0.0235 USD per GB*.

GenomSys' professional solution offers decreased storage costs by a factor up to 22.

Size comparison for a single WES dataset in FASTQ & MPEG-G.



Annual storage costs comparison of WES files in FASTQ & MPEG-G

