

Oracle® Health Sciences Omics Data Bank

Release Content Document

Release 1.0.1

E27537-02

April 2012

1 Introduction

1.1 Purpose of Document

The Release Content Document (RCD) communicates information about Oracle Health Sciences Omics Data Bank.

1.2 Release Overview

Oracle Health Sciences Omics Data Bank is a new product. It provides users a cross-platform data model to store and query omics information.

2 Product Overview

Oracle Health Sciences Omics Data Bank (ODB) is a new product from Oracle Health Sciences. ODB 1.0.1 is a cross-platform data model for omics information. It enables users to ask biologically meaningful questions. Moreover, it is fully integrated with the Oracle Health Sciences Cohort Explorer (OHSCE), bringing both phenotypic and genotypic data to researchers and clinicians.

The data model supports reference omics information and annotation, as well as result data from each individual sample. ODB 1.0.1 is packaged with a set of data loaders for reference and result data, thereby reducing your adoption barrier. It also includes Oracle SecureFile as an optional file storage mechanism.

ODB 1.0.1 includes the following components:

Data Model

- Reference tables for biological context
- Results tables for individual sample data
- Multiple omics datatypes
 - Simple variation including SNV, insertion, and deletion
 - Gene expression
 - Copy number variation

Data Loaders

- Reference data loaders for
 - Ensembl (for example, genome definition, annotations from dbSNP and COSMIC, and so on)
 - Swiss-Prot
 - HUGO
 - Pathwaycommons
- Result data loaders for
 - Variant call format (VCF v4.1)
 - Complete Genomics variant format (MasterVar)
 - Tab-delimited format for expression data
 - Mutation annotation format (MAF)

SecureFile (optional)

- Oracle SecureFile as an option for all data files that should be kept in a secure environment

3 Features

Cross-Platform Omics Data Model

ODB supports omics data management and analysis regardless of the scientific approach (for example, next generation sequencing, gene expression, and so on) and technology platform (for example, Illumina, CompleteGenomics, and so on).

Multiple Species

ODB can model reference DNA, gene, and protein information from multiple species.

Cross Reference Omics Information

ODB can handle cross reference information for DNA source, gene, gene component, variant, protein, and protein component.

Integration with Clinical Data

ODB is fully compatible with the Cohort Data Mart (part of OHSCE v1.0.0.1) to let end users integrate both clinical and omics information in their cohort selection.

Data Loaders

Easy-to-use command line data loaders with comprehensive instructions.

No-call Region

ODB differentiates no-call regions from those that are identical to the reference genome in the result data.

Links with Original Result Files

ODB links derived genomic results with their corresponding files either through Oracle SecureFile or a regular file directory.

Scalability

Result tables are partitioned by study, and several reference tables are partitioned by DNA_SOURCE.

Terminology

Table 1 Terminology

Term	Definition
COSMIC	Catalogue of Somatic Mutations in Cancer
HUGO	HUGO Gene Nomenclature Committee
MAF	Mutation Annotation Format
ODB	Omics Data Bank
OHSCE	Oracle Health Sciences Cohort Explorer
SNV	Single-nucleotide Variation
TRC	Translation Research Center
VCF	Variant Call Format

4 Disclaimer Regarding Third Party Data

Oracle makes no express or implied warranty, including but not limited to warranties regarding the accuracy, completeness, merchantability, or fitness for a particular purpose, with respect to third party data loaded into this application or the results of any functions of the application using such data. It may be used for information purposes only, and no medical, clinical or other health related decisions may be based upon such results. You are solely responsible for your use of the third party data, including your right to use the data for your purposes.

5 Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at

<http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc>.

Access to Oracle Support

Oracle customers have access to electronic support through My Oracle Support. For information, visit

<http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info> or visit

<http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs> if you are hearing impaired.

Oracle Health Sciences Omics Data Bank Release Content Document, Release 1.0.1
E27537-02

Copyright © 2012 Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate,

broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation shall be subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License (December 2007). Oracle USA, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

This software and documentation may provide access to or information on content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services.