DrugMatrix™

DrugMatrix Data Warehouse Schema Documentation

Release DM300_122004

March 2011

NIEHS/NTP

111 TW Alexander Drive
Durham, NC 27709
http://ntp.niehs.nih.gov/
http://www.niehs.nih.gov/

DrugMatrix™

Table of Contents

Conceptual Overview

The DrugMatrix data warehouse is a modified and highly-denormalized star-schema. The "hubs" of the schema are the 6 main information domains, or schema dimensions: Gene, Compound, Expression Experiment, Expression Study, Pathway, and Assay. These are represented by the gold, green, orange and blue rectangles in Figure 1. These "hubs" represent the main information domains in the DrugMatrix user interface. In addition to the domains, there are other types of data that represent either an annotation of a domain item, or an annotation that connects two domain items, both types of star-schema "fact" tables.

The singly connected annotations, as represented by ovals in Figure 1, contain additional details describing the main domain data. For example, some annotations for a specific drug would include the known adverse effects, the indications, the therapeutic class, etc. All of these sorts of annotations can be accessed from the small right panel in the compound report of the DrugMatrix user interface or from buttons near the top of each report.

The doubly-connected annotations, represented by the gray rectangles in Figure 1, connect two main information domains and contain additional information. For example, an Expression Change Data Point connects a specific Expression Experiment to a specific Gene and contains the log ratio, p-value and standard error for that expression change.

The diagram in Figure 1 provides a high level overview of the main information items in the DrugMatrix data warehouse and user interface, but is highly simplified in that it represents only 13 of the 70+ tables in the warehouse.
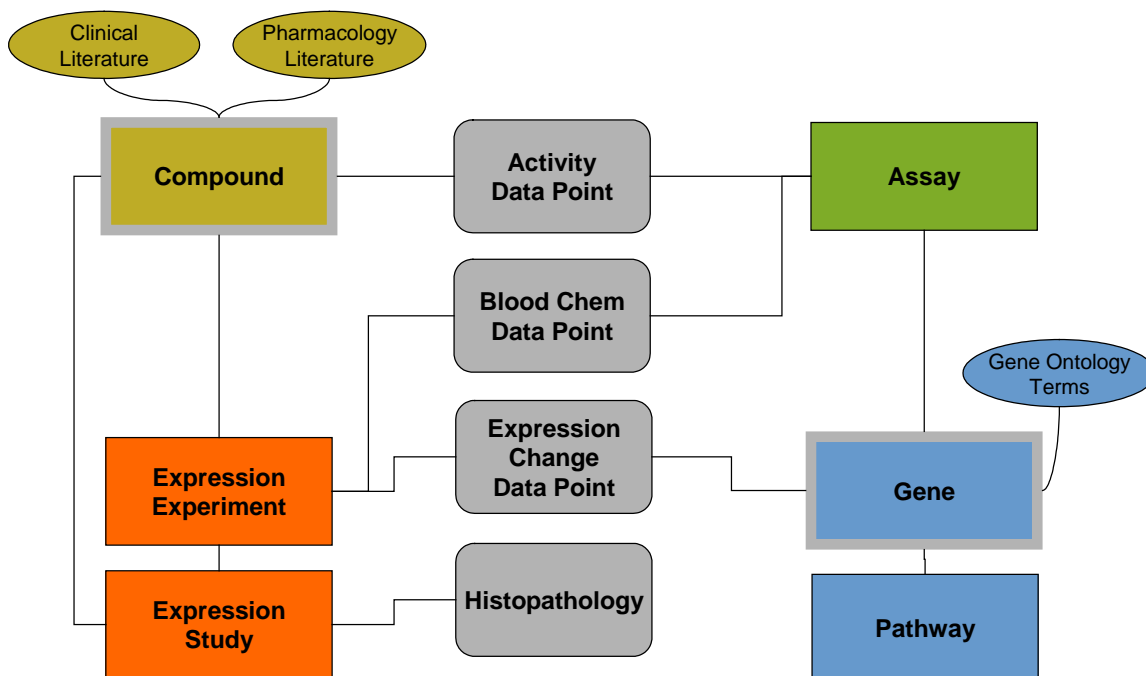
Figure 1: Conceptual Model for DrugMatrix Data Warehouse

General Principles

The 6 main domains of the conceptual model are represented by dimension tables in a star-schema (see Appendix A: Schema Diagram). Since the data warehouse dimensions are referenced by many tables in the star schema, and there are sometimes multiple tables that could be used as dimension tables, it is important to understand those tables that best represent the dimensions of the star schema. The tables that best-represent the main dimensions in the data warehouse are detailed in Table 1.

Main Domains	Tables
Compound	COMPOUND_REPORT
Gene	TARGET_REPORT
Expression Experiment	EXPERIMENT_CONDITION_REPORT
Expression Study	STUDY_REPORT
Assay	ASSAY_REPORT
Pathway	PATHWAY_REPORT

Table 1: Main Domains and Representative Data Warehouse Tables

Please note in the schema diagram in Appendix A that the COMPOUND_REPORT and TARGET_REPORT tables have the highest number of connections to other tables. In fact, these are considered the primary domains of the Chemogenomic data model in DrugMatrix.

In addition, there are other, minor dimension tables in the schema. These represent dimensions in the star schema, but have much fewer connections that the main dimension tables.

Other Domains	Tables
Signature	SIGNATURE_REPORT
Expresson	EXPRESSON_REPORT
Probe	EXPRESSON_REPORT
Chip Map	MAP_REPORT
Hybridization (single array)	HYBRIDIZATION_AND_IMAGE_REPORT
Animal Clinical Data	ANIMAL_ANNOTATION_REPORT

Table 2: Additional Dimensions and Representative Data Warehouse Tables

Please note that the DrugMatrix is designed as a data warehouse, optimized for a read-only environment supporting the DrugMatrix application. Some properties of more standard databases and data warehouses are listed in Table 3.

Databases	Data Warehouses
3rd Form Normalized	Denormalized
Transactional	Read Only
Data in one place	Multiple copies of data

Table 3: Databases and Data Warehouses

Nomenclature

Primary and foreign keys are always numeric and are represented by a simple column name. The main domain primary keys in the system are:

COMPOUND – a numeric identifier for a COMPOUND_REPORT record.
GENE or TARGET – a numeric identifier for a TARGET_REPORT record.
ASSAY – a numeric identifier for an ASSAY_REPORT record.
EXPERIMENT – a numeric identifier for an EXPERIMENT_CONDITION_REPORT record.
STUDY – a numeric identifier for a STUDY_REPORT record.
PATHWAY – a numeric identifier for a PATHWAY_REPORT record.

In the highly-denormalized warehouse schema, names of primary domain records are also provided. These columns always have compound names, with the term "_NAME" appended to the name, e.g. GENE_NAME, COMPOUND_NAME, ASSAY_NAME, EXPERIMENT_NAME, STUDY_NAME, and PATHWAY_NAME.

DrugMatrix™

Useful Sub-schemas

Gene Transcriptional Changes

Gene transcript level changes are expressed in DrugMatrix as the Log Ratio of the treated over the untreated samples. In general, using log ratio values is the simplest way of working with the gene expression changes that relate to drug treatments. The ALL_TRANSCRIPT_REPORT stores one record for each expresson for each expression experiment. Since there is a large amount of expression data in DrugMatrix, this table is quite large (~30 M rows) and any queries that use this table need to be engineered to accommodate the large table size.

ALL_TRANSCRIPT_REPORT

Raw Gene Expression Signals

All of the raw gene expression data that is used in calculating the log ratio data is also present in the DrugMatrix data warehouse. Raw data are found in the RAW_DATA_REPORT. Raw data are stored as one row for each probe for each hybridization (i.e. array). The RAW_DATA_REPORT represents the single largest table in the DrugMatrix data warehouse, with ~130 M rows. Develop and test queries that involve the RAW_DATA_REPORT carefully since these queries can create very large loads on the Oracle server. To use the RAW_DATA_REPORT data effectively, you will need to join to three other tables. The EXPERIMENT_CONDITION_REPORT contains information on the experimental treatments that can be tied to the individual array readings. To connect from the RAW_DATA_REPORT to the EXPERIMENT_CONDITION_REPORT, you will need to join through an intermediate table, the HYBRIDIZATION_AND_IMAGE_REPORT, which maps between these tables as indicated in the following diagram. In addition, you will join from the RAW_DATA_REPORT to the EXPRESSON_REPORT to map from probe-level information to expresson- and gene-level information, which is probably desired to improve interpretability.

EXPRESSON_REPORT

PROBE

RAW_DATA_REPORT

HYBRIDIZATION

HYBRIDIZATION_AND_IMAGE_REPORT

EXPERIMENT

EXPERIMENT_CONDITION_REPORT

DrugMatrix™
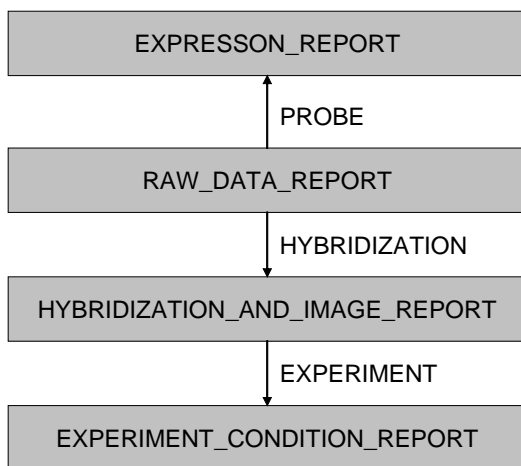
Molecular Pharmacology

The molecular pharmacology data from the MDS assays are included in two tables in the data warehouse that contain duplicate copies of the data. These tables are fully denormalized so mining these data usually only involves querying the individual tables. You can query either one of these tables since they have identical data, organized slightly differently. The denormalized reports contain information on the activity of the compound in % inhibition, IC50 and Ki, as well as the names of the compound and assay used. The GENE_ACTIVITY_REPORT has the gene_name in addition.

COMPOUND_ACTIVITY_REPORT

GENE_ACTIVITY_REPORT
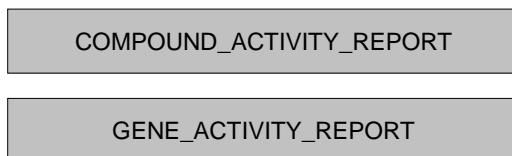
Gene Annotation

There are a number of reports that contain annotation information on the genes in DrugMatrix. These all reference the TARGET_REPORT as the dimension table for this domain.

TARGET_ACCESSION_REPORT

TARGET_EXPRESSION_SIMILARITY

TARGET_SYNONYM

↓ TARGET/GENE

TARGET_REPORT

↑ TARGET/GENE

GO_REPORT

GENE_ANNOTATION_REPORT

GENE_PATHWAY_REPORT
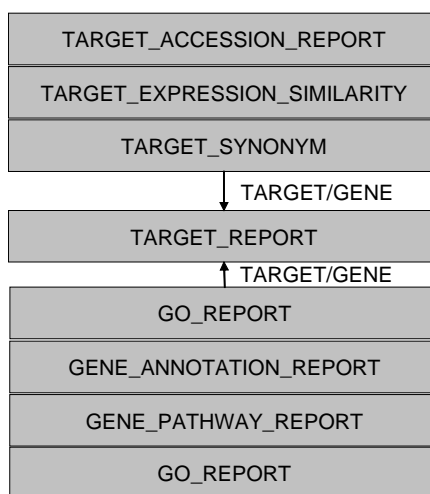
GO_REPORT

Compound Annotation

There are a number of reports that contain the annotation information (literature-based) for the drugs and compounds in DrugMatrx. These all reference the COMPOUND_REPORT as the dimension table for this domain.

COMPOUND_STRUCTURE_REPORT
COMPOUND_SYNONYM_REPORT
COMPOUND_TARGET_REPORT
COMPOUND_CURATION_REPORT
CAS_REPORT

↓ COMPOUND

COMPOUND_REPORT

↑ COMPOUND

COMPOUND_IDENTIFIER_REPORT
COMPOUND_LITERATURE_REPORT
COMPOUND_PROPERTY_REPORT
COMPOUND_SIMILARITY_REPORT
COMPOUND_SMILE_REPORT
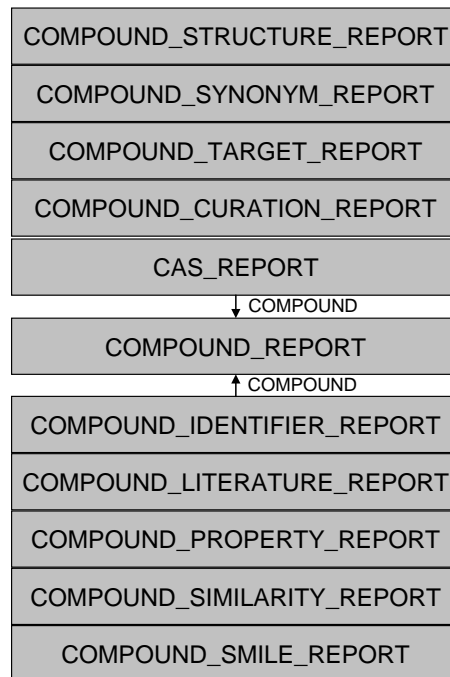
Description of Tables

ALL_TRANSCRIPT_REPORT

Type: Fact

The ALL_TRANSCRIPT_REPORT contains all of the log ratio gene expression data in DrugMatrix. It has a very large number of rows so queries designed for this table must be optimized for performance. This highly-denormalized table contains most of the information that is relevant to the gene expression experiments measured.

COLUMN_NAME	DATA_TYPE	NULLABLE	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
EXPRESSON	NUMBER	N	FK->EXPRESSON_REPORT.EXPRESSON
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
DIFFERENTNESS	NUMBER	Y	
LOG_RATIO	NUMBER	Y	
STDEV_OF_LOG_RATIO	NUMBER	Y	
SCORE	NUMBER	Y	
INTENSITY	NUMBER	Y	
TIME	NUMBER	N	
TIME_UNIT	CHAR	Y	
DOSE	NUMBER	Y	
DOSE_UNIT	VARCHAR2	Y	
ORGANISM_DESCRIPTION	VARCHAR2	N	
CHIP_NAME	VARCHAR2	N	
TYPE	VARCHAR2	N	
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
MAP	NUMBER	N	FK.MAP_REPORT.MAP
TISSUE	NUMBER	N	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE

ANIMAL_ANNOTATION_REPORT

Type: Dimension, Minor

The ANIMAL_ANNOTATION_REPORT is a minor domain table that contains information on all of the animals (biological samples) in the DrugMatrix database. Since tissue from one animal may be used in multiple experiments and on multiple arrays (hybridization), this denormalized table does not have one solitary column that can act as a unique key. The combination of the ANIMAL_ID, EXPERIMENT and HYBRIDIZATION columns do provide a compound unique key.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ANIMAL_ID	VARCHAR2	Y	
SAMPLE_ID	NUMBER	Y	
STUDY	NUMBER	Y	FK->STUDY_REPORT.STUDY
EXPERIMENT	NUMBER	Y	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
HYBRIDIZATION	NUMBER	Y	
VALID	CHAR	Y	
TYPE	CHAR	Y	

ARRAY_EXPERIMENT_REPORT

Type: Dimension (Deprecated)
This table is deprecated and all required information has been added to the
EXPERIMENT_CONDITION_REPORT, which should be used in its place.

ASSAY_REPORT

Type: Dimension, Major
As the dimension table for the Assay Domain, the ASSAY_REPORT contains information
including the name, type (BINDING, ENZYME, METABOLISM, BLOOD_CHEM, HEMATOLOGY),
species, and MDS Pharma Services catalog number for certain assay types.
For information on the assay-to-gene link, see the ASSAY_TARGET_REPORT.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ASSAY	NUMBER	N	PK
ASSAY_NAME	VARCHAR2	N	
TYPE	VARCHAR2	N	
SPECIES	VARCHAR2	Y	
CATALOG_ID	VARCHAR2	Y	
COMPOUND_ACTIVITIES	NUMBER	N	DEPRECATED

ASSAY_TARGET_REPORT

Type: Fact
The ASSAY_TARGET_REPORT contains the connecting links between the ASSAY and the
GENE domains. It contains no other information than the connection.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ASSAY	NUMBER	N	FK->ASSAY_REPORT.ASSAY
ASSAY_NAME	VARCHAR2	N	
GENE	NUMBER	Y	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	

BLOOD_REPORT

Type: Fact
The BLOOD_REPORT contains all of the averaged clinical chemistry and hematology assay data for the drug treatments in DrugMatrix. Each data point in this table represents the average of the 3 biological replicates for a particular treatment group. This denormalized table contains all of the information regarding the attributes of the treatment.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ASSAY	NUMBER	N	FK->ASSAY_REPORT.ASSAY
ASSAY_NAME	VARCHAR2	N	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
TIME	NUMBER	N	
TIME_UNIT	CHAR	Y	
DOSE	NUMBER	Y	
DOSE_UNIT	VARCHAR2	Y	
ORGANISM_DESCRIPTION	VARCHAR2	N	
CHIP_NAME	VARCHAR2	N	
TYPE	VARCHAR2	N	
SCORE	NUMBER	N	
LOG_RATIO	NUMBER	N	
STDEV_OF_LOG_RATIO	NUMBER	N	
AVG_VALUE	VARCHAR2	Y	
NORMAL_RANGE	VARCHAR2	Y	

CAS_REPORT

Type: Fact
The CAS_REPORT contains the CAS identifiers for all of the compounds in DrugMatrix.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
CAS_NUMBER	VARCHAR2	N	

COMPOUND_ACTIVITY_REPORT

Type: Fact
The COMPOUND_ACTIVITY_REPORT contains the molecular pharmacology assay results for all of the compounds in DrugMatrix. This denormalized report contains the activities measured as % inhibition, IC50 and Ki.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_LABEL	VARCHAR2	N	
COMPOUND_NAME	VARCHAR2	Y	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE
INH_ACTIVITY	NUMBER	N	
INH_UNIT	VARCHAR2	N	
IC50_ACTIVITY	NUMBER	Y	
IC50_UNIT	VARCHAR2	Y	
KI_ACTIVITY	NUMBER	Y	
KI_UNIT	VARCHAR2	Y	
ASSAY	NUMBER	N	FK->ASSAY_REPORT.ASSAY
ASSAY_NAME	VARCHAR2	N	

COMPOUND_CURATION_REPORT

Type: Fact

The COMPOUND_CURATION_REPORT contains the clinical (based on PDR) and activity class literature curation terms for each compound. The CATEGORY column describes the category of curation term and the TERM column contains the annotation term itself.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
CATEGORY	VARCHAR2	N	
TERM	VARCHAR2	N	

COMPOUND_IDENTIFIER_REPORT

Type: Fact

The COMPOUND_IDENTIFIER_REPORT contains multiple identifiers from various databases for each compound.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
COMPOUND_IDENTIFIER	VARCHAR2	Y	

COMPOUND_LITERATURE_REPORT

Type: Fact

The COMPOUND_LITERATURE_REPORT contains molecular and cellular pharmacology activity measures curated from the literature for each compound.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE
GENE	NUMBER	Y	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
TYPE	VARCHAR2	N	
ACTIVITY	NUMBER	N	
UNITS	VARCHAR2	N	
DESCRIPTION	VARCHAR2	Y	
PUBMED	NUMBER	Y	

COMPOUND_MOTIF_REPORT

Type: Fact (Beta)

The COMPOUND_MOTIF_REPORT is a beta-quality experimental table that contains motif similarity scores for each compound.

COMPOUND_PROPERTY_REPORT

Type: Fact

The COMPOUND_PROPERTY_REPORT contains computed and measured physical properties for each compound in DrugMatrix. The CATEGORY column describes the type of property.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
CATEGORY	VARCHAR2	N	
VALUE	NUMBER	Y	
UNITS	VARCHAR2	Y	

COMPOUND_REPORT

Type: Dimension, Major

As the dimension table for the Compound Domain, the COMPOUND_REPORT contains information on the compound name, molecular weight, molecular formula, ICX Number, drug development status and SMILES string.

For more compound information, also see CAS_REPORT, COMPOUND_CURATION_REPORT, COMPOUND_IDENTIFIER_REPORT, COMPOUND_PROPERTY_REPORT, COMPOUND_SMILES_REPORT, and COMPOUND_STRUCTURE_REPORT.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	PK
COMPOUND_NAME	VARCHAR2	Y	
COMPOUND_IDENTIFIER	NUMBER	N	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE
MOL_WEIGHT	NUMBER	Y	
FORMULA	VARCHAR2	Y	
ICXNUMBER	VARCHAR2	N	
DEVELOPMENT_STATUS	VARCHAR2	Y	
SMILES	VARCHAR2	Y	

COMPOUND_SIMILARITY_REPORT

Type: Fact

The COMPOUND_SIMILARITY_REPORT contains the significant computed pairwise structural similarities between all compounds in DrugMatrix using a number of chemical structure comparison algorithms. The METHOD column names the structural similarity method used.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE
PROBABILITY	NUMBER	Y	
SCORE	NUMBER	N	
METHOD	VARCHAR2	N	
QUERY_COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
QUERY_COMPOUND_NAME	VARCHAR2	Y	

COMPOUND_SMILE_REPORT

Type: Fact
The COMPOUND_SMILE_REPORT contains the SMILES string, a structural representation, for each compound in DrugMatrix.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
SMILES	VARCHAR2	Y	

COMPOUND_STRUCTURE_REPORT

Type: Fact
The COMPOUND_STRUCTURE_REPORT contains the MOLFILE, a structural representation, for each compound in DrugMatrix. The MOLFILEs are also available from MDL's ISIS product as delivered in the companion DMS database schema.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
STRUCTURE	CLOB	Y	

COMPOUND_SYNONYM_REPORT

Type: Fact
The COMPOUND_SYNONYM_REPORT contains multiple synonyms for each compound in DrugMatrix.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
SYNONYM_NAME	VARCHAR2	N	

COMPOUND_TARGET_REPORT

Type: Fact
The COMPOUND_TARGET_REPORT contains curated connections from Drug to Target as obtained from the literature and encoded in DrugMatrix.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	N	

DOSE_JUSTIFICATION_REPORT

Type: Fact
The DOSE_JUSTIFICATION_REPORT contains the dose justification document associated with each DrugMatrix expression study.

COLUMN_NAME	DATA_TYPE	NULLABLE	
STUDY	NUMBER	N	FK->STUDY_REPORT.STUDY
JUSTIFICATION	CLOB	N	

EXPERIMENT_CONDITION_REPORT

Type: Dimension, Major
As the dimension table for the Expression Experiment Domain, the
EXPERIMENT_CONDITION_REPORT contains extensive information regarding the details of the
treatment for each expression experiment, including the name, tissue, compound, dose, vehicle
(solvent), route of administration, cell culture medium, treatment duration (time), and the
percentage of genes significantly induce, repressed and changed (perturbed).
For more gene information, also see TARGET_SYNONYM, TARGET_ACCESSION_REPORT
and GO_REPORT.

COLUMN_NAME	DATA_TYPE	NULLABLE	
EXPERIMENT	NUMBER	N	
EXPERIMENT_VERSION	NUMBER	Y	
EXPERIMENT_DATE	CHAR	Y	
RATIO_CREATED_DATE	DATE	Y	
ARRAY_TECHNOLOGY	VARCHAR2	Y	New Column
EXPERIMENT_NAME	VARCHAR2	Y	
CHIP_NAME	VARCHAR2	N	Formerly SHORT_NAME
ORGANISM	VARCHAR2	N	
ORGANISM_DESCRIPTION	VARCHAR2	N	
TISSUE	NUMBER	N	
TISSUE_NAME	VARCHAR2	N	
TISSUE_PRIORITY	NUMBER	N	INTERNAL USE
COMPOUND	NUMBER	Y	
COMPOUND_NAME	VARCHAR2	Y	
ICXNUMBER	VARCHAR2	Y	
DOSE	NUMBER	Y	
DOSE_UNIT	VARCHAR2	Y	
DOSE_LEVEL	VARCHAR2	Y	
STUDY	NUMBER	N	New Column FK->STUDY_REPORT.STUDY
STUDY_NAME	VARCHAR2	N	New Column
CTLLABEL	VARCHAR2	Y	New Column
EXPLABEL	VARCHAR2	Y	New Column
CONTROL_SET	VARCHAR2	Y	
SOLVENT	VARCHAR2	Y	
ADMINISTRATION_ROUTE	VARCHAR2	Y	
ADMINISTRATION_FREQUENCY	VARCHAR2	Y	
MEDIUM	VARCHAR2	Y	
TIME	NUMBER	N	
TIME_UNIT	CHAR	Y	
TYPE	VARCHAR2	Y	
VALID	CHAR	Y	
MIN_REL_ERROR	NUMBER	Y	
PERCENT_INDUCED	NUMBER	Y	
PERCENT_REPRESSED	NUMBER	Y	
PERCENT_PERTURBED	NUMBER	Y	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE

EXPERIMENT_MOTIF_REPORT

Type: Fact (Beta)
The EXPERIMENT_MOTIF_REPORT is a beta-quality experimental table that contains motif
similarity scores for each experimentSS.

EXPERIMENT_ORGAN_RELWT_REPORT

Type: Fact
The EXPERIMENT_ORGAN_RELWT_REPORT contains the average necropsy weight for each organ for each experiment treatment group.

COLUMN_NAME	DATA_TYPE	NULLABLE	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
ORGAN	VARCHAR2	Y	
RELATIVE_WEIGHT	NUMBER	Y	

EXPERIMENT_PATHWAY_REPORT

Type: Fact (New, Deprecated)
The EXPERIMENT_PATHWAY_REPORT contains the gene expression data from all experiments for all genes in all pathways. It is a highly denormalized table that is primarily present to support the pathway impact tool. This implementation will be replaced by a more space-efficient mechanism in future releases so use of this table is deprecated. Instead, use the ALL_TRANSCRIPT_REPORT and the PATHWAY_GENE_REPORT.

COLUMN_NAME	DATA_TYPE	NULLABLE	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
PATHWAY	NUMBER	Y	FK->PATHWAY_REPORT.PATHWAY
PATHWAY NAME	VARCHAR2	Y	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
LOG_RATIO	NUMBER	Y	
SCORE	NUMBER	Y	

EXPERIMENT_PATHWAY_SCORE

Type: Data Staging (Deprecated)
The EXPERIMENT_PATHWAY_SCORE table is a temporary data table used during pathway impact computations. This table is present to support the internal data warehouse post-processing code and is not needed for any other purposes. The denormalized form of these data can be found in the EXP_PATHWAY_SCORE_REPORT which is also deprecated and will be removed in future versions. Pathway impact hypergeometric scores are no longer pre-computed and will no longer be part of the database for lookup.

EXPRESSION_CHART_REPORT

Type: Fact
The EXPRESSION_CHART_REPORT contains the contains precomputed information on the percentage of significantly changed genes in each expression experiment.

COLUMN_NAME	DATA_TYPE	NULLABLE	
REPLICATE	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
REPLICATE_NAME	VARCHAR2	Y	
REPRESSED	NUMBER	Y	
INDUCED	NUMBER	Y	
TESTED	NUMBER	Y	

DrugMatrix™

EXPRESSION_STUDY

Type: Dimension (Deleted)
This is a deleted table as of version 2.42 and should not be used. The same information can be found in the STUDY_REPORT.

EXPRESSON_REPORT

Type: Dimension, Minor
The EXPRESSON_REPORT is the minor domain table for the EXPRESSON data type. This table is most useful when mining array (hybridization) level expression data from the DrugMatrix data warehouse since it is reference by the RAW_DATA_REPORT. Since there can be more than one PROBE per EXPRESSON, this table can contain more than one row with the same EXPRESSON numerical identifier and the PROBE column is unique.

COLUMN_NAME	DATA_TYPE	NULLABLE	
EXPRESSON	NUMBER	N	NON-UNIQUE KEY
EXPRESSON_NAME	VARCHAR2	N	
SPECIES	NUMBER	N	
DNA_SEQUENCE	NUMBER	Y	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
PROBE_ACCESSION	VARCHAR2	Y	
PROBE	NUMBER	N	PK
PROBE_NAME	VARCHAR2	N	

EXP_ARRAY_PROBING_REP

Type: Fact (Deleted)
This is a deleted table as of version 2.42 and should not be used. The information that used to be in this table is no longer needed.

EXP_PATHWAY_SCORE_REPORT

Type: Fact (Deprecated)
The EXP_PATHWAY_SCORE_REPORT was used in old versions of the DrugMatrix software and is currently obsolete. The hypergeometric pathway impact scores are now computed at run time and cannot be mined from the database.

GENE_ACTIVITY_REPORT

Type: Fact

The GENE_ACTIVITY_REPORT contains the molecular pharmacology assay results for all of the compounds in DrugMatrix. This denormalized report contains the activities measured as % inhibition, IC50 and Ki. The same data is also available in the COMPOUND_ACTIVITY_REPORT, but this table has the gene identifiers in addition to the assay identifiers. Since there is a many-to-one mapping between genes and assays, respectively, this table contains multiple copies of some of the activity rows.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_LABEL	VARCHAR2	N	
COMPOUND_NAME	VARCHAR2	Y	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE
GENE	NUMBER	Y	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
INH_ACTIVITY	NUMBER	N	
INH_UNIT	VARCHAR2	N	
IC50_ACTIVITY	NUMBER	Y	
IC50_UNIT	VARCHAR2	Y	
KI_ACTIVITY	NUMBER	Y	
KI_UNIT	VARCHAR2	Y	
ASSAY	NUMBER	N	FK->ASSAY_REPORT.ASSAY
ASSAY_NAME	VARCHAR2	N	

GENE_ANNOTATION_REPORT

Type: Fact

The GENE_ANNOTATION_REPORT contains a mixture of annotations related to genes, including the GO annotation, NCBI identifiers and synonyms. The CATEGORY column contains a term for the annotation type.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	Y	FK->TARGET_REPORT.GENE
CATEGORY	VARCHAR2	Y	
VALUE	VARCHAR2	Y	

GENE_PATHWAY_REPORT

Type: Fact

The GENE_PATHWAY_REPORT is a connecting table between GENEs and PATHWAYs, mapping between these dimensions with no additional information.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	N	
PATHWAY	NUMBER	N	FK->PATHWAY_REPORT.PATHWAY
PATHWAY_NAME	VARCHAR2	Y	

GENE_TISSUE_REPORT

Type: Fact
The GENE_TISSUE_REPORT is a utility table used by the DrugMatrix application to determine the order of displaying tissues in the gene expression tissue menu. It does not have much utility for data mining.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
TISSUE	NUMBER	N	
TISSUE_NAME	VARCHAR2	N	
TISSUE_PRIORITY	NUMBER	N	

GO_REPORT

Type: Fact
The GO_REPORT contains all of the Gene Ontology terms associated with each gene in DrugMatrix. The CATEGORY column lists the GO hierarchy to which the term is associated.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	N	
CATEGORY	VARCHAR2	N	
TERM	VARCHAR2	N	
PUBMED	NUMBER	Y	
GOID	VARCHAR2	N	

HYBRIDIZATION_AND_IMAGE_REPORT

Type: Dimension, Minor
The HYBRIDIZATION_AND_IMAGE_REPORT is the dimension table for the hybridization minor domain and contains linking information between expression EXPERIMENTs and HYBRIDIZATIONs (arrays) that allow one to mine the array level data. The HYBRIDIZATION column is reference from the RAW_DATA_REPORT and the EXPERIMENT column allows retrieval of treatment condition information. The TYPE columns lists CTL or EXP to annotate each array as an untreated control or compound treated sample, respectively. The VALID column is used to flag "passed" (Y) vs. "failed" (N) arrays.

COLUMN_NAME	DATA_TYPE	NULLABLE	FK->HYBRIDIZATION_AND_IMAGE_REPORT.HYBRIDIZATION
HYBRIDIZATION	NUMBER	N	
TYPE	VARCHAR2	Y	
VALID	CHAR	Y	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT

HYBRIDIZATION_DATE_REPORT

Type: Fact (New)
The HYBRIDIZATION_DATE_REPORT contains laboratory process dates that can be used as a proxy for protocol steps and protocol changes when looking for block effects in the DrugMatrix data. These data are not displayed in the user interface but are made available for those wishing to do sophisticated analyses.

COLUMN_NAME	DATA_TYPE	NULLABLE	
HYBRIDIZATION	NUMBER	N	FK->HYBRIDIZATION_AND_IMAGE_REPORT.HYBRIDIZATION
HYB_DATE	DATE	N	
TARGET_PREP_DATE	DATE	N	
RNA_EXTRACTION_DATE	DATE	N	
EXPERIMENT_DATE	DATE	Y	

ICONIX_USERS

Type: Utility table
This table contains the DrugMatrix user account information. The information in this table is used to authenticate users when they log into DrugMatrix if you are using the default authentication. If you are using the LDAP authentication option, this table is not used. In the default implementation, this table contains clear-text passwords. Another installation option is to use unix encrypted passwords, but you will need to populate this table with the encrypted strings yourself.

COLUMN_NAME	DATA_TYPE	NULLABLE	DESCRIPTION
ID	NUMBER	N	PK
COMPANYNAME	VARCHAR2	N	
NAME	VARCHAR2	N	UK
PASSWORD	VARCHAR2	N	
ACCESSLEVEL	NUMBER	N	Must be 1 for access, 0 for no access
EMAIL	VARCHAR2	Y	

INDUCED_TRANSCRIPT_REPORT

Type: Fact
The INDUCED_TRANSCRIPT_REPORT contains all of the SIGNIFICANT UP-REGULATED (P<=0.02) log ratio gene expression data in DrugMatrix. It is based on the SIGNIFICANT_TRANSCRIPT_REPORT and only differs in that it contains only the induced rows.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
EXPRESSON	NUMBER	N	FK->EXPRESSON_REPORT.EXPRESSON
DIFFERENTNESS	NUMBER	Y	
LOG_RATIO	NUMBER	Y	
STDEV_OF_LOG_RATIO	NUMBER	Y	
SCORE	NUMBER	Y	
INTENSITY	NUMBER	Y	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
TIME	NUMBER	N	
TIME_UNIT	CHAR	Y	
DOSE	NUMBER	Y	
DOSE_UNIT	VARCHAR2	Y	
ORGANISM_DESCRIPTION	VARCHAR2	N	
CHIP_NAME	VARCHAR2	N	
TYPE	VARCHAR2	N	
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
TISSUE	NUMBER	N	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE

LOT_REPORT

Type: Fact
The LOT_REPORT contains a listing of all of the chip lots used to generate the microarray expression data in DrugMatrix.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ID	NUMBER	N	PK
NAME	VARCHAR2	N	
MAP	NUMBER	N	FK->MAP_REPORT.ID

MAP_REPORT

Type: Dimension, Minor
The MAP_REPORT contains a listing of all of the chip maps (designs) used to generate the microarray expression data in DrugMatrix.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ID	NUMBER	N	PK
NAME	VARCHAR2	N	
TYPE	VARCHAR2	N	
LAYOUT	VARCHAR2	Y	
SHORT_NAME	VARCHAR2	N	

MAP_SPOT_EXPRESSON_REPORT

Type: Fact
The MAP_SPOT_EXPRESSON_REPORT is an accessory table that maps between the PROBE, EXPRESSON and GENE elements of the genomic data model behind DrugMatrix. It is redundant with the EXPRESSON_REPORT but contains some additional information that describes each microarray spot.

COLUMN_NAME	DATA_TYPE	NULLABLE	
MAP_SPOT	NUMBER	N	
MAP	NUMBER	N	FK->MAP_REPORT.ID
CHIP_NAME	VARCHAR2	N	
X_COORD	NUMBER	Y	Grid location of spot
Y_COORD	NUMBER	Y	Grid location of spot
TYPE	CHAR	N	
PROBE	NUMBER	Y	FK->EXPRESSON_REPORT.PROBE
PROBE_NAME	VARCHAR2	Y	
GENE	NUMBER	Y	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
EXPRESSON	NUMBER	Y	FK->EXPRESSON_REPORT.EXPRESSON

PATHWAY_COMPOUND_REPORT

Type: Fact
This table links pathways to compounds and provides information on the names of each domain object.

COLUMN_NAME	DATA_TYPE	NULLABLE	DESCRIPTION
PATHWAY_COMPOUND	NUMBER	N	PK
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE
PATHWAY	NUMBER	Y	FK->PATHWAY_REPORT.PATHWAY
PATHWAY_NAME	VARCHAR2	Y	
X_COORD	NUMBER	Y	Pixel location on image
Y_COORD	NUMBER	Y	Pixel location on image

PATHWAY_GENE_REPORT

Type: Fact (Deprecated)
This table is deprecated and will be restructured to match the PATHWAY_TARGET_REPORT.

COLUMN_NAME	DATA_TYPE	NULLABLE	DESCRIPTION
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
PATHWAY	NUMBER	N	FK->PATHWAY_REPORT.PATHWAY
PATHWAY_NAME	VARCHAR2	Y	

PATHWAY_LITERATURE_REPORT

Type: Fact
This table links pathways to literature references that support the pathway.

COLUMN_NAME	DATA_TYPE	NULLABLE	DESCRIPTION
PATHWAY	NUMBER	N	FK->PATHWAY_REPORT.PATHWAY
PATHWAY_NAME	VARCHAR2	Y	
TITLE	VARCHAR2	N	
PUBMED	NUMBER	Y	

PATHWAY_PATHWAY_REPORT

Type: Fact
The PATHWAY_PATHWAY_REPORT is a connecting fact table that links related pathways.

COLUMN_NAME	DATA_TYPE	NULLABLE	DESCRIPTION
PARENT_PATHWAY	NUMBER	N	FK->PATHWAY_REPORT.PATHWAY
PARENT_PATHWAY_NAME	VARCHAR2	Y	
LINKED_PATHWAY	NUMBER	N	FK->PATHWAY_REPORT.PATHWAY
LINKED_PATHWAY_NAME	VARCHAR2	Y	

PATHWAY_REPORT

Type: Dimension, Major
As the dimension table for the Pathway Domain, the PATHWAY_REPORT contains primarily the pathway name.

COLUMN_NAME	DATA_TYPE	NULLABLE	DESCRIPTION
PATHWAY	NUMBER	N	PK
PATHWAY_NAME	VARCHAR2	Y	
SPECIES	NUMBER	Y	INTERNAL USE
IMAGE_FILE_NAME	VARCHAR2	Y	INTERNAL USE

PATHWAY_TARGET_REPORT

Type: Fact (Deprecated)
This table is deprecated and will be renamed to the PATHWAY_GENE_REPORT.

COLUMN_NAME	DATA_TYPE	NULLABLE	DESCRIPTION
PATHWAY_TARGET	NUMBER	N	PK
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	N	
PATHWAY	NUMBER	N	FK->PATHWAY_REPORT.PATHWAY
PATHWAY_NAME	VARCHAR2	Y	
X_COORD	NUMBER	N	Pixel location on image
Y_COORD	NUMBER	N	Pixel location on Image

PATHWAY_TISSUE_REPORT

Type: Fact
The PATHWAY_TISSUE_REPORT is a connecting fact table that associates pathways with the tissues they are relevant to.

COLUMN_NAME	DATA_TYPE	NULLABLE	DESCRIPTION
PATHWAY	NUMBER	N	FK->PATHWAY_REPORT.PATHWAY
PATHWAY_NAME	VARCHAR2	Y	
TISSUE	NUMBER	Y	
TISSUE_NAME	VARCHAR2	N	

PHARMACOLOGY_REPORT

Type: Fact
This table contains all of the curated pharmacology data from animal toxicology, clinical (human) findings and *in vivo* assays.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
PUBMED	NUMBER	Y	
CATEGORY	VARCHAR2	N	
SHORT_CATEGORY	VARCHAR2	Y	
TYPE	VARCHAR2	N	
VALUE	VARCHAR2	Y	
DESCRIPTION	VARCHAR2	Y	
REFERENCE	VARCHAR2	Y	

PLATFORM_FACTOR

Type: Fact
This table contains a cross-platform mapping scalar factor that can be applied to log ratio data to make different array platforms more compatible with the DrugMatrix Codelink data. Currently this table is not used but it is present for future compatibility.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ID	NUMBER	N	
PROBE	NUMBER	N	FK->EXPRESSON_REPORT.PROBE
FACTOR	NUMBER	N	

PROBE_CONSTANCY_REPORT

Type: Fact (Beta)
This table contains precomputed data describing how consistently various probes are expressed over sets of related experiments. This table supports a beta-quality feature that is in development and is not intended for use, because it may change significantly.

COLUMN_NAME	DATA_TYPE	NULLABLE
PROBE	NUMBER	N
LABEL_NAME	VARCHAR2	N
LABEL_TYPE	VARCHAR2	N
TISSUE	NUMBER	N
TISSUE_NAME	VARCHAR2	N
INDU	NUMBER	N
REPR	NUMBER	N
PERT	NUMBER	N
DIST	NUMBER	N
AVLR	NUMBER	N

PROBE_VARIANCE_REPORT

Type: Fact
This table contains precomputed global variance values for probes on the chips found in DrugMatrix. These data are used to support a Bayesian method to adjust measured probe variances based on prior knowledge, which is part of the standard error computation used when computing log ratio values for experiments in DrugMatrix and the Staging database system.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ID	NUMBER	N	PK
CHIP_NAME	VARCHAR2	N	
TISSUE_NAME	VARCHAR2	N	
TISSUE	NUMBER	N	
STRAIN_NAME	VARCHAR2	N	
STRAIN	NUMBER	N	
TYPE	VARCHAR2	N	
PROBE	NUMBER	N	
NO	NUMBER	Y	
VARO	NUMBER	Y	

RAW_DATA_REPORT

Type: Fact
This table contains all of the "raw" gene expression data used in DrugMatrix. Each row of the table represents the measurements from a single spot on a single array. Since this is the largest table in the data warehouse (well over 100 million rows), any SQL queries that access this table should be carefully constructed and optimized for performance.

COLUMN_NAME	DATA_TYPE	NULLABLE	
HYBRIDIZATION	NUMBER(8)	N	FK->HYBRIDIZATION_AND_IMAGE_REPORT.HYBRIDIZATION
PROBE	NUMBER(8)	N	FK->EXPRESSON_REPORT.PROBE
PROBE_NAME	VARCHAR2(25)	N	
NORMALIZED	NUMBER(11,3)	Y	Normalized Signal
SIGNAL	NUMBER(11,3)	Y	Raw Signal

REPLICATE_COMPOUND_REPORT

Type: Fact (Deprecated)
This table contains a mapping from expression experiments (A.K.A. replicates) to compounds.
This information is almost completely overlapping with the data in the
EXPERIMENT_CONDITION_REPORT, so use of this table is not necessary. The table basically
functions to drive a sub-component of the user interface.

COLUMN_NAME	DATA_TYPE	NULLABLE	
REPLICATE	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
REPLICATE_NAME	VARCHAR2	N	
COMPOUND	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
ICXNUMBER	VARCHAR2	N	
IMAGE_CREATED	CHAR	N	
MAP	NUMBER	N	FK->MAP_REPORT.ID
FILETYPE	VARCHAR2	N	

REPLICATE_SET_LOOKUP

Type: Fact
This utility table provides a mapping between HYBRIDIZATION, REPLICATE_SET and
EXPERIMENT, that describes the grouping of arrays into treatment sets and control sets.

COLUMN_NAME	DATA_TYPE	NULLABLE	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
REPLICATE_SET	NUMBER	N	
HYBRIDIZATION	NUMBER	N	FK->HYBRIDIZATION_AND_IMAGE_REPORT.HYBRIDIZATION
EXPERIMENTAL_TYPE	CHAR	N	
LABEL	VARCHAR2	N	

REPLICATE_SIMILARITY_REPORT

Type: Fact

This table lists pairwise similarities in expression profiles between each expression experiment in DrugMatrix and every other experiment. For each query experiment, only the top 25 scored experiments are saved. The QUERY_EXPERIMENT is the item being reported on in the user interface (Similar Expressions Tab) and the "EXPERIMENT" is the list of other experiments that have similarity to the QUERY.

COLUMN_NAME	DATA_TYPE	NULLABLE	
QUERY_COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
QUERY_COMPOUND_NAME	VARCHAR2	Y	
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE
QUERY_EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
QUERY_EXPERIMENT_NAME	VARCHAR2	Y	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
TIME	NUMBER	N	
TIME_UNIT	CHAR	Y	
DOSE	NUMBER	Y	
DOSE_UNIT	VARCHAR2	Y	
ORGANISM_DESCRIPTION	VARCHAR2	N	
CHIP_NAME	VARCHAR2	N	
PERCENT_INDUCED	NUMBER	Y	
PERCENT_REPRESSED	NUMBER	Y	
PERCENT_PERTURBED	NUMBER	Y	
TYPE	VARCHAR2	N	
SIMILARITY	NUMBER	N	Pearson's Correlation across all Genes

REPRESSED_TRANSCRIPT_REPORT

Type: Fact
The REPRESSED_TRANSCRIPT_REPORT contains all of the SIGNIFICANT DOWN-REGULATED (P<=0.02) log ratio gene expression data in DrugMatrix. It is based on the SIGNIFICANT_TRANSCRIPT_REPORT and only differs in that it contains only the repressed rows.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
EXPRESSON	NUMBER	N	FK->EXPRESSON_REPORT.EXPRESSON
DIFFERENTNESS	NUMBER	Y	
LOG_RATIO	NUMBER	Y	
STDEV_OF_LOG_RATIO	NUMBER	Y	
SCORE	NUMBER	Y	
INTENSITY	NUMBER	Y	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
TIME	NUMBER	N	
TIME_UNIT	CHAR	Y	
DOSE	NUMBER	Y	
DOSE_UNIT	VARCHAR2	Y	
ORGANISM_DESCRIPTION	VARCHAR2	N	
CHIP_NAME	VARCHAR2	N	
TYPE	VARCHAR2	N	
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
TISSUE	NUMBER	N	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE

SAMPLE_ANIMAL_LOOKUP

Type: Fact (Deleted)
This table is deleted as of version 2.42 and should not be used. This information is no longer needed.

SAMPLE_ASSAY_REPORT

Type: Fact (Deleted)
This table is deleted as of version 2.42 and should not be used. The same information is available in a non-pivoted form in the SAMPLE_ASSAY_REPORT_SC table.

SAMPLE_ASSAY_REPORT_SC

Type: Fact
This table contains the clinical and hematology assay results for each assay for each individual animal used to build the averages in the BLOOD_REPORT.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ANIMAL_ID	VARCHAR2	Y	FK->ANIMAL_ANNOTATION_REPORT.ANIMAL_ID
SAMPLE_ID	NUMBER	Y	FK->ANIMAL_ANNOTATION_REPORT.SAMPLE_ID
IDENTIFIER	VARCHAR2	Y	
ASSAY_NAME	VARCHAR2	Y	
VALUE	NUMBER	Y	

SAMPLE_HISTOPATH_REPORT

Type: Fact, Pivot (Deleted)
This table is deleted as of version 2.42 and should not be used. The same information is available in a non-pivoted form in the SAMPLE_HISTOPATH_REPORT_SC table.

SAMPLE_HISTOPATH_REPORT_SC

Type: Fact
This table contains the pathology results for each assay for each individual animal as expressed by a name and a severity, as displayed in the pathlab report.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ANIMAL_ID	VARCHAR2	Y	FK->ANIMAL_ANNOTATION_REPORT.ANIMAL_ID
SAMPLE_ID	NUMBER	Y	FK->ANIMAL_ANNOTATION_REPORT.SAMPLE_ID
IDENTIFIER	VARCHAR2	Y	
HISTOPATH_NAME	VARCHAR2	Y	
SEVERITY	NUMBER	Y	

SAMPLE_WEIGHT_REPORT

Type: Fact, Pivot (Deleted)
This table is deleted as of version 2.42 and should not be used. The same information is available in a non-pivoted form in the SAMPLE_WEIGHT_REPORT_SC table.

SAMPLE_WEIGHT_REPORT_SC

Type: Fact
This table contains the animal weight results for each individual animal used to build the pathlab report.

COLUMN_NAME	DATA_TYPE	NULLABLE	
ANIMAL_ID	VARCHAR2	Y	FK->ANIMAL_ANNOTATION_REPORT.ANIMAL_ID
SAMPLE_ID	NUMBER	Y	FK->ANIMAL_ANNOTATION_REPORT.SAMPLE_ID
IDENTIFIER	VARCHAR2	Y	
STUDY	NUMBER	Y	
TERM	VARCHAR2	Y	
WEIGHT	NUMBER	Y	

SIGNATURE_CMPD_SCORE_REPORT

Type: Fact (New)
This table contains precomputed signature match scores for all experiments by compound scored by all signatures of the tissue and chip platform appropriate to that experiment, using the scoring method appropriate to the signature algorithm type.

COLUMN_NAME	DATA_TYPE	NULLABLE	
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
SIGNATURE	NUMBER	N	FK->SIGNATURE_REPORT.SIGNATURE
MAX_SCORE	NUMBER	N	
SIGNATURE_TISSUE	NUMBER	N	
TISSUE_NAME	VARCHAR2	N	
ALGORITHM	VARCHAR2	Y	
PLATFORM	VARCHAR2	Y	
SCORE_METHOD	NUMBER	Y	
SCORE_METHOD_NAME	VARCHAR2	N	
CATEGORY	VARCHAR2	Y	

SIGNATURE_EXP_SCORE

Type: Data Staging

This table is used as a temporary repository of signature scores during the post-warehouse computations that are part of the final stage of a database build. This table is not intended for use. The values are copied to the SIGNATURE_EXP_SCORE_REPORT when the computations are complete and that report should be used.

SIGNATURE_EXP_SCORE_REPORT

Type: Fact (New)

This table contains precomputed signature match scores for all experiments scored by all signatures of the tissue and chip platform appropriate to that experiment, using the scoring method appropriate to the signature algorithm type.

COLUMN_NAME	DATA_TYPE	NULLABLE	
SIGNATURE	NUMBER	N	FK->SIGNATURE_REPORT.SIGNATURE
SIGNATURE_NAME	VARCHAR2	Y	
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
SCORE	NUMBER	N	
SCORE_METHOD	NUMBER	Y	
SIGNATURE_TISSUE	NUMBER	N	
SIGNATURE_TISSUE_NAME	VARCHAR2	N	
ALGORITHM	NUMBER	Y	
PLATFORM	NUMBER	Y	

SIGNATURE_REPORT

Type: Dimension, Minor

This table represents the basic domain information for the signature domain.

COLUMN_NAME	DATA_TYPE	NULLABLE	
SIGNATURE	NUMBER	N	PK
SIGNATURE_NAME	VARCHAR2	Y	
SIGNATURE_TYPE	VARCHAR2	Y	
IDENTIFIER	VARCHAR2	N	
ALGORITHM	VARCHAR2	Y	
PLATFORM	VARCHAR2	Y	
TISSUE	NUMBER	N	
TISSUE_NAME	VARCHAR2	N	
LABEL_DESCRIPTION	VARCHAR2	Y	
TOTAL_SIZE	NUMBER	Y	
POSITIVE_CLASS_SIZE	NUMBER	Y	
NEGATIVE_CLASS_SIZE	NUMBER	Y	
EXCLUDED_CLASS_SIZE	NUMBER	Y	
UNIVERSE_DESCRIPTION	VARCHAR2	Y	
POSITIVE_CLASS_DESCRIPTION	VARCHAR2	Y	
NEGATIVE_CLASS_DESCRIPTION	VARCHAR2	Y	
EXCLUDED_CLASS_DESCRIPTION	VARCHAR2	Y	
AVERAGE_LOG_ODDS	NUMBER	Y	
AVERAGE_TRUE_POSITIVE_PERCENT	NUMBER	Y	
AVERAGE_TRUE_NEGATIVE_PERCENT	NUMBER	Y	
CATEGORY	VARCHAR2	Y	

SIGNATURE_SCORE_REPORT

Type: Fact (Deleted)
This table has been replaced by the SIGNATURE_CMPD_SCORE_REPORT and the
SIGNATURE_EXP_SCORE_REPORT.

SIGNIFICANT_TRANSCRIPT_REPORT

Type: Fact
The SIGNIFICANT_TRANSCRIPT_REPORT contains all of the SIGNIFICANT (P<=0.02) log ratio
gene expression data in DrugMatrix. It is based on the ALL_TRANSCRIPT_REPORT and only
differs in that it contains only the significant rows.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
EXPRESSON	NUMBER	N	FK->EXPRESSON_REPORT.EXPRESSON
DIFFERENTNESS	NUMBER	Y	
LOG_RATIO	NUMBER	Y	
STDEV_OF_LOG_RATIO	NUMBER	Y	
SCORE	NUMBER	Y	
INTENSITY	NUMBER	Y	
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
TIME	NUMBER	N	
TIME_UNIT	CHAR	Y	
DOSE	NUMBER	Y	
DOSE_UNIT	VARCHAR2	Y	
ORGANISM_DESCRIPTION	VARCHAR2	N	
CHIP_NAME	VARCHAR2	N	
TYPE	VARCHAR2	N	
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
TISSUE	NUMBER	N	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE

SIGNIF_SIG_EXP_SCORE_REPORT

Type: Fact (New)
This table contains precomputed signature match scores for all experiments scored by all signatures of the tissue and chip platform appropriate to that experiment, using the scoring method appropriate to the signature algorithm type. This table is similar to the SIGNATURE_EXP_SCORE_REPORT, but the structure is slightly modified and the data has been pre-filtered for scalar product scores above 0.

COLUMN_NAME	DATA_TYPE	NULLABLE	
SIGNATURE	NUMBER	N	FK->SIGNATURE_REPORT.SIGNATURE
SIGNATURE_NAME	VARCHAR2	Y	
COMPOUND	NUMBER	Y	FK->COMPOUND_REPORT.COMPOUND
COMPOUND_NAME	VARCHAR2	Y	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE
EXPERIMENT	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
EXPERIMENT_NAME	VARCHAR2	Y	
SP_SCORE	NUMBER	N	
SPS_SCORE	NUMBER	N	
SIGNATURE_TISSUE	NUMBER	N	
SIGNATURE_TISSUE_NAME	VARCHAR2	N	
ALGORITHM	NUMBER	Y	
PLATFORM	NUMBER	Y	
CATEGORY	VARCHAR2	Y	

SIG_SIG_SIMILARITY_REPORT

Type: Fact (New)
This table contains precomputed similarity scores between the pattern of hit responses between each pair of signatures across all relevant experiments. The Score is a Pearson's similarity score.

COLUMN_NAME	DATA_TYPE	NULLABLE	
SIGNATURE	NUMBER	N	FK->SIGNATURE_REPORT.SIGNATURE
SIGNATURE_NAME	VARCHAR2	Y	
SCORE	NUMBER	N	
QUERY_SIGNATURE	NUMBER	N	FK->SIGNATURE_REPORT.SIGNATURE
QUERY_SIGNATURE_NAME	VARCHAR2	Y	

SIMILARITY

Type: Data Staging
This table is used as a temporary repository of experiment similarity scores during the post-warehouse computations that are part of the final stage of a database build. This table is not intended for use. The values are copied to the REPLICATE_SIMILARITY_REPORT when the computations are complete and that report should be used.

SIMPLE_ASSAY_LOOKUP

Type: Fact
The SIMPLE_ASSAY_LOOKUP table drives the ASSAY domain in the simple search tool in the user interface, where the TERM column is the lookup column and the ID and NAME columns are the result.

COLUMN_NAME	DATA_TYPE	NULLABLE	
TERM	VARCHAR2	N	
ID	NUMBER	N	FK->ASSAY_REPORT.ASSAY
NAME	VARCHAR2	N	

SIMPLE_COMPOUND_LOOKUP

Type: Fact

The SIMPLE_COMPOUND_LOOKUP table drives the COMPOUND domain in the simple search tool in the user interface, where the TERM column is the lookup column and the ID and NAME columns are the result.

COLUMN_NAME	DATA_TYPE	NULLABLE	
TERM	VARCHAR2	N	
ID	NUMBER	N	FK->COMPOUND_REPORT.COMPOUND
NAME	VARCHAR2	N	

SIMPLE_EXPERIMENT_LOOKUP

Type: Fact

The SIMPLE_EXPERIMENT_LOOKUP table drives the EXPRESSION EXPERIMENT domain in the simple search tool in the user interface, where the TERM column is the lookup column and the ID and NAME columns are the result.

COLUMN_NAME	DATA_TYPE	NULLABLE	
TERM	VARCHAR2	N	
ID	NUMBER	N	FK->EXPERIMENT_CONDITION_REPORT.EXPERIMENT
NAME	VARCHAR2	N	

SIMPLE_GENE_LOOKUP

Type: Fact

The SIMPLE_GENE_LOOKUP table drives the GENE domain in the simple search tool in the user interface, where the TERM column is the lookup column and the ID and NAME columns are the result.

COLUMN_NAME	DATA_TYPE	NULLABLE	
TERM	VARCHAR2	N	
ID	NUMBER	N	FK->TARGET_REPORT.GENE
NAME	VARCHAR2	N	

SIMPLE_PATHWAY_LOOKUP

Type: Fact

The SIMPLE_PATHWAY_LOOKUP table drives the PATHWAY domain in the simple search tool in the user interface, where the TERM column is the lookup column and the ID and NAME columns are the result.

COLUMN_NAME	DATA_TYPE	NULLABLE	
TERM	VARCHAR2	N	
ID	NUMBER	N	FK->PATHWAY_REPORT.PATHWAY
NAME	VARCHAR2	N	

SIMPLE_SIGNATURE_LOOKUP

Type: Fact
The SIMPLE_SIGNATURE_LOOKUP table drives the SIGNATURE domain in the simple search tool in the user interface, where the TERM column is the lookup column and the ID and NAME columns are the result.

COLUMN_NAME	DATA_TYPE	NULLABLE	
TERM	VARCHAR2	N	
ID	NUMBER	N	FK->SIGNATURE_REPORT.SIGNATURE
NAME	VARCHAR2	N	

SIMPLE_STUDY_LOOKUP

Type: Fact
The SIMPLE_STUDY_LOOKUP table drives the STUDY domain in the simple search tool in the user interface, where the TERM column is the lookup column and the ID and NAME columns are the result.

COLUMN_NAME	DATA_TYPE	NULLABLE	
TERM	VARCHAR2	N	
ID	NUMBER	N	FK->STUDY_REPORT.STUDY
NAME	VARCHAR2	N	

STUDY_IMAGEDATA

Type: Fact
The STUDY_IMAGEDATA table is a utility table for supporting the dose and time course response charts. It is probably of little value for data mining.

COLUMN_NAME	DATA_TYPE	NULLABLE	
STUDY	NUMBER	N	FK->STUDY_REPORT.STUDY
STUDY_TYPE	NUMBER	N	
XML	VARCHAR2	Y	
VALUE	VARCHAR2	Y	
TISSUE	NUMBER	N	
TISSUE_NAME	VARCHAR2	N	

STUDY_REPORT

Type: Dimension, Major
As the dimension table for the Study Domain, the STUDY_REPORT contains some basic information about each expression study (a group of various treatment, experiments, that were performed on the same study date in the same group of animals), such as the name, description, compound, compound salt form, compound purity, microarray technology, organism or strain, and fluorescent label(s) used.

COLUMN_NAME	DATA_TYPE	NULLABLE	
STUDY	NUMBER	N	
STUDY_NAME	VARCHAR2	Y	
STUDY_DESCRIPTION	VARCHAR2	N	
COMPOUND	NUMBER	Y	
COMPOUND_NAME	VARCHAR2	Y	
MOL_STRUCTURE_2D	NUMBER	Y	INTERNAL USE
SALT_FORM	VARCHAR2	Y	
PURITY	VARCHAR2	Y	
ARRAY_TECHNOLOGY	VARCHAR2	N	
ORGANISM	VARCHAR2	N	
ORGANISM_DESCRIPTION	VARCHAR2	N	
CTLLABEL	VARCHAR2	N	
EXPLABEL	VARCHAR2	N	

TARGET_ACCESSION_REPORT

Type: Fact
The TARGET_ACCESSION_REPORT contains external genomic references for genes and probes in DrugMatrix.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	N	
PROBE	NUMBER	N	FK->EXPRESSON_REPORT.PROBE
REFERENCE	VARCHAR2	N	
DATABASE	VARCHAR2	N	
TYPE	VARCHAR2	N	

TARGET_EXPRESSION_SIMILARITY

Type: Fact
This table lists pairwise similarities in expression profiles between each gene in DrugMatrix and every other gene. For each query gene, only the top 25 scored genes are saved. The QUERY_GENE is the item being reported on in the user interface (Similar Expressions Tab) and the "GENE" is the list of other genes that have similarity to the QUERY.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	N	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	Y	
METHOD	VARCHAR2	N	
SCORE	NUMBER	N	Pearson's Correlation across all Experiments
SPECIES	VARCHAR2	Y	
QUERY_GENE	NUMBER	N	FK->TARGET_REPORT.GENE
QUERY_GENE_NAME	VARCHAR2	Y	

TARGET_REPORT

Type: Dimension, Major
As the dimension table for the Gene Domain, the TARGET_REPORT contains some basic information about each gene, such as the name, symbol and a short description.
For more gene information, also see TARGET_SYNONYM, TARGET_ACCESSION_REPORT and GO_REPORT.

COLUMN_NAME	DATA_TYPE	NULLABLE	
GENE	NUMBER	N	PK
GENE_NAME	VARCHAR2	N	
SYMBOL	VARCHAR2	Y	
DESCRIPTION	VARCHAR2	Y	

TARGET_SYNONYM

Type: Fact, 1D
The TARGET_SYNONYM contains one-or-more synonyms for each gene, as well as the Gene identifier and official name.

COLUMN_NAME	DATA_TYPE	NULLABLE	
CATEGORY	CHAR	Y	
SYNONYM_NAME	VARCHAR2	N	
GENE	NUMBER	Y	FK->TARGET_REPORT.GENE
GENE_NAME	VARCHAR2	N	

Appendix A: Schema Diagram