

ACTFusion V5 “gap analysis” execution log

Bioinformatics Development

Sandy

2024.12-2025.03

Gap analysis log (tool comparison)

Tools	Command	Image (built/pulled locally on server 177)	Note	Result folder
v0.28.0 (v4), Fusion v4 pipeline	nextflow run \ /mnt/BI3/Team_workdir/sandyteng_workdir/ACTFusionV4_Torrent/repo_dev/torrent_fusion_pipeline_nextflow/Workflow_docker_illumina.nf \ -params-file \ /mnt/RD_Develop/sandyteng/ACTFusionV4/nextflow/testparams/PA043ANA_merged.verificationruns_IVTALL.json \ -c \ /mnt/BI3/Team_workdir/sandyteng_workdir/ACTFusionV4_Torrent/repo_dev/torrent_fusion_pipeline_nextflow/dockerconfigs/illumina_multi_localdocker.v9.20241125.v0.22.0.config	actgenomics/torrent_fusion_pipeline:v0.22.0	Only evaluate "AANB02_184_IDD705504_IVTALL-1-AA-21-02200" (only may keep one sample within the json file to validate the result)	/mnt/RD_Develop/sandyteng/ACTFusionV4/nextflow_outdir/20241209_PA043ANA_IVTALL_3sample_v0.28.0/AANB02_184_IDD705504_IVTALL-1-AA-21-02200/
Arriba (2.4.0)	docker run --rm \ -v \ /mnt/RD_Develop/sandyteng/FusionCaptureTools/testresult/arriba_v2.4.0_grch38:/output \ -v /mnt/RD_Develop/sandyteng/FusionCaptureTools/refdb_arriba:/references:ro \ -v \ /mnt/LDT_Lv1_Illumina/AANB02/BCL/AANB02_184/AANB02_184_IDD705504/AANB02_184_IDD705504_IVTALL-1-AA-21-02200_S28_R1_001.fastq.gz:/read1.fastq.gz:ro \ -v \ /mnt/LDT_Lv1_Illumina/AANB02/BCL/AANB02_184/AANB02_184_IDD705504/AANB02_184_IDD705504_IVTALL-1-AA-21-02200_S28_R2_001.fastq.gz:/read2.fastq.gz:ro \ -v /path/to/run_arriba.sh:/run_arriba.sh:ro \ uhrigs/arriba:2.4.0 \ sh -c "cd /output; bash /arriba_v2.4.0/run_arriba.sh \ /references/STAR_index_GRCh38_RefSeq_hg38/ \ /references/RefSeq_hg38.gtf \ /references/GRCh38.fa \ /references/blacklist_hg38_GRCh38_v2.4.0.tsv.gz \ /references/known_fusions_hg38_GRCh38_v2.4.0.tsv.gz \ /references/protein_domains_hg38_GRCh38_v2.4.0.gff3 \ 8 \ /read1.fastq.gz \ /read2.fastq.gz"	uhrigs/arriba:2.4.0	The image can be found in dockerhub (https://hub.docker.com/r/uhrigs/arriba/tags)	/mnt/RD_Develop/sandyteng/FusionCaptureTools/testresult/arriba_v2.4.0_grch38/
STAR-Fusion	docker run -v `pwd`:data -v /mnt:/mnt/ --rm trinityctat/starfusion \ STAR-Fusion \ --left_fq \ /mnt/LDT_Lv1_Illumina/AANB02/BCL/AANB02_184/AANB02_184_IDD705504/AANB02_184_IDD705504_IVTALL-1-AA-21-02200_S28_R1_001.fastq.gz \ --right_fq \ /mnt/LDT_Lv1_Illumina/AANB02/BCL/AANB02_184/AANB02_184_IDD705504/AANB02_184_IDD705504_IVTALL-1-AA-21-02200_S28_R2_001.fastq.gz \ --genome_lib_dir \ /mnt/RD_Develop/sandyteng/FusionCaptureTools/refdb_starfusion/GRCh38_gencode_v44_CTAT_lib_Oct292023.plugin-play/ctat_genome_lib_build_dir \ -O \ /mnt/RD_Develop/sandyteng/FusionCaptureTools/testresult/starfusion_gencode_v44_CTAT_lib_Oct292023_241212 \ --FusionInspector validate \ --examine_coding_effect \ --denovo_reconstruct	trinityctat/starfusion:latest	The image can be found in dockerhub (https://hub.docker.com/r/trinityctat/starfusion/)	/mnt/RD_Develop/sandyteng/FusionCaptureTools/testresult/starfusion_gencode_v44_CTAT_lib_Oct292023_241212/
CeGaT (customized STAR-Fusion)	bash /mnt/RD_Develop/sandyteng/FusionCaptureTools/cegat_starfusion/test_AANB02_184_IDD705504_IVTALL-1.sh	cegat_starfusion:latest	The image is built locally following the instructions provided on the website (https://cegat.com/fusions/)	/mnt/RD_Develop/sandyteng/FusionCaptureTools/testresult/cegat_starfusion/

Arriba workflow construction test

(Scenario 0)

- Purpose: Test if the arriba workflow can be successfully executed & compare the “**fusions.tsv**” obtained from the tool comparison result
 - Result folder (Tool comparison): /mnt/RD_Develop/sandyteng/FusionCaptureTools/testresult/arriba_v2.4.0_grch38/
 - Result folder (Pipeline): /mnt/RD_Develop/sandyteng/ACTFusionV5/nextflow_outdir/20250325_PA043ANA_IVTALL-1_arriba_test/
 - Ref. issue: <https://actg.atlassian.net/browse/ABIE-996>

ACTFusion V5 construction test

- Test scenarios
 - Arriba workflow test
 - (Scenario 1)
 - Execute pipeline using smaller datasets (multi-sample execution)
 - (Scenario 2)
 - Execute pipeline using NextSeq, hybrid capture data (8 samples)
 - Fusion v4 workflow test
 - (Scenario 3)
 - Check if fusion v4 v0.29.0 still works (2 fusion v4 test data)

Workflow construction test summary

- Ref. issue: <https://actg.atlassian.net/browse/ABIE-996>
- Ref. issue: <https://actg.atlassian.net/browse/ABIE-997>

Scenario	Workflow	Execution date	Result directory	Test summary
0	Arriba workflow	2025/03/26	/mnt/RD_Develop/sandyteng/ACTFusionV5/nextflow_outdir/20250325_PA043ANA_IVTALL-1_arriba_test/	Sample for pipeline construction see issue: https://actg.atlassian.net/browse/ABIE-996
1	Arriba workflow	2025/03/26	/mnt/RD_Develop/sandyteng/ACTFusionV5/nextflow_outdir/20250326_PA043ANA_IVTRNA_arriba_test/	Most of the in-silico (IVT) fusions are not reported by Arriba caller.
2	Arriba workflow	2025/03/26	/mnt/RD_Develop/sandyteng/ACTFusionV5/nextflow_outdir/20250326_Twist_8_NextSeq_samples_arriba_test/	The results (*fusions.tsv vs fusions.tsv) are identical as compared to "/mnt/RD_Develop/sandyteng/ACTFusionV5/20250122_Twist Bioscience/testresult/arriba_grch38/". (Executed via Arriba's containerized workflow)
3	Fusion v4 workflow	2025/03/26	/mnt/RD_Develop/sandyteng/ACTFusionV5/nextflow_outdir/20250326_PA043ANA_IVTRNA_fusionv4_test/	The workflow is copied from https://github.com/ACTGenomics/torrent_fusion_pipeline_nextflow Connect your Github account (pipeline version: v0.29.0)