# List of the changes made in cBioPortal source code to base MSSQL compliant

## Configuration updates

1. [apache-tomcat]/conf/context.xml

<Resource auth="Container" driverClassName="com.microsoft.sqlserver.jdbc.SQLServerDriver" maxActive="100" maxIdle="30" maxWait="10000" name="jdbc\_mssql/cbioportal" password="somepassword" testOnBorrow="true" type="javax.sql.DataSource" url="jdbc:sqlserver://localhost:1433;DatabaseName=cbioportal" username="someuser" validationQuery="SELECT 1"/>

1. src/main/resource/portal.properties
   1. Replace database information.

db.vendor=mssql

db.user=someuser

db.password=somepassword

db.host=localhost

db.portal\_db\_name=cbioportal

db.driver=com.microsoft.sqlserver.jdbc.SQLServerDriver

db.connection\_string=jdbc:sqlserver://127.0.0.1:1433;DatabaseName=

db.url=jdbc:sqlserver://127.0.0.1:1433;DatabaseName=cbioportal

db.tomcat\_resource\_name=jdbc\_mssql/cbioportal

1. src/main/etc/jetty-cbioportal.xml
   1. Updated jndi Resource and dbcp.BasicDataSource to use jdbc\_mssql/cbioportal and cbioportal\_test
2. src/main/etc/override-web.xml
   1. <res-ref-name>jdbc\_mssql/cbioportal</res-ref-name>
3. pom.xml
   1. Added com.microsoft.sqlserver jdbc depenecy.
   2. **[unit test]** Updated db.test.driver/url/username/password
4. .travis/settings.xml
   1. Put the proper password value.
5. Core
   1. core/pom.xml
6. **[unit test]** Added <dependency> for mssql in sql plugin for tests.
7. **[unit test]** commented out <sqlcommand> for set storage and session from <executions> for sql plugin.
8. **[unit test]**  updated <srcfiles> to use cgds\_fixed4MSSQL.sql.MS.CREATE.sql
   1. core /src/test/resources/applicationContext-dao.xml
      1. **[unit test]** updated dbcpDataSource bean for testing: points to SqlServer Drive and database
9. Business
   1. business/src/main/resources/applicationContext-business.xml
      1. Updated URL property of <bean id="businessDataSource">
      2. Updated ‘sqlSessionFactory’ bean: changed ‘mapper locations’ to mybatis-sqlserver
      3. Added bean 'vendorProperties'
10. Web
11. Model
12. Persistence
13. Persistence-Mybatis-Test is moved to Persistence-Mybatis/src/test
    1. **[unit test]** persistence/persistence-mybatis /src/test/resources/testContextDatabase.xml
       1. updated dbcpDataSource bean for testing: points to SqlServer Drive and test database database
       2. Added ‘vendorProperties’ bean.
       3. In ‘sqlSessionFactory’ bean set the mapperLocations to mybatis-sqlserver.
    2. **[unit test]** persistence/persistence-mybatis/pom.xml
       1. Replace com.h2database by MSSQL dependency

<dependency>

<groupId>com.microsoft.sqlserver</groupId>

<artifactId>mssql-jdbc</artifactId>

<version>6.1.0.jre8</version>

</dependency>

* + 1. Added log4j dependency.

<dependency>

<groupId>com.microsoft.sqlserver</groupId>

<artifactId>mssql-jdbc</artifactId>

<version>6.1.0.jre8</version>

</dependency>

* + 1. Added log4j.properties to /persistence/persistence-mybatis/src/test/resource

1. Service
2. Scripts
3. DB-scripts
4. security

Database Scripts Updates

NOTE: in MySQL Set SESSION SQL\_MODE = ANSI\_QUOTE is used. This requires to use single quotes for quoting literal strings in MS SQL.

### Portal

### Core

1. core/src/main/resources/db/cgds\_fixed4MSSQL.sql
2. core/src/main/resources/db/cgds\_fixed4MSSQL.sql.MS.CREATE.sql
   1. commented out 'GO' in two places in order to make the test scripts run
3. core/src/main/resources/db/cgds\_fixed4MSSQL.sql.MS.KEYS.sql
   1. Added ‘ON DELETE NO ACTION’ if there is a constraint ‘ON DELETE RESTRICT’ in cgds.sql (MySQL version).
4. core/src/main/resources/db/cgds\_fixed4MSSQL.sql.MS.LOAD.sql
   1. Created SqlServer compatible version of cgds.sql. Change some types and removed the some comments so that simplistic parser can handle it to create MSSQL DDLs.
5. core /src/main/resources/db/cgds\_insert\_test\_MSSQL.sql
   1. ‘SET IDENTITY\_INSERT [table\_name] ON/OFF’ in order to insert into a table with primary key with auto-increment.
6. **[unit test]** core /test/resources/seed\_mini.sql
   1. Reversed order of delete b/w gene\_panel and gene\_panel\_list, sample and sample\_list\_list due to the foreign key constraint
   2. Added ‘SET IDENTITY\_INSERT [table\_name] ON/OFF’ for inserting rows in a table if auto increment column exits in order to be compliant on MSSQL. Also added print statement.

### Persistence

### Persistence-MyBatis-Test

* 1. **[unit test]** persistence/persistence-mybatis-test/src/test/resources/testSql\_MSSQL.sql
     1. Added 'USE cbioportal\_test\_persistence' to use the correct test database for persistent testing.

Database Setup

### Test Database Setup

The complete unit tests for Core and Persistence-MyBatis-Test modules require two separate databases.

1. Security/user Setup
   1. Grant user ‘bulkadmin’ role from Security>Logins>ServerRoles.
   2. From each database security membership site, grant user the following privileges: db\_datareader, db\_datawriter, db\_ddladmin
2. Core module junit tests use cbio\_test db
   1. core/pom.xml
      1. include in dependency for MSSQL ‘groupId = org.codehaus.mojo’

<dependency>

<groupId>com.microsoft.sqlserver</groupId>

<artifactId>mssql-jdbc</artifactId>

<version>6.1.0.jre8</version>

</dependency>



<srcFiles>

<srcFile>src/main/resources/db/cgds\_fixed4MSSQL.sql.MS.CREATE.sql

</srcFile>

<srcFile>src/main/resources/db/cgds\_fixed4MSSQL.sql.MS.KEYS.sql

</srcFile>

<srcFile>src/test/resources/seed\_mini.sql</srcFile>

</srcFiles>

1. Persistence module junit tests use cbio\_test\_persistence
   1. In persistence/persistence-mybatis-test/src/test/resources/testContextDatabase.xml replace com.h2database by MSSQL database, add vendor properties, and set mapperLocation properly:

Note that in sqlSessionFactory bean, set mapperLocations pointing to SqlServer specific Mapper folders and databaseIdProvider. For setting mapperLocations do NOT use java package expression since there is no org.cbioportal.persistence.mybatis-sqlserver.

<bean id="dbcpDataSource" class="org.springframework.jdbc.datasource.DriverManagerDataSource">

<property name="url" value="jdbc:sqlserver://127.0.0.1:1433;DatabaseName=cbioportal\_test\_persistence;user=somename;password=somepassword”/>

</bean>

<bean id="vendorProperties" class="org.springframework.beans.factory.config.PropertiesFactoryBean">

<property name="properties">

<props>

<prop key="SQL SERVER">sqlserver</prop

<prop key="MySQL">mysql</prop>

</props>

</property>

</bean>

<bean id="databaseIdProvider" class="org.apache.ibatis.mapping.VendorDatabaseIdProvider">

<property name="properties" ref="vendorProperties"/>

</bean>

<bean id="sqlSessionFactory" class="org.mybatis.spring.SqlSessionFactoryBean">

<property name="dataSource" ref="dbcpDataSource" />

<property name="typeAliasesPackage" value="org.mskcc.cbio.portal.model"/>

<property name="typeHandlersPackage" value="org.cbioportal.persistence.mybatis.typehandler"/>

<property name="mapperLocations" value="classpath:org/cbioportal/persistence/mybatis-sqlserver/\*.xml" /> <property name="databaseIdProvider" ref="databaseIdProvider"/>

</bean>

* 1. In persistence/persistence-mybatis-test/pom.xml replace com.h2database by MSSQL dependency and add log4j dependency.

<dependency>

<groupId>com.microsoft.sqlserver</groupId>

<artifactId>mssql-jdbc</artifactId>

<version>6.1.0.jre8</version>

</dependency>

* 1. Add log4j.properties to /persistence/persistence-mybatis-test/resource

## Code updates ordered by maven module including Java and XML

### Portal

No updates

### Core

1. Database script updates
2. New Java classes added
   1. core/src/main/java/edu/jhu/u01/DBProperties.java
   2. Updated core/src/main/java/org/mskcc/cbio/portal/dao/JdbcUtil.java to use this DBProperties class.
   3. core/src/main/java/edu/jhu/u01/DBVendor.java which encapsulates database vendor which defines DBVendor enum
3. Updated Java Classes where MSSQL prepared statements are added
4. core/src/main/java/org/mskcc/cbio/portal/dao/DaoCancerStudy.java
5. core/src/main/java/org/mskcc/cbio/portal/dao/DaoClinicalAttributeMeta.java
6. core/src/main/java/org/mskcc/cbio/portal/dao/DaoClinicalData.java
7. core/src/main/java/org/mskcc/cbio/portal/dao/DaoClinicalEvent.java
8. core/src/main/java/org/mskcc/cbio/portal/dao/DaoCnaEvent.java
9. core/src/main/java/org/mskcc/cbio/portal/dao/DaoCoexpression.java
10. core/src/main/java/org/mskcc/cbio/portal/dao/DaoCopyNumberSegment.java
11. core/src/main/java/org/mskcc/cbio/portal/dao/DaoCopyNumberSegmentFile.java
12. core/src/main/java/org/mskcc/cbio/portal/dao/DaoCosmicData.java
13. core/src/main/java/org/mskcc/cbio/portal/dao/DaoDrug.java
14. core/src/main/java/org/mskcc/cbio/portal/dao/DaoDrugInteraction.java
15. core/src/main/java/org/mskcc/cbio/portal/dao/DaoGene.java
16. core/src/main/java/org/mskcc/cbio/portal/dao/DaoGeneticAlteration.java
17. core/src/main/java/org/mskcc/cbio/portal/dao/DaoGeneticProfile.java
18. core/src/main/java/org/mskcc/cbio/portal/dao/DaoGeneticProfileSamples.java
19. core/src/main/java/org/mskcc/cbio/portal/dao/DaoGistic.java
20. core/src/main/java/org/mskcc/cbio/portal/dao/DaoInteraction.java
21. core/src/main/java/org/mskcc/cbio/portal/dao/DaoMutSig.java
22. core/src/main/java/org/mskcc/cbio/portal/dao/DaoMutation.java
23. core/src/main/java/org/mskcc/cbio/portal/dao/DaoPatient.java
24. core/src/main/java/org/mskcc/cbio/portal/dao/DaoPdbUniprotResidueMapping.java
25. core/src/main/java/org/mskcc/cbio/portal/dao/DaoPfamGraphics.java
26. core/src/main/java/org/mskcc/cbio/portal/dao/DaoProteinArrayData .java
27. core/src/main/java/org/mskcc/cbio/portal/dao/DaoProteinArrayInfo.java
28. core/src/main/java/org/mskcc/cbio/portal/dao/DaoProteinArrayTarget.java
29. core/src/main/java/org/mskcc/cbio/portal/dao/DaoSample.java
30. core/src/main/java/org/mskcc/cbio/portal/dao/DaoSampleList.java
31. core/src/main/java/org/mskcc/cbio/portal/dao/DaoSampleProfile.java
32. core/src/main/java/org/mskcc/cbio/portal/dao/DaoSangerCensus.java
33. core/src/main/java/org/mskcc/cbio/portal/dao/DaoTextCache.java
34. core/src/main/java/org/mskcc/cbio/portal/dao/DaoTypeOfCancer.java
35. core/src/main/java/org/mskcc/cbio/portal/dao/DaoUser.java
36. core/src/main/java/org/mskcc/cbio/portal/dao/DaoUserAuthorities.java
37. core/src/test/java/org/mskcc/cbio/portal/scripts/TestImportClinicalData.java
38. core/src/test/java/org/mskcc/cbio/portal/scripts/TestImportExtendedMutationData.java
39. core/src/test/java/org/mskcc/cbio/portal/scripts/TestNormalizeExpressionLevels.java
40. Other Java Code Updates
41. core/src/main/java/org/mskcc/cbio/portal/network/NetworkIO.java
    1. added string trim to entrez\_gene in addNote and addDrugNode.
    2. addNode method add a node to a map whose key is entrez\_gene. And if entrez\_gene = '62232' it has trails of spaces eg., '62232 '.
    3. Therefore, in a subsequent method cannot find a node with entrez\_gene='62232' since its key in the map is '62232 '.
42. core/src/main/java/org/mskcc/cbio/portal/dao/DaoCancerStudy.java
    1. replaced IMPORT\_DATE = NOW() by IMPORT\_DATE =CURRENT\_TIMESTAMP
43. core/src/main/java/org/mskcc/cbio/portal/dao/DaoTextCache.java
    1. replaced NOW() by CURRENT\_TIMESTAMP in insert statement
    2. Updated purgeOldKeys:
       1. setting date time parameter for delete statement for mssql: datetime format from 'yyyyMMddHHmmss' to 'yyyy-MM-dd HH:mm:ss'.
       2. added MS SQL version of setting current timestamp value in a delete statement.
    3. Updated sql for getAllTypesOfCancer() in order to pass JUnit test, TestWebService: in SQL added 'order by type\_of\_cancer\_id' in order to make the order of the query results the same as in MySQL case.
44. core/src/main/java/org/mskcc/cbio/portal/dao/DaoMutation.java
    1. Replaced XML\_PATH method by SQL function dbo.group\_concat(col, max\_length) in order to avoid too many parameters passed to SQL error.
    2. //JK-FUTURE-TODO: Remove hard coded max length
45. core\src\main\java\org\mskcc\cbio\portal\model\GeneticProfile.java
    1. Added an extra setShowProfileInAnalysisTab(byte[] v) in case needed.
    2. In <select id=getSampleByStableIdAndStudyId>: PUBLIC -> [PUBLIC], STATUS->[STATUS]
46. core/src/main/java/org/mskcc/cbio/portal/dao/MySQLbulkLoader.java
    1. added instance variable "debugMode" to turn on/off some debugging related codes.
    2. added WITH (KEEPIDENTITY) condition in MS SQL bulk insert statement in order to insert a row if a value is provided for an auto-increment column.
    3. Added 2 static variables representing NULL value in a data file used in bulk insert statement for each DB vendor.
    4. Their values are "\\N" and ""(empty string) for MySQL and MS SQL, respectively
47. core/src/main/java/org/mskcc/cbio/portal/dao/DaoDrugInteraction.java
    1. 'INTERACTION\_TYPE' returned from the database since the column length is char(50) and sqlserver jdbc returns a value padded to fill up to char(50). But MySQL jdbc returns trimmed value.
    2. Add trim to setInteractionType since 'Interaction\_Type' column data type is char(50) instead of varchar(50) in the DB. In MSSQL jdbc brings char(50) length string padded by spaces. in the contrary, mysql jdbc returns trimmed value. So in order to make string comparison in the javacode, trimmed the value.
48. core/src/main/java/org/mskcc/cbio/portal/dao/DaoGeneticProfile.java
    1. The column data type of 'ShowProfileInAnalsisTab' is binary(1), and MSSQL jdbc returns 'bytes[1]', but MYSQL jdbc allows getBoolean, which fails in MSSQL jdbc. SO updated code to getBytes instead of getBoolean in extract method
49. core/src/main/java/org/mskcc/cbio/portal/dao/DaoMutation.java you
    1. Fixed issue with different handling of null value in bulk insert for MSSQL and MySQL. Created static strings ("" for MSSQL, "\\N" for MySQL) in MySQLBulkLoader, one for each DB vendor to replace null value in data file to be loaded in the DB.

### Business

1. Java Code Updates
   1. model/src/main/java/org/cbioportal/model/TypeOfCancer.java
      1. Updated setDedicatedColor since dedicatedColor column in the database is char(31) and MySQL returns trimmed value while SqlServer returns Char(31) length string value.
2. XML updates
   1. Tried to create MSSQL designated Mapper folder and to update mapper location to point to the MSSQL Mapper folder. But didn’t work out. Instead updated ProfileDataMapper.xml.
   2. business\src\main\resources\org\mskcc\cbio\portal\persistence\ProfileDataMapper.xml
      1. Replaced MySQL function NOW() by current\_timestamp for SqlServer.

### Web

### Persistence

1. Java Code Updates
   1. /persistence-mybatis/src/main/java/org/cbioportal/persistence/mybatis/MutationMyBatisRepository.java
      1. Updated method hasAlleleFrequenceData method to include

if (mutationMapper.hasAlleleFrequencyData(geneticProfileId, sampleId) == null) return false;

* + 1. in method 'getSignificantlyMutatedGenes' set the value of parameter for setting groupConcat to false since MS SQL does not support setting system parameter groupconcat as in MySQL.

1. XML updates
   1. created a separate folder for mssql MyBatis Mapper folder: persistence/persistence-mybatis/src/main/resources/org/cbioportal/persistence/mybatis-sqlserver
      1. this results in changes mapper locations in /business/src/main/resources/applicationContext-business.xml and persistence/persistence-mybatis-test/src/test/resources/testContextDatabase.xml
   2. [183] persistence\persistence-mybatis\src\main\resources\org\cbioportal\persistence\mybatis-sqlserver\ClinicalAttributeMapper.xml
      1. If there is no order by clause the result order in MSSQL may not be the same in MySQL therefore added an order by clause if both sortBy and direction are null.

<if test="sortBy == null and direction == null and projection != 'ID' and limit == null"> ORDER BY clinical\_sample.ATTR\_ID ASC</if>

* 1. persistence\persistence-mybatis\src\main\resources\org\cbioportal\persistence\mybatis-sqlserver\ClinicalDataMapper.xml
     1. If there is no order by clause the result order in MSSQL may not be the same in MySQL therefore added an order by clause if both sortBy and direction are null.

<if test="sortBy == null and direction == null and projection != 'ID' and limit == null"> ORDER BY clinical\_sample(or clinical\_patient).ATTR\_ID ASC </if>

* 1. GeneMapper.xml
  2. persistence\persistence-mybatis\src\main\resources\org\cbioportal\persistence\mybatis-sqlserver\GenePanelMapper.xml
     1. In select id= getGeneticProfileByStableId updated select statement to convert binary value into bit so that in the jdbc result set it can be treated as a Boolean: cast(genetic\_profile.SHOW\_PROFILE\_IN\_ANALYSIS\_TAB as bit)
  3. persistence\persistence-mybatis\src\main\resources\org\cbioportal\persistence\mybatis-sqlserver\GeneticProfileMapper.xml
     1. In <sql id="select"> cast(genetic\_profile.SHOW\_PROFILE\_IN\_ANALYSIS\_TAB as bit)
  4. persistence\persistence-mybatis\src\main\resources\org\cbioportal\persistence\mybatis-sqlserver\StudyMapper.xml
     1. In <sql id="select"> : PUBLIC -> [PUBLIC], STATUS->[STATUS]
     2. In sql replace "" by []
     3. In sql cast binary datatype column to tinyint in order to treat it as a boolean in java: min(cast(cancer\_study.[PUBLIC] as tinyint)
     4. Added aggregation function to all the selected columns other than the group by column
     5. Added <sql id="cancerTypeMapperSelect"> instead of referring org.cbioportal.persistence.mybatis.CancerTypeMapper.select
     6. In order to select aggregated column value for group by.
     7. In order by clause replace referring column name by column index(eg., order by 1) since column alias includes $prefix property which cannot be set in order by clause.
     8. Added <sql id="selectWithoutGroupBy"> for other \*Mapper.xml referring to <sql id=select>.
     9. In order to return date type column in 'UTC' time zone, used "AT TIME ZONE 'UTC'" which is a new feature available in SqlServer 2016. Eg., select importDate AT TIME ZONE 'UTC'
  5. persistence\persistence-mybatis\src\main\resources\org\cbioportal\persistence\mybatis-sqlserver\GeneticProfileMapper.xml
     1. Updated <include refid = "org.cbioportal.persistence.mybatis.StudyMapper.selectWithoutGroupBy"> instead of referring select.
  6. persistence\persistence-mybatis\src\main\resources\org\cbioportal\persistence\mybatis-sqlserver\SampleListMapper.xml
     1. Updated <include refid =" org.cbioportal.persistence.mybatis.StudyMapper.selectWithoutGroupBy"> instead of referring select.
  7. persistence\persistence-mybatis\src\main\resources\org\cbioportal\persistence\mybatis-sqlserver\MutationMapper.xml
     1. In sql id=getMutationsCounts
        1. Added aggregate ftn to the select list.
        2. Added cancer\_study to 'from' clause of subselect clause
     2. Fixed bug in getSignificantlyMutatedGenes: removed distinct from STUFF clause and added missing sample id constraint in STUFF clause.
     3. Added order by count(distinct SAMPLE\_ID) clause in countSamplesWithMutatedGenes, and countSamplesWithKeywords in order to match the order of query results as in MySQL
     4. Rewrote getGenesOfMutations: replaced distinct A by group by A and order by count(A) in order to match the order of query results as in MySQL

### Persistence-MyBatis-Test

1. Java Code Updates
   1. persistence/persistence-mybatis-test/src/test/java/org/cbioportal/persistence/mybatis/ClinicalDataMyBatisRepositoryTest.java
      1. Commented out fetchClinicalDataNullAttributeSummaryProjection() and added fetchClinicalDataNullAttributeSummaryProjection\_SQLServer() since without proper 'order by' clause getting the same order of results set in MSSQL vs MySQL is not guaranteed. Therefore, updated the unit testing method instead of changing the query itself.

### Service

### Scripts

# CBioPortal JUnit Test

## Overview

1. Uses Maven Plugin Surefire.( http://maven.apache.org/surefire/maven-surefire-plugin/)
2. Unit test defined in each module. Its configuration is defined in each module’s own pom.xml.
3. NOTE: in MySQL Set SESSION SQL\_MODE = ANSI\_QUOTE is used. This requires to use single quotes for quoting literal strings in MS SQL.

## How to debug unit test in eclipse

1. Debug all unit tests:
   1. From command line

>cd C:\eclipse\_genomics\Git\cbioportal-sqlserver\portal\target\

>mvn -Dmaven.surefire.debug test or run >debug-unittest.bat

* 1. From eclipse create a remove java application with host=localhost, port=5005

1. Debug a single unit test
   1. From command line

>cd C:\eclipse\_genomics\Git\cbioportal-sqlserver\portal\target\

>mvn -e -DfailIfNoTests=false -Dtest=TestDaoDrugInteraction -Dmaven.surefire.debug test

or run >debug-singletest.bat

* 1. From eclipse create a remove java application with host=localhost, port=5005

1. Note: for normal debugging mode use host = localhost, port=8000 for eclipse remote Java application.

## How to skip unit test of a module.

1. Add <skipTests>true</skipTests> to Surefire plugin's configuration in a module’s pom.xml.

<project>

[...]

<build>

<plugins>

<plugin>

<groupId>org.apache.maven.plugins</groupId>

<artifactId>maven-surefire-plugin</artifactId>

<version>2.4.2</version>

<configuration>

<skipTests>true</skipTests>

</configuration>

</plugin>

</plugins>

</build>

[...]

</project>

1. Alternatively, create a profile that will disable the tests (still the pom.xml of the module) (haven’t tried). With this solution if you run mvn clean package, it will run all tests. If you run mvn clean package -DnoTest=true, it will not run the tests for this module.

<project>

[...]

<profiles>

<profile>

<id>noTest</id>

<activation>

<property>

<name>noTest</name>

<value>true</value>

</property>

</activation>

<build>

<plugins>

<plugin>

<groupId>org.apache.maven.plugins</groupId>

<artifactId>maven-surefire-plugin</artifactId>

<version>2.4.2</version>

<configuration>

<skipTests>true</skipTests>

</configuration>

</plugin>

</plugins>

</build>

</profile>

</profiles>

[...]

</project>