HOWARD RELEASE NOTES

1 RELEASE NOTES

1.1 0.12.0

This release introduce prioritization options and transcripts view, improve samples managment, annotation databases generation and operations, and python packages stability.

1.1.1 News

- Add annotation method 'BigWig'
- Add prioritization options:
    - SQL syntax available to define filters
    - New 'Class' priorisation field
- New transcripts view:
    - Create a transcript view, using a structure from multiple source type (e.g. snpEff, external annotation databases)
    - Mapping between multiple transcript ID source (e.g. refSeq, Ensembl)
    - Transcripts prioritization, using same prioritization process than variants
    - Export transcripts table as a file, in multiple format such as VCF, TSV, Parquet
- Export with a specific sample list
- Configuration and parameters files in YAML format allowed
- Add dynamic transcript column for NOMEN calculation (using transcript priorization column)
- Add plugins:
    - 'update_databases'

1.1.2 Updates

- Improve snpEff annotations operations
- New option 'uniquify' for dbSNFP generation, identification of columns type
- Managment, check and export of samples columns
- Improve query type mode
- Improve splice annotation
- Improve NOMEN generation

1.1.3 Fixes

- Genotype format detection
- Fix packages releases
- Fix parameters and configuration files options
- Fix calculations list and parametrization
- More explicite log messages

1.2 0.11.0

This release introduce splice annotation tool, and update duckDB python package for improve stability.

1.2.1 News

- Add splice tool with docker image
- Add plugins:
    - 'genebe' (GeneBe annotation using REST API)

– 'minimalize' (Minimalize a VCF file, such as removing INFO/Tags or samples)

1.2.2   Updates

- DuckDB 1.0.0 stable Snow Duck (Anas Nivis) release
- Add API Documentation
- Improve tests

1.2.3   Fixes

- Paths parameters check fixed (genome and genomes-folders)
- Fix snpEff download error with databases list

1.3   0.10.0

This release is a refactor of HOWARD (Highly Open Workflow for Annotation & Ranking toward genomic variant Discovery) in Python, using Parquet and duckDB.

HOWARD annotates and prioritizes genetic variations, calculates and normalizes annotations, translates files in multiple formats (e.g. vcf, tsv, parquet) and generates variants statistics.

See README and gitHub for more explanations.

1.4   Previous releases

See HOWARD gitHub for more information about previous releases.

1.4.0.1   0.9b-07/10/2016:

- Script creation #### 0.9.1b-11/10/2016:
- Add Prioritization and Translation #### 0.9.1b-11/10/2016:
- Add snpEff annotation and stats #### 0.9.8b-21/03/2017:
- Add Multithreading on Prioritization and Translation #### 0.9.9b-18/04/2017:
- Add Calculation step #### 0.9.10b-07/11/2017:
- Add generic file annotation through --annotation option
- No need to be in configuration file
- Need to be in ANNOVAR database folder (file 'ASSEMBLY_ANN.txt' for annotation 'ANN')
- Add options: --force , --split
- Add options for VCFanotation.pl: --show_annoataion, --show_annotations_full
- Add database download option nowget in VCFanotation.pl
- Fixes: multithreading, VAF calculation, configuration and check dependencies #### 0.9.11b-07/05/2018:
- Replace VCFTOOLS command to BCFTOOLS command
- Release added into the output VCF
- Update SNPEff options
- Add VARTYPE, CALLING_QUALITY and CALLING_QUALITY_EXPLODE option on calculation
- Add description on calculations #### 0.9.11.1b-14/05/2018:
- Improve VCF validation
- Fix snpEff annotation bug #### 0.9.11.2b-17/08/2018:
- Add --vcf input vcf file option
- Create Output file directory automatically
- Improve Multithreading #### 0.9.12b-24/08/2018:
- Improve Multithreading
- Input VCF compressed with BGZIP accepted
- Output VCF compression level
- Add VCF input sorting and multiallele split step (by default)
- Add VCF input normalization step with option --norm
- Bug fixes #### 0.9.13b-04/10/2018:
- Multithreading improved
- Change default output vcf
- Input vcf without samples allowed
- VCF Validation with contig check

- Add multi VCF in input option
- Add --annotate option for BCFTOOLS annotation with a VCF and TAG (beta)
- Remove no multithreading part code to multithreading with 1 thread
- Remove --multithreading parameter, only --thread parameter to deal with multithreading
- Replace --filter and --format parameters by --prioritization and --translation parameters
- Add snpeff options to VCFannotation.pl #### 0.9.14b-21/01/2019:
- Reorganization of folders (bin, config, docs, toolbox...).
- Improve Translation (TSV or VCF, sort on fields, selection of fields, filtering on fields), especially memory efficiency
- Change Number/Type/Description of new INFO/FORMAT header generated
- Remove snpEff option --snpeff and --snpeff_hgvs. SnpEff is used through --annotation option
- Add '#' to the TAB/TSV delimiter format header
- Update dbNSFP config annotation file script
- Change default configuration files for annotation (add dbSNFP 3.5a, update mcap and regspintron) and prioritization
- Bug fixed: file identification in annotation configuration
- Bug fixed: calculation INFO fields header, snpeff parameters options on multithreading
- Bug fixed: snpeff parameters in command line #### 0.9.15b-19/09/2019:
- Rename HOWARD.sh to HOWARD.
- Add --nomen_fields parameter and update NOMEN calculation.
- Add --bcftools_stats and --stats parameter.
- Change PZScore, PZFlag, PZComment and PZInfos generation, adding default PZ and all PZ filters.
- Bug fixed: translation fields identification. #### 0.9.15.1b-28/05/2020:
- Bug fixed: NOMEN calculation clear previous NOMEN values if using force option. #### 0.9.15.2-12/10/2020:
- Change --norm parameter by adding '--check-ref=s' in bcftools command.
- Add --norm_options parameter.
- force translation VCF by default.
- Change VAF_stats and add DP_stats. #### 0.9.15.4-11/04/2021:
- Remove --snpeff_threads parameter (for snpEff 5.0e compatibility) and improve --snpeff_stats.
- Add a check and rehead INFO fields if necessary (prevent some incorrect INFO header format).
- Fix --compress parameter and add --index parameter. #### 0.9.15.5-21/07/2021:
- Add INFO description Type option, with autodetection.
- Add prioritization mode 'VaRank'/'max' for score calculation.
- Add calculation DP, AD, GQ and associated stats.
- Fix INFO field type for VCF. #### 0.9.15.6-16/09/2021:
- Structural variant compatibility.
- NOMEN extraction and generation improved, with --nomen_pattern option (only for SNV and InDel).
- Improve translation with variant sorting.
- Improve error catching.