

AWS HealthOmics ??? ? ?? ??

_____ , _____

AWS HealthOmics 是 一个 托管 的 基因组 数据 分析 平台 , 它 允许 用户 在 云上 存储 基因组 数据 并 进行 分析 。 它 支持 各种 格式 的 基因组 数据 , 包括 FASTQ , BAM , CRAM 等 。 用户 可以 使用 各种 分析 工具 对 数据 进行 分析 , 并 将 结果 存储 在 云上 。

HealthOmics 是 一个 托管 的 基因组 数据 分析 平台 。 Healthomics Storage 是 一个 托管 的 基因组 数据 存储 平台 。 Healthomics Analytics 是 一个 托管 的 基因组 数据 分析 平台 。 Healthomics Workflows 是 一个 托管 的 基因组 数据 分析 平台 。

AWS HealthOmics? ???????

?? ??

HealthOmics 是 一个 托管 的 基因组 数据 分析 平台 , 它 允许 用户 在 云上 存储 基因组 数据 并 进行 分析 。 它 支持 各种 格式 的 基因组 数据 , 包括 FASTQ , BAM , CRAM 等 。 用户 可以 使用 各种 分析 工具 对 数据 进行 分析 , 并 将 结果 存储 在 云上 。

HealthOmics 是 一个 托管 的 基因组 数据 分析 平台 , 它 允许 用户 在 云上 存储 基因组 数据 并 进行 分析 。 它 支持 各种 格式 的 基因组 数据 , 包括 FASTQ , BAM , CRAM 等 。 用户 可以 使用 各种 分析 工具 对 数据 进行 分析 , 并 将 结果 存储 在 云上 。

- **Storage** - ?? (FASTQ, BAM, CRAM? ?? ????? ?? ????? ?? ???)
- **Analytics** - ?????? ??? (? , VCF? ??? ?? ???)
- **Workflows** - ?????? (Nextflow, WDL, CWL ? ????? ?????? ????? ??? ? ??? ?? ????? ?? ??? ???)

??

- 它 支持 各种 格式 的 基因组 数据 , 包括 FASTQ , BAM , CRAM 等 。 用户 可以 使用 各种 分析 工具 对 数据 进行 分析 , 并 将 结果 存储 在 云上 。

- [IAM](#) [Amazon CloudWatch](#)
- [S3](#) [Intelligent Tiering](#)
- [EBS](#) [EFS](#)

Storage

- Reference store [S3](#)
- Sequence store [S3](#) [Intelligent Tiering](#)

Workflows

Ready2Run

Ready2Run (R2R) [GATK](#), [Singlecell](#), [AlphaFold](#)

- AWS HealthOmics
- Getting started
- Storage
 - Reference store
 - Sequence stores
- Analytics
 - Variant stores
 - Annotation stores
- Workflows
 - Ready2Run workflows
 - Private workflows
 - Runs
 - Run groups
- Data shares
 - All data shares
- Documentation



Step 1
Select a Ready2Run workflow - new
Ready2Run workflows are pre-built pipelines from third-party software companies and open-source communities.



Step 2
Add run parameters
Create a run by setting workflow run parameter values, run priority, and output location.



Step 3
Start a run
Choose Start run to perform a Ready2Run workflow your run in the Runs list page.

Ready2Run workflows (36) [info](#) [View](#)

	Workflow ID	Workflow name	Version	List price/run	Publisher	Subscription	Est. run time (hh:mm)	Last published
<input type="radio"/>	6094971	AlphaFold for 601-1200 residues	1.0	\$9.00	DeepMind	Not required	11:15	May 15th, 2023
<input type="radio"/>	4885129	AlphaFold for up to 600 residues	1.0	\$6.00	DeepMind	Not required	07:30	May 15th, 2023
<input type="radio"/>	2374431	Bases2Fastq for 2x150	1.0	\$2.62	Element Biosciences	Not required	01:45	May 15th, 2023
<input type="radio"/>	1578479	Bases2Fastq for 2x300	1.0	\$2.57	Element Biosciences	Not required	01:30	May 15th, 2023
<input type="radio"/>	4137328	Bases2Fastq for 2x75	1.0	\$0.84	Element Biosciences	Not required	00:45	May 15th, 2023
<input type="radio"/>	1830181	ESMFold for up to 800 residues	1.0	\$0.25	Meta Research	Not required	00:15	May 15th, 2023
<input type="radio"/>	3768383	GATK-BP fq2bam	1.0	\$8.00	Broad Institute	Not required	10:10	May 15th, 2023
<input type="radio"/>	5454617	GATK-BP Germline bam2vcf for 30x genome	1.0	\$4.58	Broad Institute	Not required	02:45	May 15th, 2023
<input type="radio"/>	9500764	GATK-BP Germline fq2vcf for 30x genome	1.0	\$10.00	Broad Institute	Not required	12:30	May 15th, 2023
<input type="radio"/>	5562080	GATK-BP Somatic WES bam2vcf	1.0	\$4.68	Broad Institute	Not required	01:30	May 15th, 2023

- VCF files can be converted to TSV/CSV or GFF3 files
- HealthOmics can convert VCF files to TSV/CSV or GFF3 files
- Amazon Athena can query VCF files (Athena supports 3 file formats)

Annotation files: VCF (Variant Effect Predictor (VEP) annotations ?? ??.)

Annotation files: TSV, VCF, or GFF

?? ??

AWS CLI — HealthOmics can be accessed via the AWS CLI.

AWS CLI (AWS CLI) — AWS CLI is a command-line tool that allows you to interact with AWS services. It is available on Windows, macOS, and Linux. For more information, see the AWS CLI documentation.

AWS SDK — AWS SDKs are available for various programming languages (Java, Python, Ruby, .NET, iOS, and Android). They provide a programmatic way to interact with AWS services. For more information, see the AWS SDK documentation.

AWS API — The AWS API is a RESTful interface that allows you to interact with AWS services. HealthOmics can be accessed via the AWS API.

?? ??? ? ? Quota

<https://docs.aws.amazon.com/general/latest/gr/healthomics-quotas.html>

?? ??

HealthOmics Private workflow can be accessed via the AWS CLI.

- WDL workflows
 - [GATK Best Practice workflows](#)
 - [Analysis ready germline BAM to VCF](#)
 - [CRAM to BAM](#)
 - [FASTQs to analysis ready BAM](#)

- [Germline FASTQs to VCF](#)
- [Somatic SNPs and InDELS](#)
- [Protein folding workflows](#)
 - [AlphaFold](#)
 - [ESMFold](#)
- [Other WDL workflows](#)
 - [HISAT-Genotype HLA Caller](#)
- Nextflow workflows
 - [NF-Core workflows](#)
 - [FASTQC](#)
 - [RNAseq](#)
 - [scRNAseq-cellranger](#)
 - [scRNAseq](#)
 - [TaxProfiler](#)
 - [Other Nextflow workflows](#)
 - [VEP](#)

<https://github.com/aws-samples/amazon-omics-tutorials/tree/main/example-workflows>

Learn more about HealthOmics from these workshops and tutorials:

- HealthOmics workshop – [HealthOmics end to end workshop](#)
 - [Migrating Nf-core Workflows Into AWS HealthOmics](#)
- AWS genomics resources – [Public Amazon ECR repositories](#) related to genomics
- Python tutorials – [Jupyter notebook tutorials](#) on GitHub, covering HealthOmics storage, analytics, and workflows

Become familiar with additional HealthOmics tools that AWS provides:

- WDL linter – [HealthOmics linter for WDL](#)
- Nextflow linter – [HealthOmics linter for Nextflow](#)
- HealthOmics Amazon ECR helper tool – [Amazon ECR helper tool for HealthOmics](#)
- HealthOmics tools on GitHub – [Tools for working with HealthOmics](#) (Transfer manager, URI parser, Omics rerun, Run analyzer).

?? ???

- Secure Your Genomic Workflows and Data with AWS HealthOmics
- Category: AWS HealthOmics / AWS for Industries

Revision #9

Created 11 April 2024 01:44:55 by Hyunmin Kim

Updated 19 February 2025 06:19:35 by Hyunmin Kim