Genome Enhancer release 3.0

New features

The 3.0 release of Genome Enhancer comes with the following new features:

- **Integration of somatic variants info into the analysis report**
- **New table of prospective drugs approved for the studied oncology**
- **Extended list of drugs approved for neoplasms with predicted drug scores assigned**
- **Extended clinical trials info**
- **Epigenomics analysis launch on CpG loci identifiers**
- **Improvement of genomics data analysis algorithm**
- **Database updates**
- **New and updated demo reports**

Integration of somatic variants info into the analysis report

We introduced the new '*Somatic variants*' column in two drug tables of the Genome Enhancer report: FDA approved drugs and repurposed drugs. This information is added to the drug tables when Genome Enhancer analyzes somatic mutations/variations in the tumor DNA (genomic data - exome, full genome) and reveals a significant association between predicted drugs and the identified DNA variations.

The column '*Somatic variants*' is generated with the help of the MTB (Molecular Tumor Board) algorithm incorporated in Genome Enhancer. MTB report can be generated for cases when genomics data is studied for a certain list of cancer-related pathologies. The MTB report is based on the mutation-drug predictive association method. Somatic variants of the patient under study are searched in curated databases of predictive biomarkers (GKDB, CIViC) and are reported according to their clinical evidence. The new '*Somatic variants*' column of Genome Enhancer reports only mutation-drug associations with the clinical evidence of levels 1 and 2. Level 1 means the evidence is supported by drug approval organizations or clinical guidelines, level 2 contains general clinical evidence (clinical trials, case reports). The evidence level 3 (preclinical

evidence) is ignored in the main Genome Enhancer report, but can be easily viewed in the general MTB report generated by the system. Both levels A and B are taken into account for construction of the '*Somatic variants*' column (level A means evidence was found in the same cancer type as for the patient under study, level B means evidence was found in any other cancer type).

Example visualization of new column '*Somatic variants*' is shown below:

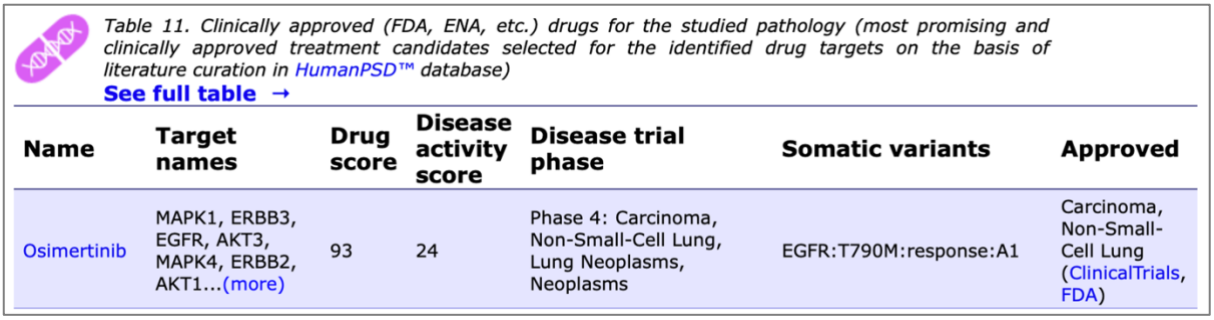

Table 11. Clinically approved (FDA, ENA, etc.) drugs for the studied pathology (most promising and clinically approved treatment candidates selected for the identified drug targets on the basis of literature curation in HumanPSD™ database)

See full table →

Name	Target names	Drug score	Disease activity score	Disease trial phase	Somatic variants	Approved
Osimertinib	MAPK1, ERBB3, EGFR, AKT3, MAPK4, ERBB2, AKT1...(more)	93	24	Phase 4: Carcinoma, Non-Small-Cell Lung, Lung Neoplasms, Neoplasms	EGFR:T790M:response:A1	Carcinoma, Non-Small-Cell Lung (ClinicalTrials, FDA)

The provided info includes the gene, in which the mutation was found, the mutation/variant(s) ID, the predicted type of associated drug action (resistance / response) and the level of evidence on the scale of A1-A2-B1-B2 briefly described above. Further details on the respective variant-drug associations can be found in the MTB report itself, which will be linked to the main analysis report.

You can check how somatic variants info is now integrated into the main analysis report of Genome Enhancer by viewing the new demo report Non-Small Cell Lung Carcinoma (NCI-H1975) --- Genomics, VCF, which is based on the analysis of NCI-H1975 cell line exhibiting epithelial morphology that was isolated in 1988 from the lungs of a nonsmoking female with non-small cell lung cancer. Respective VCF data that was loaded to Genome Enhancer for further automated processing was obtained from Cancer Cell Line Encyclopedia (CCLE).

New table of prospective drugs approved for the studied oncology

If all input diseases in the Genome Enhancer analysis appeared to be neoplasms, a new drug table '*Drugs approved in clinical trials for Oncology*', containing only the drugs that were clinically approved for the pathologies under study, will be constructed in addition to the standard drug tables provided by Genome Enhancer report. This brings the prospective drug candidates classification for the respective cases to 5 categories of drugs:

Clinically approved (FDA, ENA, etc.) drugs for the studied pathology (most promising and clinically approved treatment candidates selected for the identified drug targets on the basis of literature curation in HumanPSD™ database)

Drugs used in clinical trials for the studied pathology (most promising treatment candidates selected for the identified drug targets on the basis of literature curation in HumanPSD™ database)

Repurposed drugs used in clinical trials for other pathologies (prospective drugs against the identified drug targets on the basis of literature curation in HumanPSD™ database)

Prospective drugs, predicted by PASS software to be active against the identified drug targets with predicted activity against the studied disease(s) (drug candidates predicted with the chemoinformatics tool PASS)

Prospective drugs, predicted by PASS software to be active against the identified drug targets, though without cheminformatically predicted activity against the studied disease(s) (drug candidates predicted with the chemoinformatics tool PASS)

Extended list of drugs approved for neoplasms with predicted drug scores assigned

If at least one of the input diseases appeared to be a neoplasm, a new table '*Supplementary drug info'* will be provided below the tables with prospective drug candidates identified by Genome Enhancer for the studied pathology. This table includes an extended list of drugs generally used in clinic for treatment of neoplasms. In the '*Predicted Drug Score'* column of this table respective drug score on the scale from 1 to 100 predicted by Genome Enhancer for the studied case is shown (if drug targets of the respective treatment were not found by the system in the molecular mechanism of the studied pathology, the N/I (not available) value will be displayed). If MTB report was generated and any of the drugs from this supplementary table received a matching gene-drug predictive association, the '*Somatic variants*' column will provide respective info.

Example of updated prospective treatments classification for oncology and supplementary info on extended drugs used for treatment of neoplasms can be viewed in Genome Enhancer demo report Colorectal Cancer (Personalized patient data) --- Genomics, VCF or in the new demo report Non-Small Cell Lung Carcinoma (NCI-H1975) --- Genomics, VCF.

Extended clinical trials info

New columns with clinical trials info were added to the drug tables in Genome Enhancer report: '*Disease trial phase*' for FDA approved drugs or drugs used in clinical trials for the studied pathology and '*Maximum trial phase*' for repurposing drugs. The '*Disease trial phase*' column displays the maximum clinical trials phase in which the drug was tested for the studied pathology. The '*Maximum trial phase*' column displays the maximum clinical trials phase in which the drug was tested for any pathology.

Epigenomics analysis launch on CpG loci identifiers

Starting from this release, Genome Enhancer is able to analyze DNA methylation data provided in the format of CpG loci IDs with respective numerical values without any additional omics data needed for the analysis. The Illumina Human Methylation BeadChip data can be submitted to Genome Enhancer in table format containing cg IDs and respective ß-values or other numerical data corresponding to the respective CpG sites. Details on processing of such data can be found in the new demo report Hypertension (GSE157131) --- Epigenomics, cg lists or in the extended description of Genome Enhancer analysis algorithm. Data used in the demo report is from GSE157131 - methylation data from stored peripheral blood leukocytes of African American participants in the GENOA study.

Improvement of genomics data analysis algorithm

If genomics data was submitted to Genome Enhancer analysis, additional weights reflecting the transcription factor binding affinity change caused by the mutation will be calculated. The list of 40 matrices most affected by variations will be further used in composite modules search (CMA algorithm). New weights will be also summed up with previously calculated weights which are dependent on the mutation localization (exon region / promoter region / other location) and the summed weight will be used to find the regulatory regions of genes most affected by variations (a sliding window of 1100 bp is used to scan through the intronic, 5' and 3' regions of genes and a region with highest sum of the mutation weights will be selected). More details on the new variations weighting procedure can be found in the extended description of Genome Enhancer analysis algorithm or inside the Methods section of demo reports constructed on genomics data: the Colorectal Cancer (Personalized patient data) --- Genomics, VCF and new demo project Non-Small Cell Lung Carcinoma (NCI-H1975) --- Genomics, VCF.

Database updates

TRANSFAC®, TRANSPATH® and HumanPSD™ databases used in the Genome Enhancer analysis were updated to the release 2022.1. The CIViC database used for MTB report construction was updated to version 01 March 2022.

New and updated demo reports

All demo reports of Genome Enhancer pipeline were updated to the analysis results provided by Genome Enhancer release 3.0. Two new demo reports were added: hypertension study based on DNA methylation data (epigenomics, CpG loci IDs) and non-small cell lung carcinoma study based on the genomics VCF input. All demo reports can be freely accessed with a demo account at https://ge.genexplain.com or on Genome Enhancer product page of our web site.

Genome Enhancer Expert

Starting from release 2.0 Genome Enhancer offers to its users a powerful synergism between the automatic pipeline for multi-omics data processing of Genome Enhancer and the comprehensive bioinformatics toolbox of the geneXplain® platform.

You can contact us via info@genexplain.com to find out how to upgrade your current geneXplain license to the Genome Enhancer Expert solution.