affymetrix

User Manual

Affymetrix®
Chromosome Analysis Suite (ChAS) 2.1

For Research Use Only. Not for use in diagnostic procedures.

P/N 702943 Rev 7

Trademarks

Affymetrix®, GeneChip®, NetAffx®, Command Console®, Powered by Affymetrix™, GeneChip-compatible™, Genotyping Console™, DMET™, GeneTitan®, Axiom®, CytoScan®, and GeneAtlas® are trademarks or registered trademarks of Affymetrix, Inc. All other trademarks are the property of their respective owners.

All other trademarks are the property of their respective owners.

This database/product contains information from the Online Mendelian Inheritance in Man® (OMIM®) database, which has been obtained under a license from the Johns Hopkins University. This database/product does not represent the entire, unmodified OMIM® database, which is available in its entirety at http://www.omim.org/

Limited License Notice

Limited License. Subject to the Affymetrix terms and conditions that govern your use of Affymetrix products, Affymetrix grants you a non-exclusive, non-transferable, non-sublicensable license to use this Affymetrix product only in accordance with the manual and written instructions provided by Affymetrix. You understand and agree that except as expressly set forth in the Affymetrix terms and conditions, that no right or license to any patent or other intellectual property owned or licensable by Affymetrix is conveyed or implied by this Affymetrix product. In particular, no right or license is conveyed or implied to use this Affymetrix product in combination with a product not provided, licensed or specifically recommended by Affymetrix for such use.

Patents

Software products may be covered by one or more of the following patents: U.S. Patent Nos. 5,733,729; 5,795,716; 5,974,164; 6,066,454; 6,090,555; 6,185,561; 6,188,783; 6,223,127; 6,228,593; 6,229,911; 6,242,180; 6,308,170; 6,361,937; 6,420,108; 6,484,183; 6,505,125; 6510,391; 6,532,462; 6,546,340; 6,687,692; 6,607,887; 7,062,092 and other U.S. or foreign patents.

Copyright

© 2009-2014 Affymetrix, Inc. All rights reserved.

Contents

Chapter 1 Introduction 9
About Recent ChAS Updates 9
About this Manual 11
Technical Support 11

Chapter 2 Installing Chromosome Analysis Suite 13
Minimum and Recommended Requirements for ChAS 2.1 13
Recommended Requirements and Prerequisites for Arrays 15
Zip File Contents 15
Installing ChAS 2.1 16
New Installation 16
Upgrade Installation 16
Copying Analysis Files 18
Analysis File Locations 19
Viewing Hidden Files and Folders 21
Windows 7 and XP Users 21
Windows 8.1 Users 21
Analysis File Download 23
Setting Proxy Server Access 27
Un-installing ChAS 29

Chapter 3 Getting Started with Chromosome Analysis Suite. 30
Starting Chromosome Analysis Suite 30
Analysis Workflow Module 32
Setting Up the Analysis Workflow Module (First Time) 32
File Types and Data Organization in Chromosome Analysis Suite 33
Chromosome Analysis Suite File Types 33
Data Organization in ChAS 35
Basic Workflow for Cytogenetics Analysis 35
Array Processing Workflow (Using AGCC) 36
Probe-Level Analysis of CEL File Data 36
Loading Data into ChAS for Display 38
Viewing Data and Features of Interest using the ChAS Display Controls 39
Working with ChAS 40
Accessing Functions in ChAS 40
Changing the Size of Panes in ChAS 40
Opening Panes in Separate Windows 41

Chapter 4 CN/LOH/Mosaicism Analysis. 43
Single Sample Analysis 44
ChAS and NetAffx Analysis File Compatibility 44
Introduction to Single Sample Analysis 44

LOH Segments on the X and Y Chromosomes 47
Performing a Single Sample Analysis 49
Reference Files 58
Creating Custom QC Views in the Analysis Workflow 59
Exporting QC Tables and Probe-Level Data Using Generate Report 63
Analysis Workflow Troubleshooting 64
Log Rollover 65
Log Collection 65

Chapter 5 Loading Data 66

Introduction to Loading Data 66
Loading Files 68
Loading Files using the Search Feature 70
Segment Smoothing and Segment Joining (Optional) 73
Segment Tab Options 76
About Smoothing 77
About Joining 78
Setting QC Parameters in the ChAS Browser 79
QC Thresholds Tab Options 80
Adding a QC Property 81
Editing an Existing QC Threshold 82

Chapter 6 Viewing Data in ChAS 83

Overview of ChAS Window Components 84
Files List 85
Data Types List 86
Named Settings 87
Status Bar 87
Display Area 88
Upper Panes 89
Selected Chromosome View 89
Lower Panes 90
Changing the ChAS Browser NetAffx Genomics Annotation File Version 90

Chapter 7 Displaying Data in Graphic Views. 91

The Graphic Display Views 92
Karyoview 93
Whole Genome View 97
Selected Chromosome View 101
Detail View 104
Selecting a Chromosome Section for Display 109
Edit Configurations in the Miscellaneous (Misc) Tab 111
Autosave 112
Coordinated Box Format 112
Zoom Buffer 113

CHP File Colors 113
Controlling the Display of Data 115
Selecting Data for Display 116
Selecting Data Types for Display 118
Changing the Grouping of Samples and Data Types 119
Selecting Light or Dark Schemes for Display 124
Expanding and Contracting Annotations 124
Changing Graph Appearance 125
Learning More About Features 132
Pop-ups 132
Right-click Menu Options 133
Selection Details Table 135
Going to External Websites 140

Chapter 8 Filtering Segments 142
Applying Segment Parameter Filters 142

Chapter 9 Modification and Interpretation of Segment Data 145
Editing Segment Data 145
Using Edit Mode 146
Types of Segment Editing 146
Merging Segments 147
Merging Segment Groups 147
Segment to Segment Merge 150
Un-Merging Merged Segments 152
Deleting Segments 153
Un-Deleting a Deleted Segment(s) 155
Editing the Start/End Coordinates of a Segment 155
Un-Doing the Edited Start/End Coordinates of a Segment 157
De Novo Segment Drawing 157
Un-Doing a Segment De Novo Drawing 159
Changing the Copy Number 161
Un-Doing a Copy Number Change 161
Modified Segments in the Segments Table 162
Revert All Edits for a Sample 163
Setting Up and Using the Calls Feature 164
Setting Up Calls 164
Using the Calls Feature 166
Setting Up and Using the Interpretations and Snippets Feature 168
Setting Up the Interpretation Snippets Feature 168
Using the Interpretation Snippets Feature 170
Tracking (Reviewing the Log File) 173

Chapter 10 Using CytoRegions 178
Selecting a CytoRegions Information File 178

Viewing CytoRegions 180
CytoRegions in the Graphic Views 180
CytoRegions Table 182
Searching CytoRegions 185
Using Filters with CytoRegions 185
Using Restricted Mode 186

Chapter 11 Using the Overlap Map 188
Selecting the Overlap Map File 188
Viewing Overlap Regions 190
Viewing Overlap Map Regions in the Graphic Displays 190
Viewing the Overlap Map Table 192
Setting the Overlap Filter 194

Chapter 12 Creating and Editing AED Files 196
Creating an AED File of Annotations 196
Adding Annotations to an AED File 198
Creating an AED File of Regions 198
Adding Regions to an Existing AED File 201
Deleting Regions from an AED File 203
Viewing or Editing AED File Properties 203
Viewing the Genome Assembly Version 204
Viewing and Editing Annotations 207
Entering General Information 208
Customizing Properties 209
New User Annotations 215
Viewing and Batch Editing AED File Contents 216
AED/BED Color Rules 219
Exporting Information in AED or BED Format 224

Chapter 13 Displaying Data in Table Views 228
Common Table Operations 228
The Standard Tool Bar Controls 228
Sorting by Columns 229
Ordering Columns 230
Selecting Columns for Display 230
Sum, Mean, and Median Calculator 232
Saved Table States 232
Modifying Table States 233
Applying Previously Saved Table States 235
Deleting Previously Saved Table States 236
Segments Table 236
Tool Bar 237
Segments Table 237
Obtaining All Annotations Associated With a Segment 240

Exporting Selected Segments to a PDF File 243
Graphs Table 248
Tool bar 248
Graphs Table 249
Somatic Mutations Table 253
Tool Bar 253
Somatic Mutation Table 254
QC and Sample Info Tab 255
QC and Sample Information Table 256
Loaded AED/BED Files Table 257
Chromosome Summary Data 258
Auto-Generated Autosome LOH Percentage 260
Searching Results 260
Finding Intersections 265

Chapter 14 Reporting Results 269

Exporting Graphic Views 269
Exporting as PDF Reports 270
Export as PNG 273
Exporting Table Data 274
Export Table Data in a PDF File 275
Export Tables as TXT File 280
Exporting a Segments Table with Modified Segments to a .TXT File 282
Transfer to Clipboard 283
Combining PDFs into a Single Report 284

Chapter 15 User Profiles and Named Settings 289

Types of Settings 289
User Profiles 289
Named Settings 289
Creating and Using User Profiles 290
Creating and Using Named Settings 292
Exporting and Importing Preferences 294

Appendix A Analysis Parameters 297

Analysis Parameters for Single Sample Analysis 297
Reference Model File Creation 297

Appendix B AED File Format 298

AED File Description 298
Header row 299
Metadata Records 300
Annotations 300
Property Name Elements 301

Namespaces 301
Properties 301
Types 303
Compatibility 304
UCSC Browser Extensible Data (BED) 304
Microsoft Excel and Other Spreadsheet Applications 305
Microsoft Notepad and Other Text Editors 305
References 306

Appendix C ChAS Properties and Types 307
Identifying Properties within ChAS 307
Standard AED Property Style 307
ChAS Property Style 308
Full URI Property Style 308
Identifying Value Types within ChAS 308
Automatic Conversion of CxCHP Headers to Properties 309
Header Name Conversion 309
Converted Properties 310
Derived Properties 310

Appendix D ChAS Browser NetAffxGenomic Annotations 311
Homo Sapiens Database Files 311
NetAffxGenomic Annotation Files 311

Appendix E Genomic Position Coordinates 313
Genome Assemblies 313
SNP and Marker Positions 313
Segment Positions 313
Difference Between ChAS 1.0 and ChAS 1.1 314

Appendix F BED Files 315
Editing BED Files 315

Appendix G CytoScan® Algorithms and QC Metrics 316
Algorithm Overview 316
Feature Identification and Signal Extraction 316
Single Sample CytoScan® Workflow 316
Signal-Level Covariate Adjustors 318
Dual Quantile Normalization 319
Copy Number Workflow 319
Log2 Ratio Calculation 319
High Pass Filter Image Correction 319
Log2 Ratio-Level Covariate Adjustors 320

Median Autosome Normalization 320
Systematic Residual Variability Removal 320
Signal Restoration 320
Segmentation 320
Segment Table Output 322
Mosaicism Segment Algorithm 322
Limitations of the Mosaicism Algorithm for CytoScan Arrays 323
SNP Marker Workflow 323
Signal Summarization 323
Allelic Signal Computation 323
Genotyping 324
Allelic Difference GC Correction 324
Allele Peak Track Computation 324
Detection of LOH 324
Array Data QC Metrics 324
Median of the Absolute values of all Pairwise Differences (MAPD) 324
Waviness SD 325
SNPQC 325
MAPD – Detailed Description 325
Effect of MAPD on Functional Performance 326
Waviness-SD – Detailed Description 326
SNPQC – Detailed Description 327
Effect of SNPQC on Functional Performance 328

Introduction

The Chromosome Analysis Suite (ChAS) software for cytogenetic analysis enables you to view and summarize chromosomal aberrations across the genome. Chromosomal aberrations may include copy number gain or loss, mosaicism, or loss of heterozygosity (LOH).

ChAS provides tools to:

- Perform single sample analysis of CEL files from Affymetrix® CytoScan® Arrays.
- Analyze segment data at different levels of resolution.
- View results data (CYCHP files) that summarize chromosomal aberrations in table and graphical formats.
- Display CNCHP data from Affymetrix® Genome-Wide Human SNP Array 6.0.
- Display OSCHP data from Affymetrix OncoScan FFPE Assay arrays.
- Customize and load your own annotations and regions for focused analysis.
- Apply separate filters to the entire genome and user-specified regions of interest to remove irrelevant information such as segments in areas that are not of interest.
- Perform detailed comparisons between different samples.
- Directly access external databases such as NCBI, UCSC Genome Browser, Ensembl, and OMIM.
- Generate reports of user-selected data in graphical and tabular formats.
- Edit and add Interpretation to Segment Data.

IMPORTANT: **The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.**

Chromosome Analysis Suite is not a secondary analysis package. However, it does create CYCHP files and tab-delimited text files required for secondary analysis packages available from companies in the Affymetrix® GeneChip® Compatible Program.

About Recent ChAS Updates

ChAS 2.1:

- Whole Genome View.
- OncoScan OSCHP file visualization including somatic mutation calls.
- Auto-Dots graph type enables the transition from points to big dots graph type based on a user-defined region.
- New Analysis and Visualization Library files (NA33) for CytoScan arrays. Allows for simultaneous viewing of OncoScan and CytoScan files.
- Automatic Autosome Percentage LOH calculation based on user defined LOH filter setting.
- New column in Segment Table (Full Location) containing genomic coordinates in 1 column for direct use in external websites.
- Updates to the Analysis Workflow including export of probe level data for the entire sample.
- New library file folder location outside any hidden folders.
- Y axis coordinates no longer have fixed boundaries in the Detail View.
- Disabled mosaic segment calls for the Y chromosome.

- Median Log2 data values are calculated for each CN Segment in both the Segments table and *mouse over*.

ChAS 2.0:

- Full package Installer (Analysis and Visualization on 64bit/8+GB RAM systems).
- Browser-only Installer (Visualization-only, 32bitwith 3+GB RAM systems, 64bit with 8+GB RAM systems).
- New Analysis and Visualization Library files for CytoScan Arrays (NA32.3).
- Mosaicism Algorithm: GainMosaic and LossMosaic events called, and visualized in Browser as new segment data types.
- New Analysis Workflow module: for setting up and running analyses, monitoring QC trending over time, log collection.
- ChAS 2.0 allows you to edit CYCHP and CNCHP data segments in the ChAS Browser. Merge, delete, draw de novo, adjust boundaries or undo your changes. Enables visualization and tracking of changes made.
- Array Nomenclature: ISCN-based description for each segment.
- Zoom Buffer around segments: user configurable click-free breakpoint visualization
- Enhanced Zooming: 3X Zoom in and out buttons, “Shift+click+draw-a-box” zooming to any coordinates.
- Color of CHP data: User configurable colors for sample data color assignment upon loading into Browser.
- AED File Editor: Allows editing and searching of AED file contents.
- “Calls” column driven by user-configured drop list.
- “Interpretation Snippets”: Enables user-configurable canned text to be inserted directly into “Interpretations” field.
- Named Saved Table States: Enables switching between multiple user-configurable saved custom column sets in tables. Allows for switching between large numbers of table columns that might be used during interpreting, to a subset of columns for non-cluttered reporting, without manually adding, removing, or rearranging columns.
- Genes and OMIM information: Updated data tracks, number/count of OMIM genes and RefSeq genes in new Segments table columns.
- Directional Arrows to indicate transcription direction on annotations.
- Enhanced “Interpretations” column editing: enables insertion of user-configured “snippets” of text.
- Gender symbol now displayed with CHP file name.
- Failing QC metrics are now color-highlighted. Metrics that have failed thresholds are highlighted in the Analysis Workflow Dashboard and in the Browser.
- Auto-calculation of the Sum, Median, Mean of any table data numbers selected.
- Improved *Search* function to allow searches by text that include additional columns in the results.
- Intuitive Segment Names: Name is now generated by: the Segment Type, State and Sample filename for each segment.
- Flanking markers/positions listed in Segments Table: Enables enhanced breakpoint reporting.
- Mean Log2 and mean Weighted Log2 data values calculated for each CN Segment in Segments table.
- Removed Segment “confidence” value in all tables and visualizations.
- New RUO-blue color scheme. ChASDx software will adopt a red color scheme when launched.

ChAS 1.2.2 and 1.2.1 contain improved AllelePeaks track data for rare challenging samples, and bug fixes.

ChAS 1.2 versions include the following new features and enhancements:

- Genotype calling for Affymetrix® CytoScan® HD Array and CytoScan® 750K Array results, exporting genotype calls.
- Status bar displays the NetAffx database build loaded in the ChAS Browser.
- Load a different NetAffx Genomic Annotations Browser database into the ChAS Browser without restarting the software.
- View sample and results information in the Sample Information window.
- New Named Settings.
- Easily identify User Profiles available for import.
- Apply user-selected colors to reference annotations (except DGV annotations, which contain a pre-defined set of differential color codes) in the Detail View.
- Segments table updates:
  - When multiple reference annotations overlap a segment, the genes in the Segments table row are listed in the order that they appear from left to right in the Detail View.
  - The “Types” column sorts segments in the same order that they appear in the Data Types window pane.
- ChAS verifies the hg version of AED or BED files and will not load files with hg version different from that currently loaded in the ChAS Browser. If the hg version information is not found in the AED or BED file, ChAS warns you before loading the file.
- AED files created in ChAS 1.2 include the hg version.
- All loaded AED or BED files are automatically reloaded in a subsequent session with the same user profile.
- Reports include CytoRegion and Overlap Map file names.
- ChAS remembers the destination folder for exported reports for a user profile during a session and in subsequent sessions.

About this Manual

This manual provides step-by-step instructions for performing the procedures required to use ChAS. The content of this manual is also available in the online help for ChAS. The steps outlining procedures are frequently supplemented with screen captures to further illustrate the instructions. The screen captures in this manual may not exactly match the windows displayed on your screen.

Technical Support

Affymetrix provides technical support to all licensed users via phone or E-mail. To contact Affymetrix Technical Support:

AFFYMETRIX, INC.

3420 Central Expressway

Santa Clara, CA 95051 USA

Tel: 1-888-362-2447 (1-888-DNA-CHIP)

Fax: 1-408-731-5441

sales@affymetrix.com

support@affymetrix.com

AFFYMETRIX UK Ltd.,
Voyager, Mercury Park,
Wycombe Lane, Wooburn Green,
High Wycombe HP10 0HH
United Kingdom
UK and Others Tel: +44 (0) 1628 552550
France Tel: 0800919505
Germany Tel: 01803001334
Fax: +44 (0) 1628 552585
saleseurope@affymetrix.com
supporteurope@affymetrix.com

AFFYMETRIX JAPAN K.K.
Mita NN Bldg. 16F
4-1-23 Shiba Minato-ku,
Tokyo 108-0014 Japan
Tel. 03-5730-8200
Fax: 03-5730-8201
salesjapan@affymetrix.com
supportjapan@affymetrix.com

Installing Chromosome Analysis Suite

Chromosome Analysis Suite (ChAS) is a stand-alone application.

ChAS 2.1 supports the analysis and/or visualization of CytoScan® CYCHP results data files.

ChAS 2.1 supports the visualization of Genome-Wide Human SNP 6.0 CNCHP results data files.

ChAS 2.1 does not support the analysis or visualization of Cytogenetics 2.7M Array files.

ChAS 2.1 supports the visualization of OncoScan FFPE Assay OSCHP files.

To process CytoScan® HD Arrays and analyze the intensity (CEL file) data, you need to install:

- CytoScan® HD Array software plug-in for AGCC for processing arrays.
- Chromosome Analysis Suite 2.1 for data analysis and visualization.
- Analysis files for Chromosome Analysis Suite.

To process OncoScan FFPE Assay arrays and analyze the intensity (CEL file) data:

- Please refer to the OncoScan Console User Manual (P/N 703195).

NOTE: ChAS 2.1 is for OSCHP file visualization only. Analysis of CEL file data is performed in OncoScan Console.

IMPORTANT: Due to the amount of memory that ChAS needs to operate, Affymetrix VERY STRONGLY recommends that you DO NOT install the ChAS software on production AGCC computers being used for scanning and operating fluidics systems.

This chapter includes:

- *Minimum and Recommended Requirements for ChAS 2.1*
- *Zip File Contents* (on page 15)
- *Installing ChAS 2.1* (on page 16)
- *Viewing Hidden Files and Folders* (on page 21)
- *Setting Proxy Server Access* (on page 27)

Minimum and Recommended Requirements for ChAS 2.1

- ChAS 2.1 can be installed 2 ways:
    1) Full Suite (Analysis and Visualization modules)
    2) Visualization-only version.
- The Full Suite package of Chromosome Analysis Suite 2.1 software must be installed on a 64-bit analysis workstations to create CYCHP results data files.
- This software takes the CytoScan Array CEL files as input to the Analysis engine and allows researchers to perform Copy Number, LOH, Genotyping, and other types of genetic analysis.

- Chromosome Analysis Suite 2.1 is also available as a Visualization package which can be installed on 32bit or 64bit workstations, to visualize CYCHP file data.
- The Visualization package is not capable of creating CNCHP, CYCHP, or OSCHP files, but can view them and create reports.

Recommended System Requirements

ChAS FULL Package		ChAS Browser ONLY	
Processor	3 GHz (or greater) Pentium Dual Core Processor	**Processor**	3 GHz (or greater) Pentium Dual Core Processor
64-bit Windows Operating System and Web Browser	Windows 7 Professional SP1 and Windows 8.1 Internet Explorer 8.0 or above	**64-bit Win7 and Win8.1 Operating System and Web Browser**	Windows 7 Professional SP1 and Windows 8.1 Internet Explorer 8.0 or above
Available Disk Space	250 GB HD + data storage	**Available Disk Space**	250 GB HD + data storage
Free Disk Space Required at Installation	≥ 5 GB	**Free Disk Space Required at Installation**	≥ 5 GB
RAM	16 GB	**RAM**	16 GB

Minimum System Requirements

ChAS FULL Package		ChAS Browser ONLY	
Processor	2 GHz (or greater) Pentium Dual Core Processor	**Processor**	2 GHz (or greater) Pentium Dual Core Processor
64-bit Windows Operating System and Web Browser	Windows 7 Professional SP1 and Windows 8.1 Internet Explorer 8.0 or above	**32-bit XP and 64-bit Win7 and Win8.1 Operating System and Web Browser**	Windows XP SP3 (32-bit) Windows 7 SP1 (64-bit) Windows 8.1 Internet Explorer 8.0 or above
Available Disk Space	150 GB HD + data storage	**Available Disk Space**	150 GB HD + data storage
Free Disk Space Required at Installation	≥ 5 GB	**Free Disk Space Required at Installation**	≥ 5 GB
RAM	8 GB	**RAM**	3 GB (32-bit System) 8 GB (64-bit System)

Recommended Requirements and Prerequisites for Arrays

A 64-bit system is recommended for all array types. System requirements are specific for the array type that you plan to analyze (CEL> CYCHP).

NOTE: Chromosome Analysis Suite requires AGCC 3.2.4 or higher to produce CytoScan® CEL files.

NOTE: A 64-bit system is required to generate CytoScan® CYCHP data files, however CytoScan® CYCHP data files can be viewed on a 32-bit or 64-bit system.

- Operating system on 32-bit
  - Windows XP with service pack 3
- Operating system on 64-bit
  - Windows 7 Professional with service pack 1.
  - Windows 8.1 Professional

The Chromosome Analysis Suite has been verified on the operating systems listed above.

ChAS may work on other Windows operating systems, however operating systems (not listed) have not been verified.

NOTE: The larger file sizes associated with the CytoScan® HD Array should be taken into account when calculating the necessary free space requirement. A CytoScan® HD Array CYCHP file is ~120 MB.

Zip File Contents

There are 4 items you will need to run the CytoScan® HD Array and visualize the data in ChAS 2.1.

- 2 AGCC workstation files, from the CytoScan® HD product web page at Affymetrix.com
  - cytoscan_hd_agcc.zip
  - CytoScanHD Fluidics Script (see array product page for details)
- 2 Data Analysis (ChAS) workstation files, from ChAS product web page (www.affymetrix.com/chas):
  - Chromosome_Analysis_Suite_2.1.zip (instructions below)
  - CytoScan_HD_Analysis_Files_NA33.zip (see below or download within ChAS 2.1)

The ChAS software installer is in a zipped package with an executable.

The zipped package can be downloaded from **www.affymetrix.com/chas**

After you extract the files from the “Chromosome_Analysis_Suite_2.1.zip¨ zipped package, a folder named “Chromosome Analysis Suite 2.1" appears.

- The Chromosome Analysis Suite 2.1 folder contains:
  - ChAS 2.1 Analysis and Browser (64-bit REQUIRED)
  - ChAS 2.1 Browser (32-bit)
  - ChAS 2.1 Manual.pdf
  - ChAS 2.1 Release Notes.pdf

Installing ChAS 2.1

NOTE: The setup process installs additional required components, which includes Java components and Visual C++ runtime.

New Installation

- Download the **Chromosome_Analysis_Suite_2.1.zip** package from the Affymetrix website.

Analysis and Browser Package Installation:

1. To install the Analysis and Browser package, double click on the ChAS 2.1 Analysis&Browser (64bit).exe file from the “ChAS 2.1 Analysis&Browser” folder. The Install Shield Wizard for Chromosome Analysis Suite will begin.
2. Follow the on-screen instructions on your system to complete the ChAS installation.
3. After the installation is completed, you will need to download new NA33 or higher Analysis files from the NetAffx site, or copy them into your ChAS Library folder. If you are unable to connect to the internet, then go to the section of this document called "Copying Analysis Files" below.

Browser Package Installation:

1. To install the Browser package, double click on the ChAS2.1_Browser.exe file from the "ChAS 2.1 Browser" Folder. The Install Shield Wizard for Chromosome Analysis Suite will begin.
2. Follow the on-screen instructions on your system to complete the ChAS installation.
3. After the installation is completed, you will need to download new Analysis files from the NetAffx site, or copy them into your ChAS Library folder. If you are unable to connect to the internet, then go to the section of this document called “Copying Analysis Files” below.

Upgrade Installation

NOTE: The ChAS 2.1 installer supports upgrading from ChAS versions 2.0, 1.2.2, and 1.2.1. A new Library path will be created by the ChAS 2.1 installer and Browser annotations and array-type-specific annotation files will be copied to the new library folder. NA33 and/or NA32.3 library files are required to run an array analyses in ChAS 2.1.

Starting with ChAS 1.2, the ChAS installer installs a copy of Java JRE into the program directory for its own use. The installation proceeds regardless of whether there are other versions of Java JRE or Java SDK installed on your system. Adding or removing other instances of Java will not affect the copy of Java JRE in the ChAS program directory.

All pending or canceled Workflows from ChAS 2.0 Analysis Manager should be completed or removed prior to upgrading to ChAS 2.1.

Analysis and Browser Installation from ChAS 2.0:

1. Double-click on the ChAS 2.1 Analysis&Browser (64bit).exe icon to uninstall ChAS 2.0 from the workstation.
2. To install the Analysis and Browser package, double click on the ChAS 2.1 Analysis&Browser (64bit).exe file from the "ChAS 2.1 Analysis&Browser” folder. The Install Shield Wizard for Chromosome Analysis Suite will begin.
3. Follow the on-screen instructions on your system to complete the installation.

4. During installation, the NA32.3 analysis library files will be copied to the new library file location. These files can be used for analysis in ChAS 2.1 If you prefer NA33 or higher for analysis, you will need to download these analysis files from the NetAffx site, or copy them into your ChAS Library folder. If you are unable to connect to the internet, then go to the section of the document called "Copying Analysis Files" below.

Analysis and Browser Installation from ChAS 1.2.1 or 1.2.2:

1. Before you install/upgrade the software close the Analysis Manager icon which is located on the bottom right hand side of the task bar.
2. To install the Analysis and Browser package, double click on the ChAS 2.1 Analysis&Browser (64bit).exe file from the "ChAS 2.1 Analysis&Browser" folder. The Install Shield Wizard for Chromosome Analysis Suite will begin.
3. Follow the on-screen instructions on your system to complete the installation.
4. After the installation is completed, you need to download new NA33 or higher Analysis files from the NetAffx site, or copy them into your ChAS Library folder. If you are unable to connect to the internet, then go to the section of this document called "Copying Analysis Files" below.

NOTE: If NA32.3 files are required, please go to affymetrix.com/chas to download the *CytoScan_HD_Analysis_Files_NA32.3.zip* file.

Browser Installation from ChAS 2.0:

1. Double-click on the ChAS 2.1 Analysis&Browser (64bit).exe icon to uninstall the ChAS 2.0 Browser from the workstation.
2. To install the Browser package, double click on the ChAS 2.1 Analysis&Browser (64bit).exe file from the "ChAS 2.1 Analysis&Browser" folder. The Install Shield Wizard for Chromosome Analysis Suite will begin.
3. Follow the on-screen instructions on your system to complete the installation.
4. After the installation is completed, you need to download new NA33 or higher Analysis files from the NetAffx site, or copy them into your ChAS Library folder. If you are unable to connect to the internet, then go to the section of this document called "Copying Analysis Files" below.

NOTE: If NA32.3 files are required, please go to affymetrix.com/chas to download the *CytoScan_HD_Analysis_Files_NA32.3.zip* file.

Browser Package Installation from ChAS 1.2.1, ChAS 1.2.2:

1. Before you install the software close the Analysis Manager icon which is located on the bottom right hand side of the task bar.
2. To install the Browser package, double click on the ChAS2.1_Browser.exe file from the "ChAS 2.1 Browser" Folder. The Install Shield Wizard for Chromosome Analysis Suite will begin.
3. Follow the on-screen instructions on your system to complete the installation.
4. After the installation is completed, you will need to download new Analysis files from the NetAffx site, or copy them into your ChAS Library folder. If you are unable to connect to the internet, then go to the section of this document called "Copying Analysis Files" below.

Upgrading From ChAS 1.2 or Earlier:

- Upgrading from ChAS 1.2 or earlier to ChAS 2.1 is not supported. Please uninstall the earlier version of ChAS from the Add/Remove Programs in the Control Panel. After ChAS is uninstalled, then run the ChAS 2.1 installation. See page 16 for more information.

Copying Analysis Files

The CytoScanHD Analysis Files zip package download contains the Analysis Files required to process CytoScanHD Array CEL files into CYCHP files.

Also included are the visualization files for GenomeWideSNP_6 files (required in ChAS to view GenomeWideSNP_6 files) and OncoScan FFPE Assay (required in ChAS to view these array types).

If you are unable to connect to the Internet to download the Analysis files through ChAS, download the analysis files from the affymetrix.com/chas webpage, then extract the file contents to the following location:

Windows XP Users:

Copy the contents from the Analysis Files folder to:

`C:\Affymetrix\ChAS\Library`

The path shown above may vary on non-US-English systems.

Windows 7 and 8.1 Users:

Copy the contents from the Analysis Files folder to:

`C:\Affymetrix\ChAS\Library`

Analysis Files Package Contents:

The list of Analysis Files to copy from the unzipped CytoScanHD Analysis Files package are named:

1. CytoScanHD_Array.ExportProbesetTable.r1.ReportScript
2. CytoScanHD_Array.cdf
3. CytoScanHD_Array.CelGenderCheck.r2.cytoruo.WFStepTemplate
4. CytoScanHD_Array.chasarrinfo
5. CytoScanHD_Array.chasdataprocessconfig
6. CytoScanHD_Array.chasqcconfig
7. CytoScanHD_Array.chrXprobes
8. CytoScanHD_Array.chrYprobes
9. CytoScanHD_Array.hg19.v3.refcovar
10. CytoScanHD_Array.na33.annot.db
11. CytoScanHD_Array.na33.r1.REF_MODEL
12. CytoScanHD_Array.probe_tab
13. CytoScanHD_Array.QCMetricsCollection.r2.cytoruo.WFStepTemplate
14. CytoScanHD_Array.r2.cytoruo.QCHistoryLog
15. CytoScanHD_Array.r1.qca
16. CytoScanHD_Array.r1.qcc
17. CytoScanHD_Array.reference_model.def.NA33.r1.chasparam
18. CytoScanHD_Array.RefModelCreation.NA33.r1.cytoruo.WFScriptTemplate
19. CytoScanHD_Array.RMCCleanup.r2.cytoruo.WFStepTemplate
20. CytoScanHD_Array.RMCCopyFromSource.r2.cytoruo.WFStepTemplate
21. CytoScanHD_Array.RMCFileCheck.r2.cytoruo.WFStepTemplate
22. CytoScanHD_Array.RMCMoveToLibPath.r2.cytoruo.WFStepTemplate
23. CytoScanHD_Array.single_sample.def.NA33.r1.chasparam
24. CytoScanHD_Array.SingleSampleAnalysis.NA33.r1.cytoruo.WFScriptTemplate
25. CytoScanHD_Array.snplist.txt

26. CytoScanHD_Array.SSACleanup.r2.cytoruo.WFStepTemplate
27. CytoScanHD_Array.SSACopyFromSource.r2.cytoruo.WFStepTemplate
28. CytoScanHD_Array.SSAMoveToOutputPath.r2.cytoruo.WFStepTemplate
29. CytoScanHDMetrics-Basic View.r1.awqcconfig
30. CytoScanHDMetrics-Detail View.r1.awqcconfig
31. GenomeWideSNP_6--Basic View.r1.awqcconfig
32. GenomeWideSNP_6.chasarrinfo
33. GenomeWideSNP_6.chasdataprocessconfig
34. GenomeWideSNP_6.chasqcconfig
35. NetAffxGenomicAnnotations.Homo_sapiens.hg19.na33.db
36. OncoScan.chassarrinfo
37. OncoScan.chasdataprocessingconfig
38. OncoScan.chasconfig
39. OncoScan.ExportAllTables.r1.ReportScript
40. OncoScan.ExportProbesetTable.r1.ReportScript
41. OncoScan.ExportSegmentTable.r1.ReportScript
42. OncoScan.ExportSigSNPS.r1.ReportScript
43. OncoScan.ExportSOMTable.r1.ReportScript
44. OncoScan.na33.r1.annot.db
45. OncoScan.na33.r2.som_annotations
46. OncoScanMetrics-Basic View.r1.awqcconfig
47. OncoScanMetrics-Detail View.r1.awqcconfig

Analysis File Locations

To see the location of the analysis files:

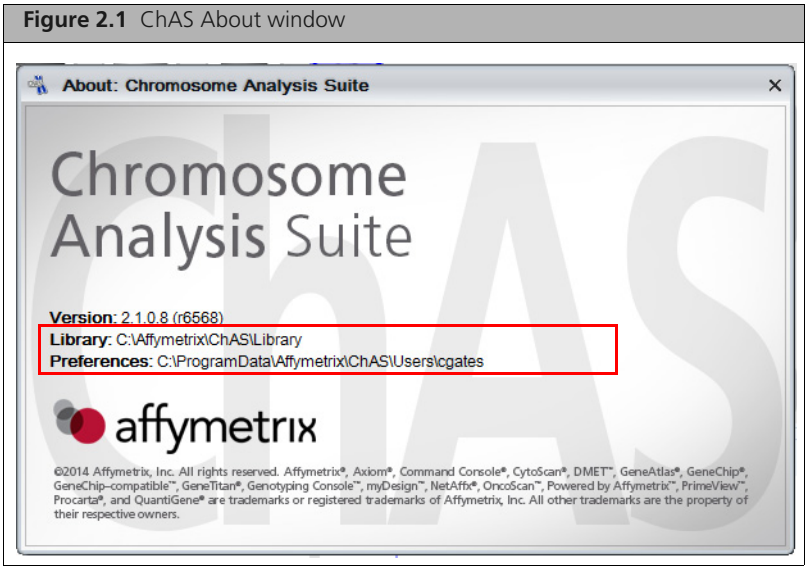
- From the Help menu, click **About...**

  OR

- Click the **About** button on the main tool bar.

The About this Software window opens. (Figure 2.1)

Figure 2.1 ChAS About window



At the bottom it displays the path and location for both the library and preference files.

For Windows XP:

- Library: `C:\Affymetrix\ChAS\Library`
- Preferences: `C:\Documents and Settings\All Users\Application Data\Affymetrix\ChAS\users`

For Windows 7 and 8.1:

- Library: `C:\Affymetrix\ChAS\Library`
- Preference file: `C:\ProgramData\Affymetrix\ChAS\preferences.xml`
- All other user profile related preference files and saved settings: `C:\ProgramData\Affymetrix\ChAS\users`

NOTE: File locations are different in Windows XP and Windows 7/8.1.

If you cannot see the folder locations listed above, go to *Viewing Hidden Files and Folders* (on page 21).

Viewing Hidden Files and Folders

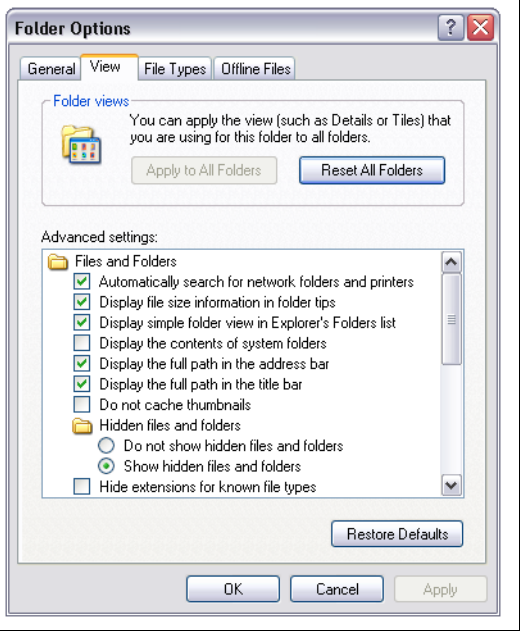
The ChAS preference files may be placed in folders and files that are normally hidden from the user in Windows.

Windows 7 and XP Users

To display hidden files and folders in Windows 7 and XP:

1. (Windows 7) Click the **Start** button, then select **Control Panel**.
   (Windows XP) Click the **Start** button, then select **Settings** → **Control Panel**.
2. At the Control Panel window, double-click **Folder Options** icon.
   The Folder Options window appears. (Figure 2.2)

Figure 2.2 Folder Options window (Windows XP)



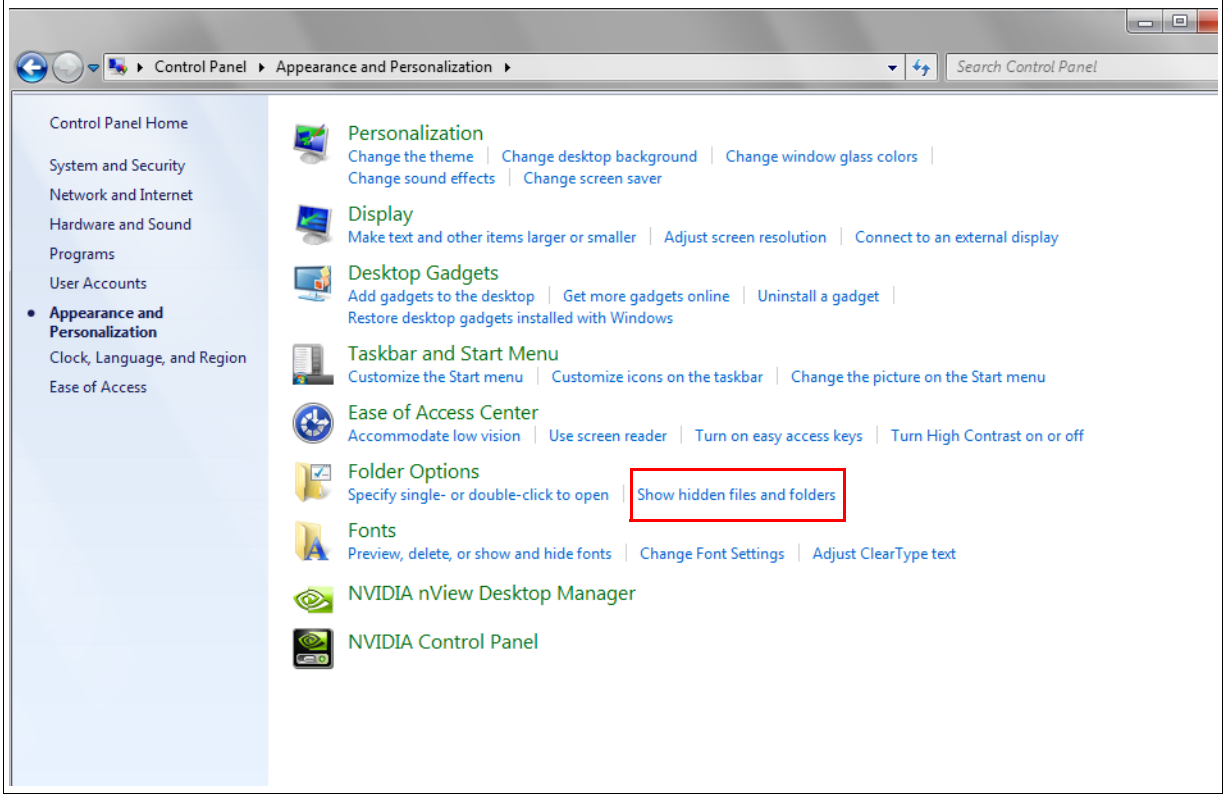
3. On the View tab, under Hidden files and folders, click **Show hidden files and folders**.
   Hidden files and folders are dimmed to indicate they are not typical items.
   If you know the name of a hidden file or folder, you can search for it.

Windows 8.1 Users

To display hidden files and folders in Windows 8.1:

1. In the Start menu, click **Control Panel**.
   The Control Panel window appears. (Figure 2.3)

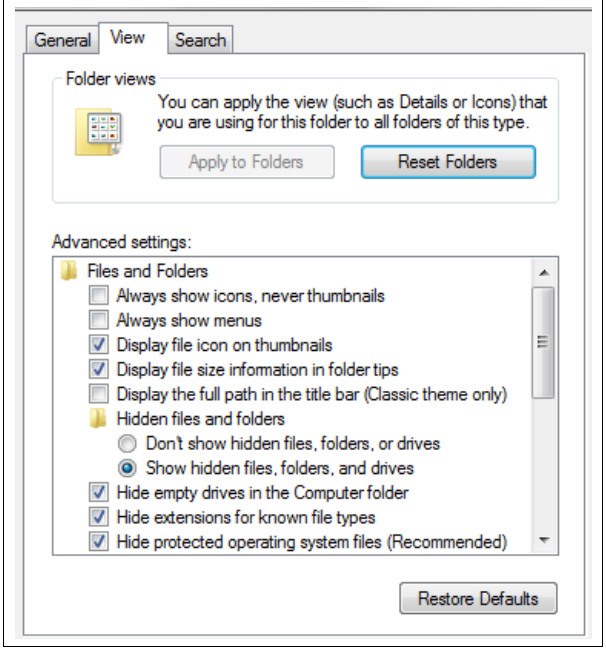
Figure 2.3 Control Panel (Windows 8.1)



2. In the Folder Options window that appears (Figure 2.4), click the **View** tab. Under Hidden files and folders, click **Show hidden files and folders.**

   Hidden files and folders are dimmed to indicate they are not typical items. If you know the name of a hidden file or folder, you can search for it.

Figure 2.4 Folder Options window



3. Click **OK**.
4. Close all open windows.

Analysis File Download

When you start ChAS for the first time, you will be prompted to:

1. Create a user profile (see *Creating and Using User Profiles* (on page 290)).
2. If ChAS is being installed for the first time, you must download the library files from NetAffx that ChAS uses to analyze and annotate data.

NOTE: To process the CytoScan® HD Array in AGCC, you will also need to install the appropriate library files for AGCC on the AGCC workstation (see the specific array product page at www.affymetrix.com for details).

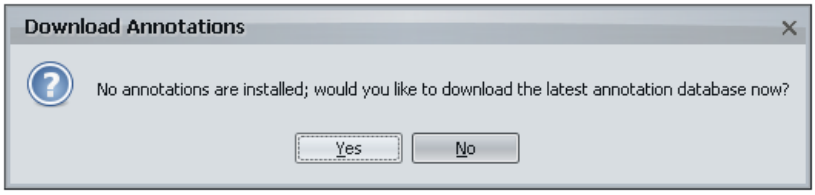
You can download the ChAS analysis files from NetAffx using either the ChAS Browser or the Analysis Workflow. The files are saved into the same library folder whether downloading through the ChAS Browser or the Analysis Workflow.

To download ChAS analysis files from NetAffx through the ChAS Browser:

1. Create a NetAffx account with a user name and password (go to www.affymetrix.com and click **Register** at the top of the page).
2. Start ChAS Browser.

   If no annotations are installed, a Download Annotations notice appears (Figure 2.5).

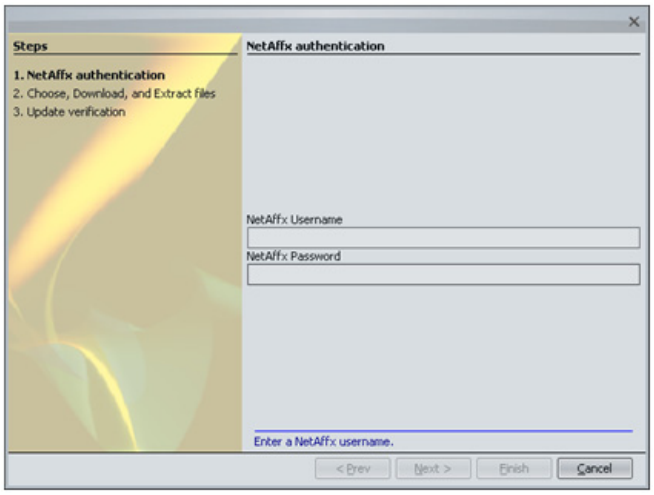
Figure 2.5 Download Annotations notice



3. Click **Yes**.

   The NetAffx Authentication window opens (Figure 2.6).

Figure 2.6 NetAffx Authentication window

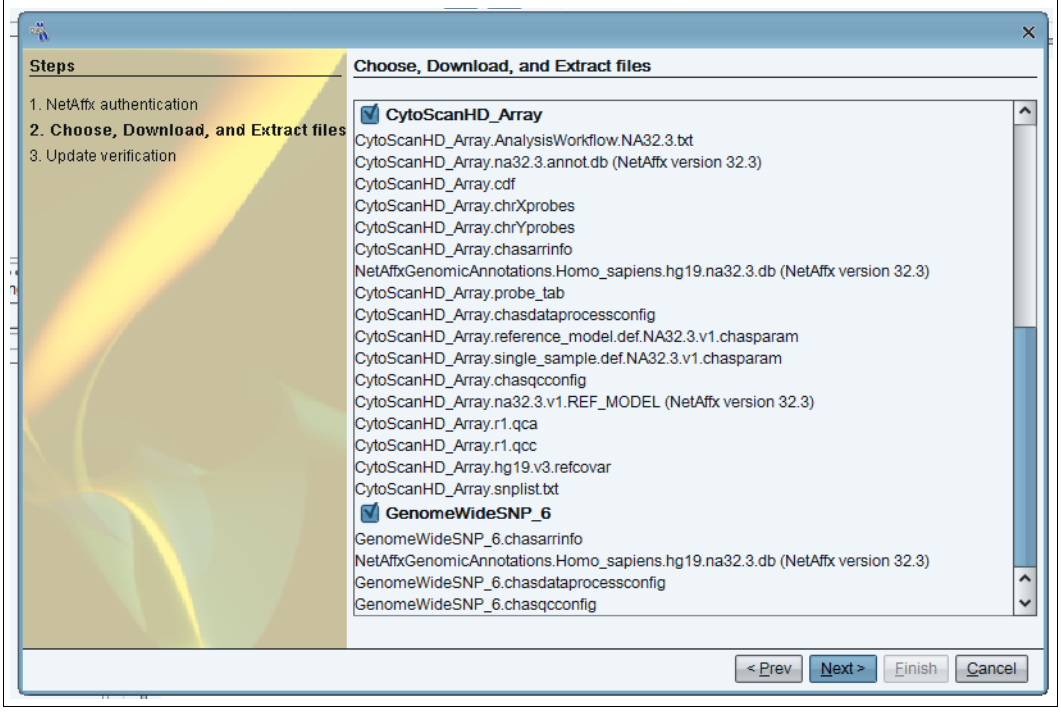


NOTE: You can also open the NetAffx Authentication window by selecting "Update Library and Annotation Files" from the Help menu.

4. Enter your NetAffx user name and password and click **Next**.

   The Choose Files window (Figure 2.7) opens with a list of library files that your local hardware can support. Some array types require more memory or processing power. If your hardware does not meet the power requirements, the library files for that array will not appear.

Figure 2.7 Choose Files dialog window



5. Select the library and annotation files you want to download.

   You can choose **Select All** to download all the available files.

6. Click **Next**.

   The Download Progress window displays the progress of the downloading and unpacking of the files.

7. Click **Finish** when the download is complete.

To download ChAS analysis files from NetAffx using the Analysis Workflow:

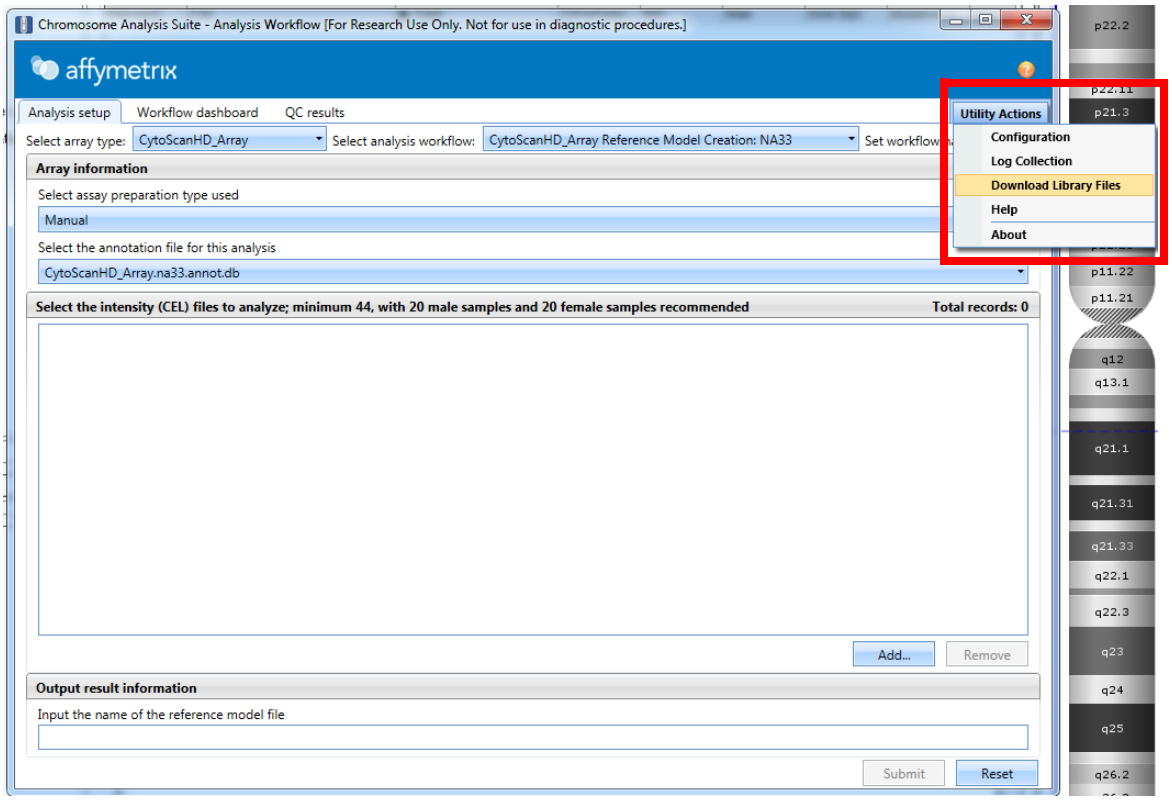
1. Go to www.affymetrix.com to log into an existing NetAffx account.

   To create a new NetAffx account, click **Register** (at the top of the NetAffx page). Alternatively, follow steps 2 and 3 below, then click the **Register Now** link. (Figure 2.9)

2. Launch the Analysis Workflow by selecting **Analysis → Perform Analysis Setup**.

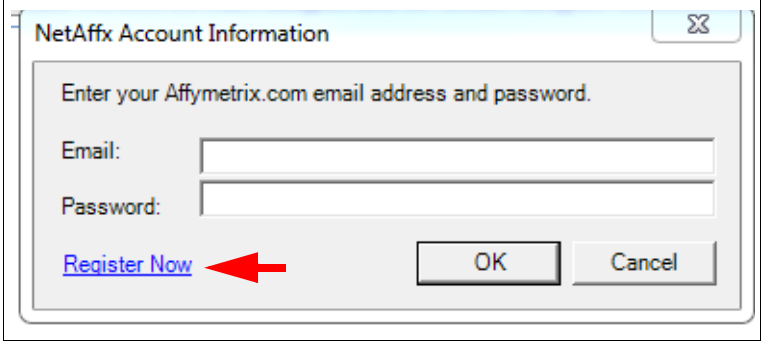
3. Select **Utility Actions → Download Library Files** (Figure 2.8)

Figure 2.8 Download Library Files



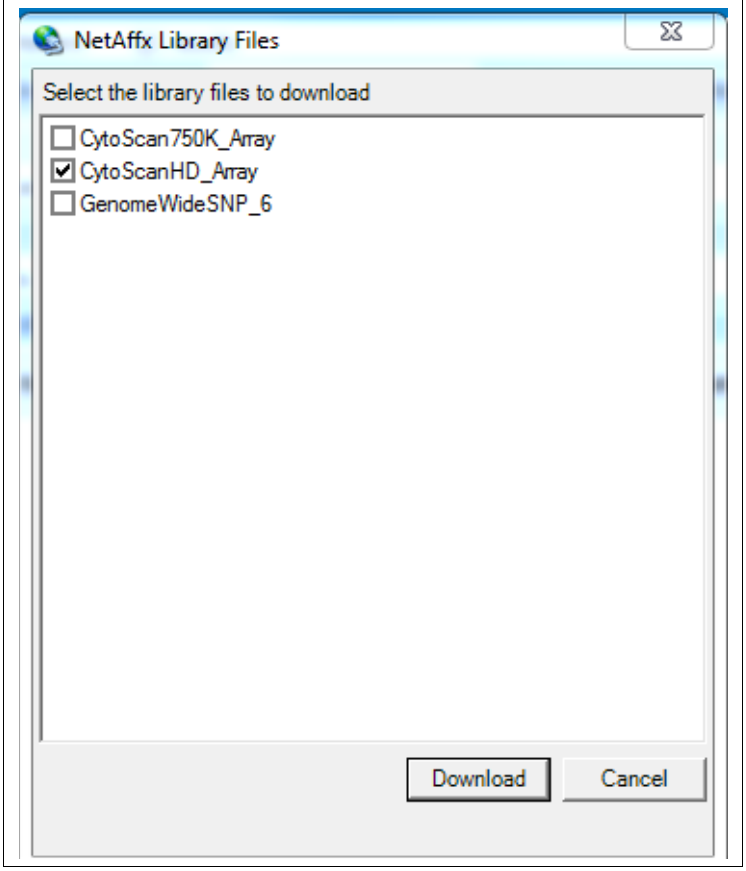
4. Enter your NetAffx user name and password (Figure 2.9), then click **OK**.

Figure 2.9 NetAffx sign in



5. Select the array type check box(es) for the analysis files that you want to download. (Figure 2.10)

Figure 2.10 Download Library Files



6. Click **Download**.

Setting Proxy Server Access

This configuration should only be done if the user's system has to pass through a proxy server to access Affymetrix NetAffx server.

In most cases, when a customer requires the use of a proxy, they can set a system-level proxy on their machine using their default internet browser and leave the setting (inside ChAS on the default choice) at **Use System Proxy**.

NOTE: You can set the Proxy Server Access through either the ChAS Browser or the Analysis Workflow.

If you require a custom Proxy Server, it is recommended that you set the Custom Proxy Server in both the Analysis Workflow (page 28) and the ChAS Browser (below).

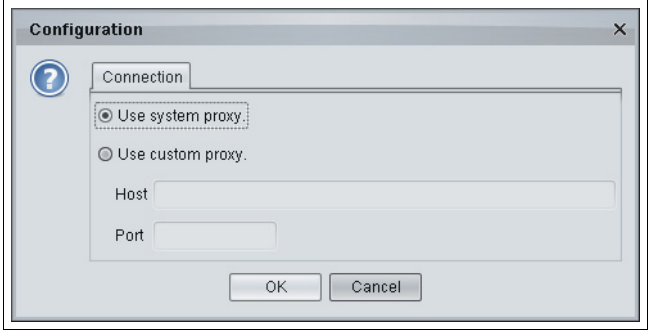
Please contact your IT department if you do not know what proxy settings you need.

To Set the Proxy Server Access through the ChAS Browser:

1. From the Preferences drop-down menu, select **Edit Application Configuration…**

   The Configuration window opens (Figure 2.11).

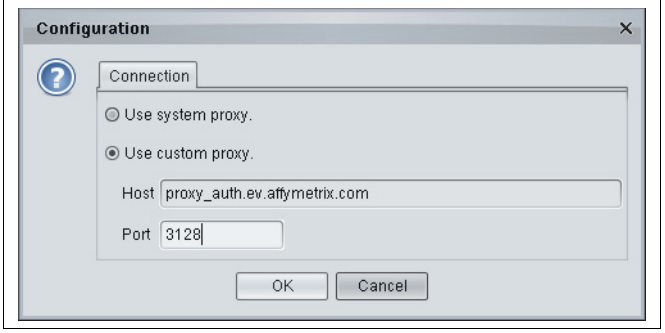
Figure 2.11 Configuration window



   By default it has the **Use System Proxy** option selected.

2. Select **Use Custom Proxy** and enter the Host and Port information (Figure 2.12).

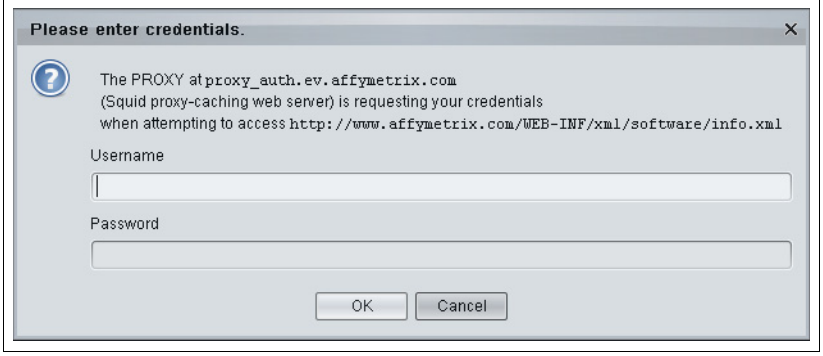
Figure 2.12 Configuration window with custom proxy information



3. Click **OK**.

   When ChAS needs to connect with the internet, a window may ask for the credentials (Figure 2.13).

Figure 2.13 Credentials window



This can happen when:

- The software is started.
- You request an update for the software and annotation files.

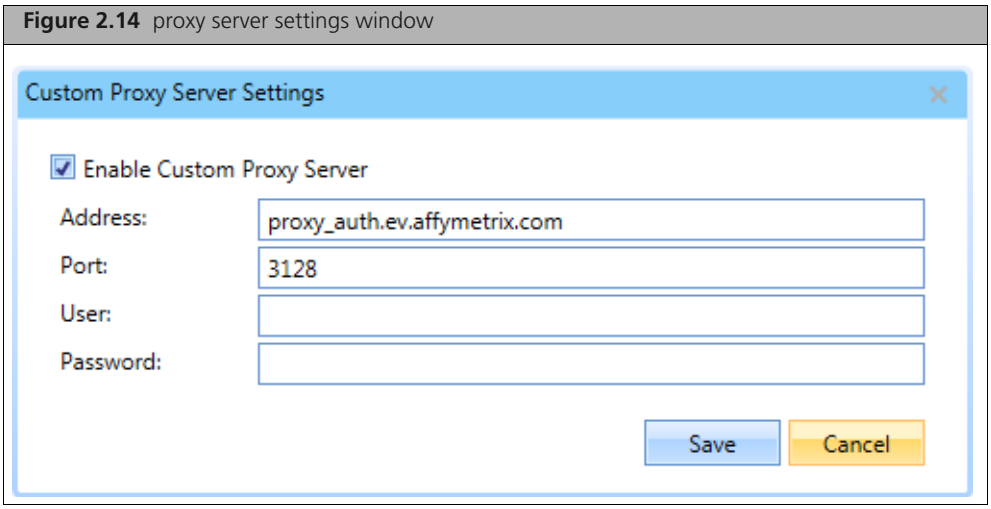
4. Enter the credentials for the proxy server and click **OK**.

NOTE: This proxy user ID and password is NOT the same ID and password used to connect to the Affymetrix NetAffx server.

To Set the Proxy Server through the Analysis Workflow:

1. From the ChAS Browser, launch the Analysis Workflow by clicking **Analysis → Perform Analysis Set Up**
2. Click **Utility Actions → Configuration**.
3. Click the **Set Custom Proxy Server** button.
4. Click the **Enable Custom Proxy Server** checkbox.
5. Enter the Host Address, Port, Username and Password. (Figure 2.14)

Figure 2.14 proxy server settings window



NOTE: This proxy user ID and password is NOT the same ID and password used to connect to the Affymetrix NetAffx server.

Un-installing ChAS

1. From the Windows Start Menu, navigate to the Windows Control Panel,
2. Navigate to the **Uninstall or change a program** (Win7/8.1) or **Add or Remove Programs** (XP) window.
3. Locate the **Chromosome Analysis Suite** application, then perform the un-install as you normally would.

TIP: You can also un-install ChAS by re-running its .exe installation file.

Getting Started with Chromosome Analysis Suite

This chapter provides some context for using ChAS. It covers:

- *Starting Chromosome Analysis Suite*
- *File Types and Data Organization in Chromosome Analysis Suite* (on page 33)
- *Basic Workflow for Cytogenetics Analysis* (on page 35)
- *Working with ChAS* (on page 40)

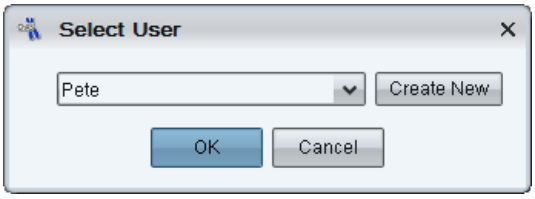
Starting Chromosome Analysis Suite

To start Chromosome Analysis Suite:

1. Double-click the ChAS icon

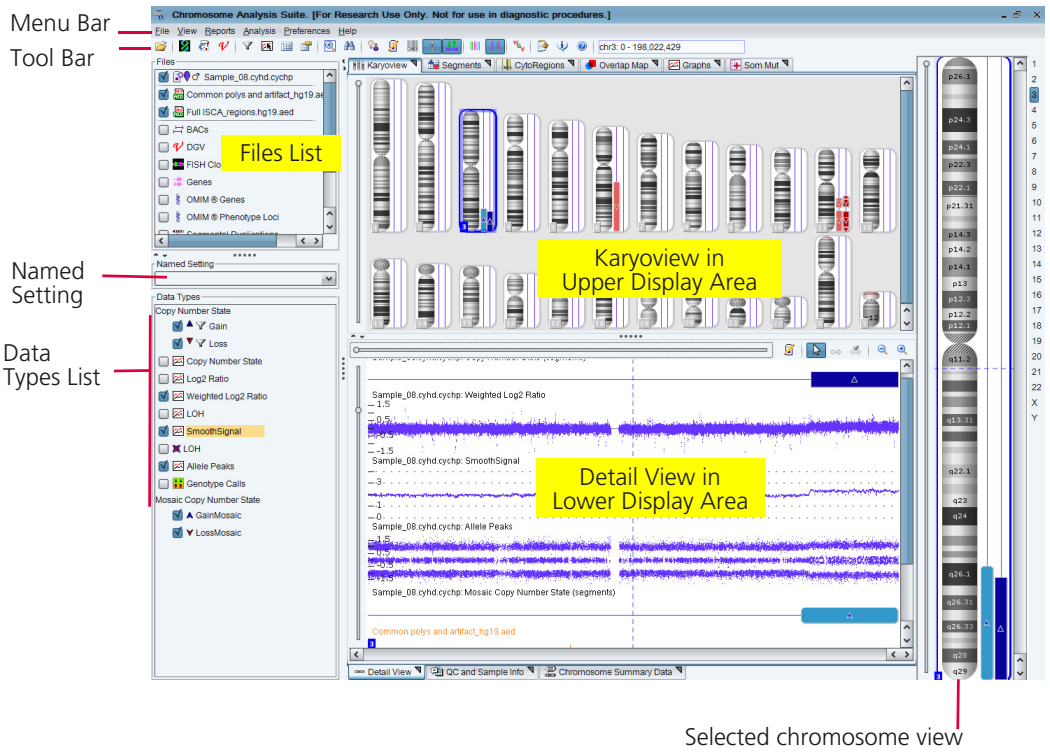
on the desktop.
   The Select User window appears. (Figure 3.1).

Figure 3.1 Select User window



2. Use the drop-down button to select a user or click Create New to create a new user profile. For more information see *Creating and Using User Profiles* (on page 290).
3. Click **OK**.
   The Chromosome Analysis Suite opens. (Figure 3.2 on page 31)

Figure 3.2 Chromosome Analysis Suite with .cychp file loaded



The ChAS browser window has the following components:

- Menu Bar - Provides access to the functions of the software.
- Tool Bar - Provides quick access to commonly used functions.
- Files List - Shows the data and annotation files that are available for display. *See Files List(on page 85)*
- Data Types List - Displays the type of data available in the files. *See Data Types List(on page 86)*
- Named Settings - Displays a list of the previously saved display settings for ChAS. *See Named Settings(on page 87)*
- Status Bar - Displays information on the status of the software, the ChAS Browser NetAffx Genomic Annotation file version, the hg version, information about the annotation or probe that the mouse pointer is nearest to in the Detail View, and the user profile name. *See Status Bar(on page 87)*
- Display Area - Displays the following data in graphical and table formats:
  - CYCHP, CNCHP, and/or OSCHP marker data
  - Detected segments
  - Region information file data
  - Reference annotations

  See Viewing Data in ChAS(on page 83) for more information.

Analysis Workflow Module

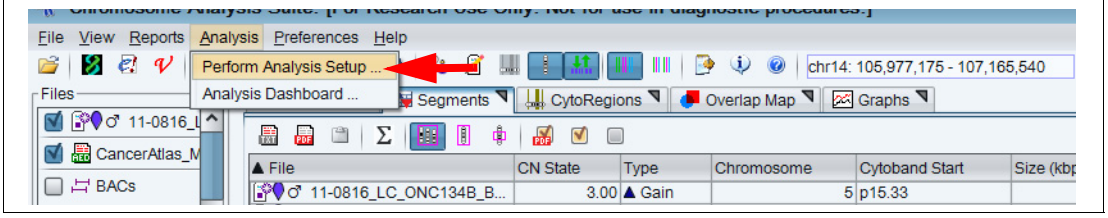
The Analysis Workflow tracks ongoing analysis tasks for ChAS. After this module is set up, its icon appears in the Windows taskbar when it is in use.

Setting Up the Analysis Workflow Module (First Time)

After installing ChAS 2.1, it is necessary to configure the data paths. The Analysis Workflow requires the following steps:

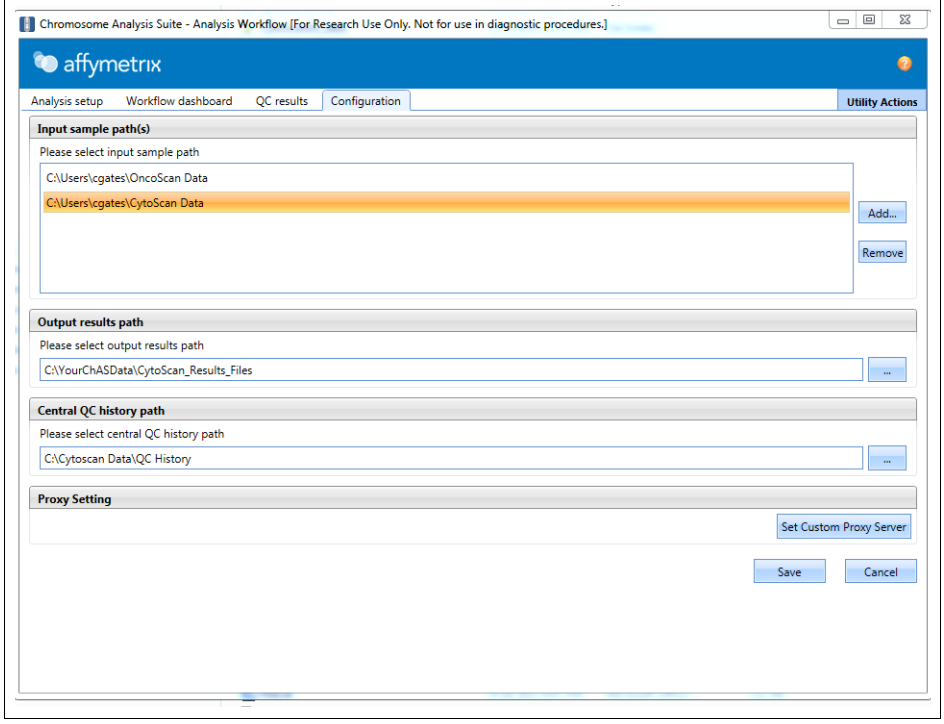
1. From the Analysis menu, select **Perform Analysis Setup.** (Figure 3.3)

Figure 3.3 Analysis drop-down menu



The Analysis Workflow Configuration window tab appears. (Figure 3.4)

Figure 3.4 Analysis Workflow Module - Setup example



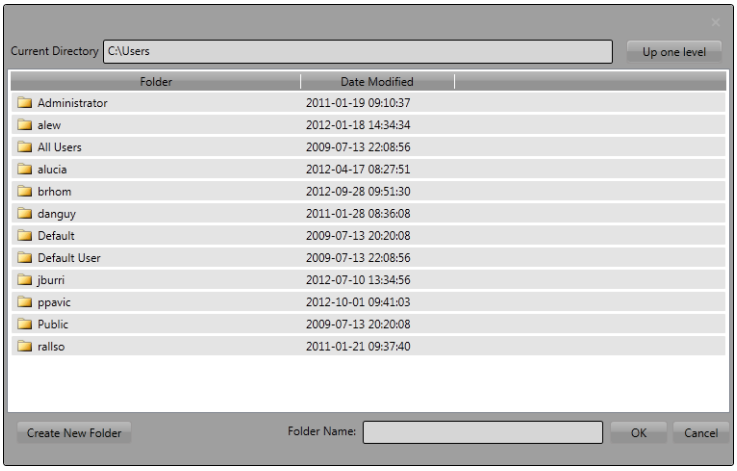
To assign an Input sample path(s):

Affymetrix recommends a minimum of 2 sample path folders.

1. Click **Add**.

The following window appears: (Figure 3.5)

Figure 3.5 Add Input sample files window



2. Click the **Up one level** button to navigate to the recommended C:\Users directory, then click the **Create new folder** button to label a new input sample folder. (Example: `C:\Users\YourName\CytoScan Data`)
3. Repeat Step 2 to create a second input folder in an easy to access area. (Example: `C:\Users\YourName\Collaborator_OncoScan Data`)

To assign an Output results path:

1. Click the browse button. Affymetrix recommends you navigate to your current ChAS data folder location, then click **Create New Folder** to create an output results path folder inside this folder. (Example: `C:\Users\YourName\CytoScan Data\CytoScan_Results_Files`)
2. Click **Save**.

To assign a Central QC history path:

1. Click the **Browse** button, then navigate to a folder in which to to store the QC history file. (For Example `C:\Cytoscan_data\`)
2. Click **Create New Folder** to create a central QC history path folder. (Example: QC_History)

File Types and Data Organization in Chromosome Analysis Suite

To fully use the capabilities of ChAS, you need to understand the ChAS file types and data organization in ChAS.

Chromosome Analysis Suite File Types

ChAS uses the following types of files.

- Data files
- Region Information files
- Support files

Files Types Supported in ChAS

Some of the data files that ChAS uses are generated by other Affymetrix software.

Table 3.1 File types used by ChAS

File Type	Created In	ChAS…
Sample file (ARR)	AGCC	Uses this information to associate sample attribute information with CEL and CYCHP or CNCHP files.
Intensity Data file (CEL)	AGCC	Analyzes the intensity data in the CEL file, then generates a CYCHP file. Note: A 64-bit system is required to analyze CytoScan® HD intensity data.
Analysis Results (CYCHP) CytoScan® HD array: CYCHP contains copy number, LOH, mosaicism, and genotype call information	ChAS	Displays results in graphical and tabular formats.
Analysis Results (OSCHP) OncoScan FFPE Assay array, OSCHP contains copy number, LOH, and somatic mutation information	OncoScan Console (OC)	Displays the probe-level analysis, segment level data, and somatic mutation data.
CopyNumber/LOH (CNCHP)	GTC	Displays probe-level analysis data and generates segment data on-the-fly.
CHP Change Archive (CHPCAR)	ChAS	Stores user-annotated Calls, Interpretations and Modifications made to CHP file Segment data.
Region Information File (BED or AED)	ChAS or Text Editor	Allows users to display their own custom data and optionally use the information to define CytoRegions or an Overlap Map. ChAS can export data in BED format for use with the UCSC Browser and other programs which understand this format.
Tab-separated values (TSV, TXT)	ChAS	Exports data in this format for use in a spreadsheet program or other user-defined uses. This format is for export only. ChAS does not import TSV or TXT files.

Region Information Files

The region information files in Browser Extensible Data (BED) and Affymetrix Extensible Data (AED) format provide lists of regions in the genome with position information and other annotations. To open a BED or AED file, click the button or select **File → Open** on the menu bar. All BED or AED files that are opened during a session will reload when you start a new session with the same user profile.

NOTE: You can also use the reference annotations to provide region information.

NOTE: You can use the Export feature to export data in existing BED files to an AED file. See *Exporting Information in AED or BED Format* (on page 224).

Analysis and Visualization Library Files

IMPORTANT: Every feature in ChAS 2.1 requires support files.

- Array-type specific Library file sets with files for running Copy Number/LOH/Mosaicism analysis and Reference Creation workflows (Analysis files)
- Files for visualizing and exporting data from xxCHP results data files.
- Reference Annotation files (Browser Annotation files named with the following format: <NetAffxGenomicAnnotations.Homo_sapiens.hgxx.naxx.db>)

Data Organization in ChAS

ChAS allows you to keep your CEL and CYCHP, CNCHP, and OSCHP files in any folder on your computer, instead of being kept in a workspace. As long as you know where the files are, you can load them from anywhere and move them around at your convenience.

IMPORTANT: Affymetrix recommends that you perform analysis operations with all files stored on a local disk drive.

Basic Workflow for Cytogenetics Analysis

IMPORTANT: The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.

NOTE: When referring to steps that apply to CYCHP, CNCHP and OSCHP, the term "xxCHP" is used.

ChAS can be used to:

- Perform probe-level analysis of CEL file data for CytoScan® HD Arrays
- Display probe-level analysis data (xxCHP) from:
  - CytoScan® Arrays (CYCHP)
  - Genome-Wide Human SNP Array 6.0 (CNCHP)
  - OncoScan FFPE Assay arrays (OSCHP)

  There are some differences in the way the ChAS handles these different types of arrays and how it treats the data from these 3 types of files.

The basic cytogenetic analysis workflow includes the following steps:

1. *Array Processing Workflow (Using AGCC)* (on page 36).
2. *Probe-Level Analysis of CEL File Data* (on page 36)
   - For the CytoScan® HD Array, this analysis is performed in ChAS and produces CYCHP files. (See *CN/LOH/Mosaicism Analysis* (on page 43)).

   NOTE: A 64-bit system (with > 8GB RAM) is required to analyze CytoScan® HD Array CEL files.

For the Genome-Wide Human SNP Array 6.0, this analysis is performed in Genotyping Console (GTC) software and produces CNCHP files. For more details, please refer to the ***GTC User Guide***.

For the OncoScan FFPE Assay files, the analysis is performed in OncoScan Console software and produces OSCHP files. For more details. please defer to the ***OncoScan Console User Guide***.

3. *Viewing Data and Features of Interest using the ChAS Display Controls* (on page 39)

Array Processing Workflow (Using AGCC)

Array processing is performed in AGCC 3.2.4 or higher for the CytoScan HD Array, OncoScan FFPE Assay, and Genome-Wide Human SNP 6.0 Array.

NOTE: You need to have the appropriate AGCC library files installed on the instrument control workstation to perform these analyses for the different array types.

The array processing includes the following steps:

1. Registering samples and arrays.
2. Washing and staining the arrays.
3. Scanning arrays and generating intensity (CEL) file data.

The following file types are produced:

- Sample (ARR files)
- DAT Files
- CEL Files
- Audit
- JPG

See the *AGCC User's Guide* for more information.

Probe-Level Analysis of CEL File Data

Cytogenetic analysis data is handled differently from genome-wide genotyping data in this step.

NOTE: You need to have the appropriate ChAS library files installed to perform these analyses for different array types. A 64-bit system is required to analyze CytoScan® CEL files.

- CytoScan® HD Array: The probe level analysis on CEL file data is performed in ChAS and produces CYCHP files which contain the data shown in Table 3.2. *See CN/LOH/Mosaicism Analysis(on page 43)*
- Genome-Wide Human SNP Array 6.0 Data: The probe level analysis on CEL file data is performed in GTC and produces the CNCHP file data types shown in Table 3.2. See the *GTC User Guide* for more information.
- OncoScan FFPE Assay Data: The probe level analysis on CEL file data is performed in OncoScan Console and produces the OSCHP file data types show in Table 3.2. See the *OncoScan Console User Guide* for more information.

Table 3.2 Analysis results data by array type

	CytoScan® HD Array[1]	Genome-Wide Human SNP Array 6.0[2]	OncoScan FFPE Assay[3]
Graph Data for the individual CN and SNP probes			
Copy Number State	yes	yes	yes
Log2 Ratio	yes	yes	yes
Weighted Log2 Ratio	yes	no	yes
LOH	yes	yes	yes
Allele Peaks	yes	no	no
Allele Difference	no	yes	yes
Genotype Calls	yes	no	no
Smooth Signal	yes	yes	yes
Somatic Mutations	no	no	yes
B-allele Frequency	no	no	yes
Segment data			
Gain and Loss segments based on runs of aberrant Copy Number State data	yes	yes	yes
Mosaic Gain and Loss segments of non-integer Copy Number States between CN=1 and CN=3	yes	no	no
LOH (Loss of Heterozygosity) based on runs of SNPs where heterozygote calls are absent	yes	yes	yes

* **CYCHP for CytoScan® HD, CNCHP for Genome-Wide Human SNP 6.0 Array, and OSCHP for OncoScan FFPE Assay.**

1) For more details on CytoScan® HD Array data, see Table 7.1 on page 106.

2) For more details on Genome-Wide SNP Array 6.0 data, see Table 7.2 on page 107.

3) For more details on OncoScan FFPE Assay Data, see Table 7.3 on page 107.

NOTE: Segment types drawn with flat ends (Gain and Loss) are the result of algorithms which can ascertain precise marker-to-marker breakpoints. Segment Types drawn with rounded ends (LOH, GainMosaic, LossMosaic) are the output of algorithms which closely approximate breakpoints based on the data.)

Loading Data into ChAS for Display

You perform the same steps for the different types of array data (CYCHP, CNCHP, or OSCHP), but ChAS handles the 3 types of data differently.

CytoScan® HD Array (CYCHP files)

When loading CYCHP files into ChAS for viewing, the software:

1. Selects the segments in the CYCHP file to display as segments.
2. Applies any smoothing or joining that would alter the length and other properties of segments.
3. Displays the segments and graph data:
   - Segment Data
     - Copy Number Gain/Loss
     - Loss of Heterozygosity (LOH)
     - Mosaic Gain/Loss
   - Graph Data
     - Copy Number State
     - Log2 Ratio
     - Weighted Log2 Ratio
     - Smooth Signal
     - Loss of Heterozygosity (LOH)
     - Allele Peaks
     - Genotype calls

Genome-Wide SNP Array 6.0 (CNCHP files)

When loading CNCHP files into ChAS for viewing, the software:

1. Performs segment generation by analyzing the CN and LOH graph data in the CNCHP file.
2. Applies any smoothing or joining that would alter the length and other properties of Copy Number segments.

   In GTC software, these steps were performed in the Segment Reporting Tool.
3. Displays the segments and graph data:
   - Segment data
     - Copy Number Gain/Loss
     - Loss of Heterozygosity (LOH)
   - Graph Data
     - Copy Number State
     - Log2 Ratio
     - Allele Difference
     - SmoothSignal

- Loss of Heterozygosity (LOH)

OncoScan FFPE Assay (OSCHP files)

When loading OSCHP files into ChAS for viewing, the software:

1. Selects the segments in the OSCHP file to display as segments.
2. Displays the segments and graph data:
   - Segment data
     - Copy Number Gain/Loss
     - Loss of Heterozygosity
   - Graph Data
     - Copy Number State
     - Log2 Ratio
     - Weighted Log2 Ratio
     - Allele Difference
     - B-allele frequency
     - Smooth Signal
     - Loss of Heterozygosity
     - Somatic Mutations

Viewing Data and Features of Interest using the ChAS Display Controls

ChAS provides several options for viewing and studying the loaded CYCHP, CNCHP, or OSCHP data:

- Graphic Displays

  See Displaying Data in Graphic Views(on page 91)
- Tables

  See Displaying Data in Table Views(on page 228)

After the data is loaded, you can:

- Filter the segments by Segment Parameters to hide segments that do not meet your requirements for significance.

  See Filtering Segments(on page 142)
- Select a region information file for use as a CytoRegion file and:
  - Perform differential filtering for segments in CytoRegions and in the rest of the genome.
  - Display only segments that appear in CytoRegions using Restricted Mode.

  See Using CytoRegions(on page 178)
- Select a region information file for use as an Overlap Map and use the Overlap filter to identify or conceal segments that appear in the Overlap Map regions.

  See Using the Overlap Map(on page 188)
- Add selected features of the genome to new or existing Region (AED) files, and edit annotation data on existing annotations. (To open a BED or AED file, click the button or select **File → Open** on the menu bar.)

  See Creating and Editing AED Files(on page 196)
- Prepare reports on your findings by exporting graphics and table data in PDF and other formats.

  See Reporting Results(on page 269)
- Save setups of ChAS for different tasks in user profiles and named settings.

  See User Profiles and Named Settings(on page 289)

Working with ChAS

This chapter covers:

- *Accessing Functions in ChAS* (on page 40)
- *Changing the Size of Panes in ChAS* (on page 40)
- *Opening Panes in Separate Windows* (on page 41)

Accessing Functions in ChAS

Commands in ChAS can be accessed in multiple ways. You can access commands from:

- Main menus
- Tool bar
- Right-click options in the:
  - Files list
  - Data Types list
  - Karyoview
  - Selected Chromosome View
  - Detail View
  - Table headers

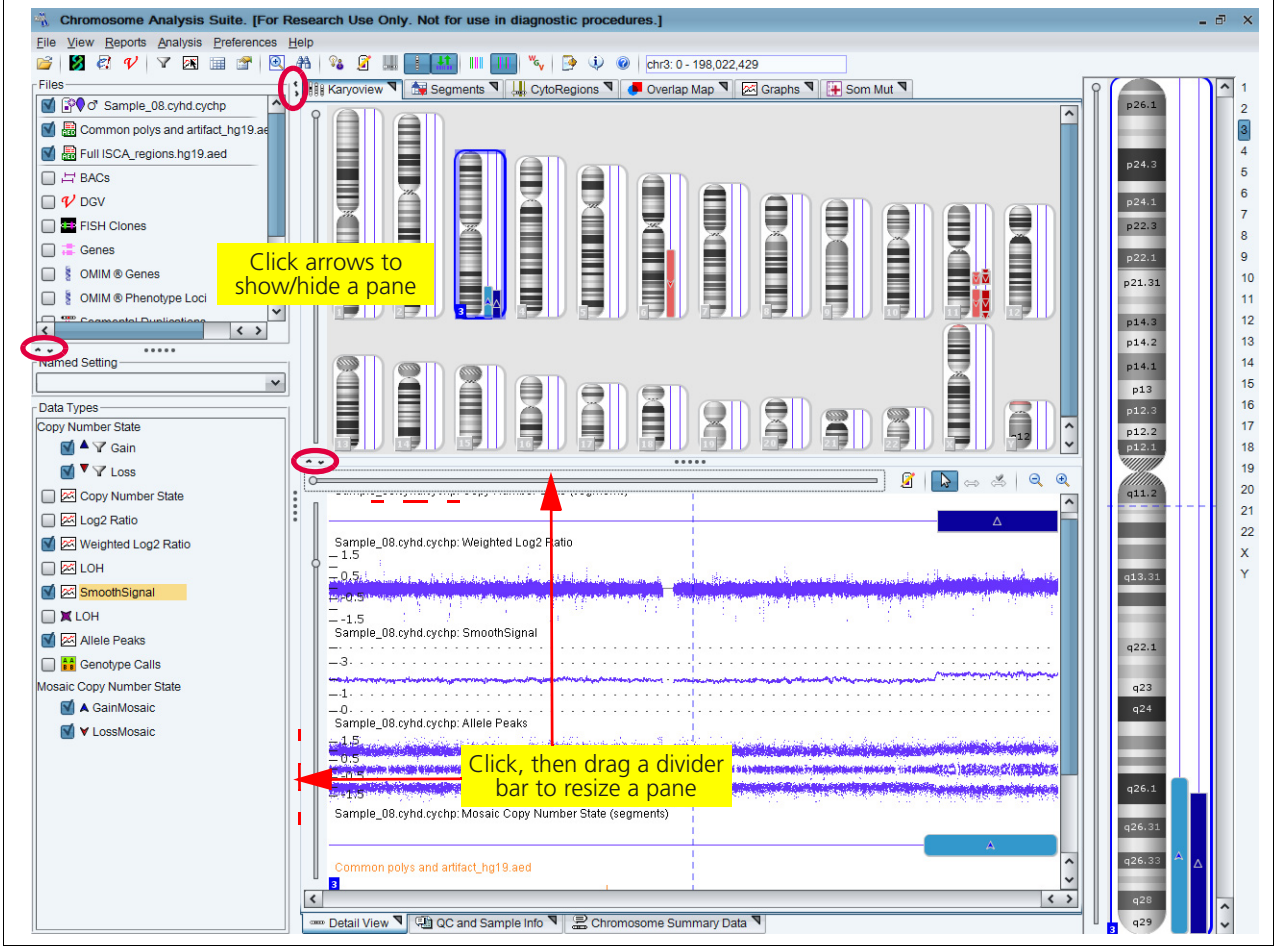
The alternatives for a particular function are described in the step-by-step instructions for using the function.

Changing the Size of Panes in ChAS

You can change the size of the panes in the ChAS window (Figure 3.6):

- Click and drag the dividers between panes.
- Click the arrows in the dividers (or) to hide or maximize an entire pane.

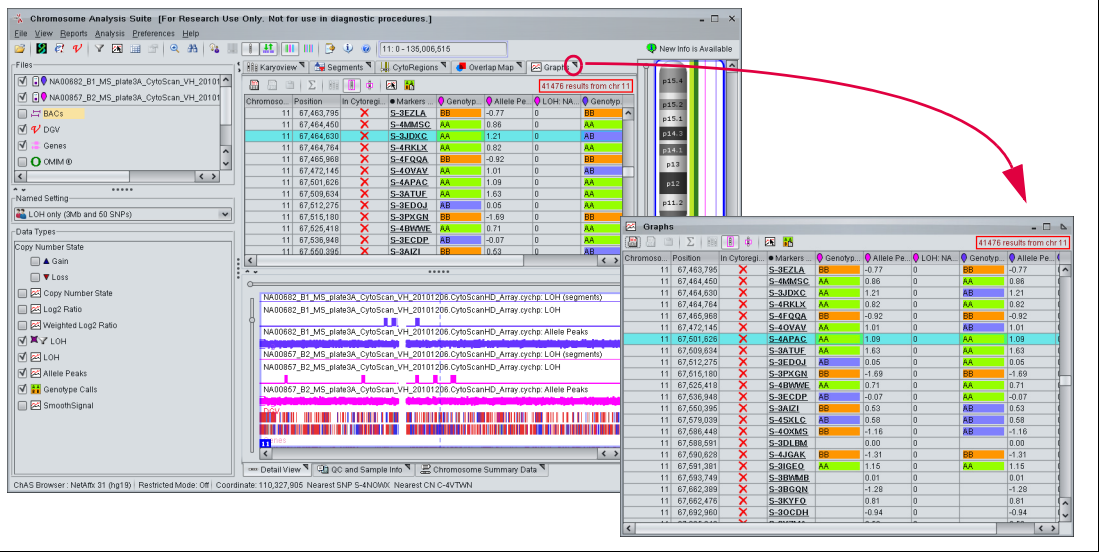
Figure 3.6 Resize or show/hide panes



Opening Panes in Separate Windows

You can display a pane in a separate window (Figure 3.7 on page 42) by clicking the icon on the tab. To close the window and return the information to the tab panel, click the icon in the window.

Figure 3.7 Graphs pane opened in new window



CN/LOH/Mosaicism Analysis

ChAS analyzes the intensity data (CEL file) from CytoScan® HD Arrays. The software performs a single sample analysis which compares the data in a CEL file to a previously created reference file, using analysis parameters specified in the .chasparam file. The analysis generates a CYCHP data file that you load and view in ChAS. The analysis detects segments that exhibit:

- Copy number state gain or loss – Regions of integer copy number gain or integer copy number loss
- Mosaic gain and loss – Regions of non-integer copy number gain or loss (CN states between 1 and 3)

NOTE: Mosaicism analysis is currently only available for CytoScan® Array CEL files. However, CnCHP files contain the Smooth Signal data type which displays non-integer copy number changes.

- Loss of Heterozygosity (LOH) – Regions where the preponderance of SNPs do not display heterozygosity

For more details on loading and viewing CYCHP data, see *Loading Data* (on page 66).

IMPORTANT: The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.

Affymetrix provides .chasparam and reference files as part of complete Library file packages. You can also create your own reference file using ChAS.

NOTE: You can also load Genome-Wide Human SNP Array 6.0 CNCHP and OncoScan FFPE Assay OSCHP files into ChAS to display and detect Copy Number and Loss of Heterozygosity segments. See *Loading Data* (on page 66).

IMPORTANT: Affymetrix recommends that you perform analysis operations with all files stored locally.

This chapter has the following sections:

- *Performing a Single Sample Analysis* (on page 49)
- *Creating Custom QC Views in the Analysis Workflow* (on page 59)

Single Sample Analysis

Single Sample Analysis compares the values in one or more user-selected CEL files with the values in a reference file that is created from a set of sample files. You can use the factory-loaded reference file provided by Affymetrix or create your own (For more details, see *Creating Custom QC Views in the Analysis Workflow* (on page 59))

This section covers:

- *ChAS and NetAffx Analysis File Compatibility* (on page 44)
- *Introduction to Single Sample Analysis* (on page 44)
- *LOH Segments on the X and Y Chromosomes* (on page 47)
- *Performing a Single Sample Analysis* (on page 49)

ChAS and NetAffx Analysis File Compatibility

The table below (Table 4.1) lists the compatibility between ChAS and NetAffx Analysis file versions for the CytoScan® HD Array.

NOTE: ChAS automatically prevents you from selecting an incompatible NetAffx Analysis file version for analysis or when viewing analysis results.

Table 4.1 Compatibility between ChAS version and NetAffx analysis file set version

	ChAS Software Version						
CytoScan® HD Array NetAffx Analysis File Set Version	ChAS 2.1	ChAS 2.0	ChAS 1.2.2/ 1.2.1	ChAS 1.2	ChAS 1.1	ChAS 1.0.1	ChAS 1.0
NA33(hg19)	Yes	No	No	No	No	No	No
NA32.3(hg19)	Yes	Yes	No	No	No	No	No
NA32.1(hg19)	No	No	Yes	No	No	No	No
NA32(hg19)	No	No	No	Yes	No	No	No

NOTE: Refer to the ChAS Release Notes for data equivalency information between the ChAS software and the Library file versions used to create CHP files.

Introduction to Single Sample Analysis

Single Sample Analysis requires:

- ChAS analysis files for the array

  See *Analysis File Download* (on page 23).

- A previously created reference model file

  You can use the reference model file provided by Affymetrix or create your own using your own CEL file data and the Reference File creation function. The Reference Model file in the CytoScan HD™ Array set of NA33 and NA32.x Analysis includes 380 microarrays which were run as part of a larger set of microarrays by nine operators processing ~48 unique samples in two rounds each, with randomization of the placement of sample DNAs across the PCR plates and randomization of the reagents and instruments used. The source DNA includes:

- 284 HapMap samples including at least one replicate of each of 270 HapMap samples: 90 from each of the Yoruban, Asian, and Caucasian ethnic groups, from cell-line derived DNAs from the Coriell Institute of Medical Research.
- 96 DNA samples from blood of phenotypically healthy male and female individuals obtained from BioServe Biotechnologies.

- CEL file data

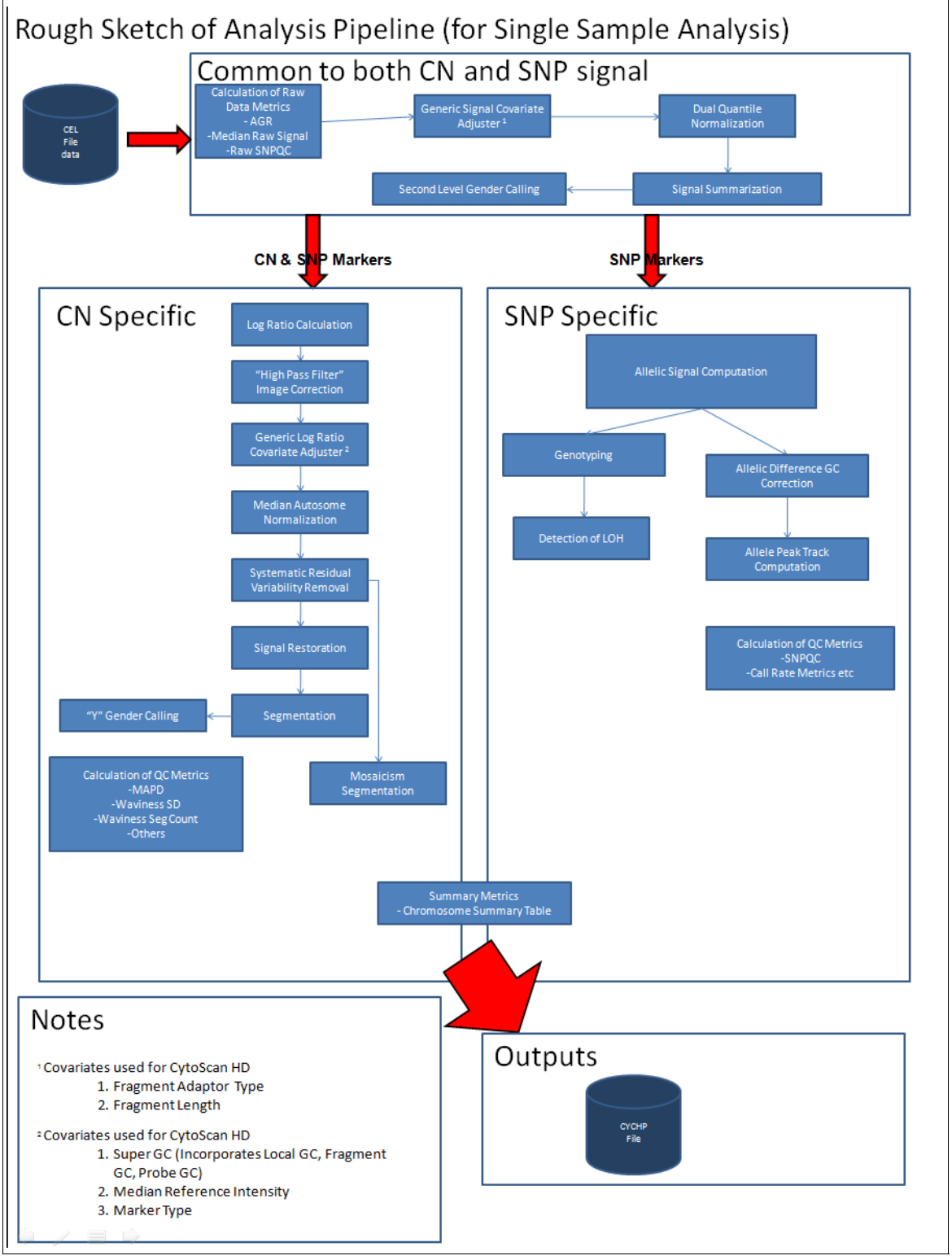
During the analysis, ChAS generates CYCHP files with:

- Graph Data
  - Copy Number State
  - Log2 Ratio
  - Weighted Log2 Ratio
  - LOH
  - Allele Peaks
  - Smooth Signal
  - Genotype Calls
- Segment Data
  - Copy Number Gain/Loss
  - Mosaicism Gain/Loss
  - Loss of Heterozygosity (LOH)

The CYCHP files can be loaded into ChAS for viewing and study.

The figure below (Figure 4.1 on page 46) shows an overview of single sample analysis for the CytoScan® HD Array.

Figure 4.1 CytoScan® HD Single Sample Analysis



Copy Number Segments on the X and Y Chromosomes

The expected Copy Number State on the X chromosome in normal males is not constant over its entire length. This is due to the structure of the sex chromosomes, and the fact that they share extensive homology with each other only in the Pseudo Autosomal Regions (PARs) that they each have at either end. PAR1 is at the top of the p-arm and PAR2 at the bottom of the q-arm.

Markers occurring in the PAR regions are mapped exclusively to the X Chromosome. Therefore, in normal males the PAR regions of the X are expected to be CN=2 (probes on the X and Y contribute to the signal), while the rest of the Chr X is expected CN=1 for normal males. As a result, we treat the two X PARs in males as independent units (CN=2 expected) from the rest of the X chromosome (CN=1 in males) when generating Copy Number Segments.

Aberrant segments that cross PAR/non-PAR boundaries may be normalized into one segment if they have equivalent type (Gain or Loss) and CN State. During this normalization process, ChAS will not combine an aberrant (Gain or Loss) segment with a normal segment across PAR/non-PAR boundaries, even if they have the same CN State. If smoothing is subsequently applied, aberrant segments with different Copy Number State may be combined. If joining is subsequently applied, aberrant segments separated by a non-aberrant segment may be combined.

Because only Y-specific probes are mapped to the Y chromosome, the expected state of the entire Y chromosome is 1 for males and is 0 for females.

Mosaic Copy Number Segments on the X Chromosome

The expected Copy Number State on the X chromosome in normal males is not constant over its entire length. This is due to the structure of the sex chromosomes and the fact that they share extensive homology with each other only in the Pseudo Autosomal Regions (PARs) that they each have at either end. PAR1 is at the top of the p-arm and PAR2 at the bottom of the q-arm.

Markers occurring in the PAR regions are mapped exclusively to the X Chromosome. Therefore, in normal males the PAR regions of the X are expected to be CN=2 (probes on the X and Y contribute to the signal), while the rest of the Chr X is expected CN=1 for normal males.

Mosaic Segments whose boundaries start and end entirely in one of the PAR regions will use CN=2 as normal to determine the type (GainMosaic or LossMosaic) of Mosaic segment to draw.

Because the Mosaicism algorithm can generate segments which cross the PAR boundaries, Mosaic Segments that touch the non-PAR region of the X chromosome use the gender call of the sample to determine the Type of Mosaic segment to draw.

Because only Y-specific probes are mapped to the Y chromosome, the expected state of the entire Y chromosome is 1 for males and is 0 for females.

IMPORTANT: Because the Mosaicism algorithm only calls segments between CN=1 and CN=3, LossMosaic segments on the non-PAR regions of sex chromosomes in males will NOT be present in this version of software. Use of the SmoothSignal data track is recommended for analyzing male samples for Mosaicism on the sex chromosomes.

Mosaic segment detection is not provided for the Y chromosome. However, the allele peaks and smooth signal tracks can be used to assess mosaicism.

LOH Segments on the X and Y Chromosomes

CytoScan® HD Array

For normal XY male samples, the X chromosome will have single-copy based LOH calls (CN = 1). Male samples with more than one X chromosome (for example, XXY) may have LOH calls on the X chromosome, depending on the constitution of the X chromosomes' origins.

The tables below briefly describe how the array-specific algorithms call LOH segments for the X or Y chromosome.

Table 4.2 Expected LOH calls on the X and Y chromosomes for the CytoScan® HD and 750K Arrays

	X Chromosome	Y Chromosome
Normal male sample (XY)	LOH calls that are single copy-based LOH call (CN = 1).	No LOH calls are made for the Y chromosome. Genotype calling is not performed on the Y chromosome.
Male sample with multiple X chromosomes (for example, XXY)	LOH calls are possible, depending on the constitution of the X chromosomes' origins.	
Normal female sample (XX)	LOH calls are possible, depending on the constitution of the X chromosomes' origins.	
Female sample with a single X chromosome (X0)	LOH calls on X regions which have only a single copy. Heterozygous SNP genotypes are possible, but are due to the low inherent Heterozygote call error rate noise, not the true presence of two alleles.	

Table 4.3 Expected LOH calls on the X and Y chromosomes for the Genome-Wide Human SNP Array 6.0

		Y Chromosome
Normal male sample (XY)	LOH calls on the non-PAR region of the X chromosome resulting from "forced" homozygote-only calls due to the presence of the Y chromosome. Heterozygous calls are ignored on the X chromosome in males.	LOH calls that are due to single copy genotyping calls (CN = 1).
Male sample with multiple X chromosomes (for example, XXY)	LOH calls are possible, depending on the constitution of the X chromosomes' origins. SNP genotypes are not constrained to homozygous calls. Heterozygous calls are ignored on the X chromosome in males.	
Normal female sample (XX)	LOH calls are possible, depending on the constitution of the X chromosomes' origins.	LOH analysis is not performed on the Y chromosome since it is assumed that there not substantial Y chromosomal material.
Female sample with a single X chromosome (X0)	LOH calls on X regions with only a single copy. Heterozygous SNP genotypes are possible, but are due to the low inherent Heterozygote call error rate noise, not the true presence of two alleles.	

Table 4.4 Expected LOH calls on the X and Y chromosomes for the OncoScan FFPE Assay

	X Chromosome	Y Chromosome
Normal male sample (XY)	LOH calls that are single copy-based LOH call (CN = 1).	No LOH calls are made for the Y chromosome.
Male sample with multiple X chromosomes (for example, XXY)	LOH calls are possible where there is either loss or low heterozygosity.	
Normal female sample (XX)	LOH calls are possible depending on the constitution of the X origins or in regions of either loss or low heterozygosity.	
Female sample with a single X chromosome (X0)	LOH regions on X which have only a single copy. Will be called LOH where there is single copy X.	

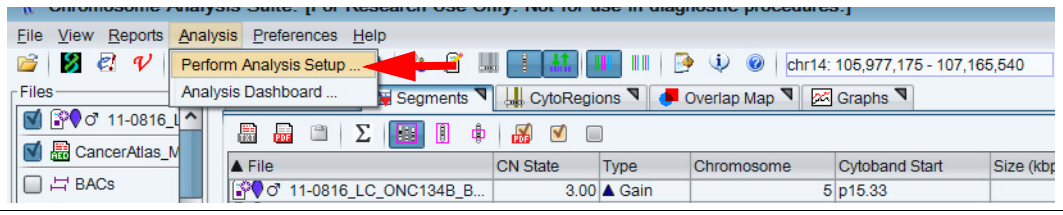
Performing a Single Sample Analysis

NOTE: Many of the options in the following section only need to be selected once, as these options are retained for your convenience in future single sample analysis runs.

To perform a single sample analysis:

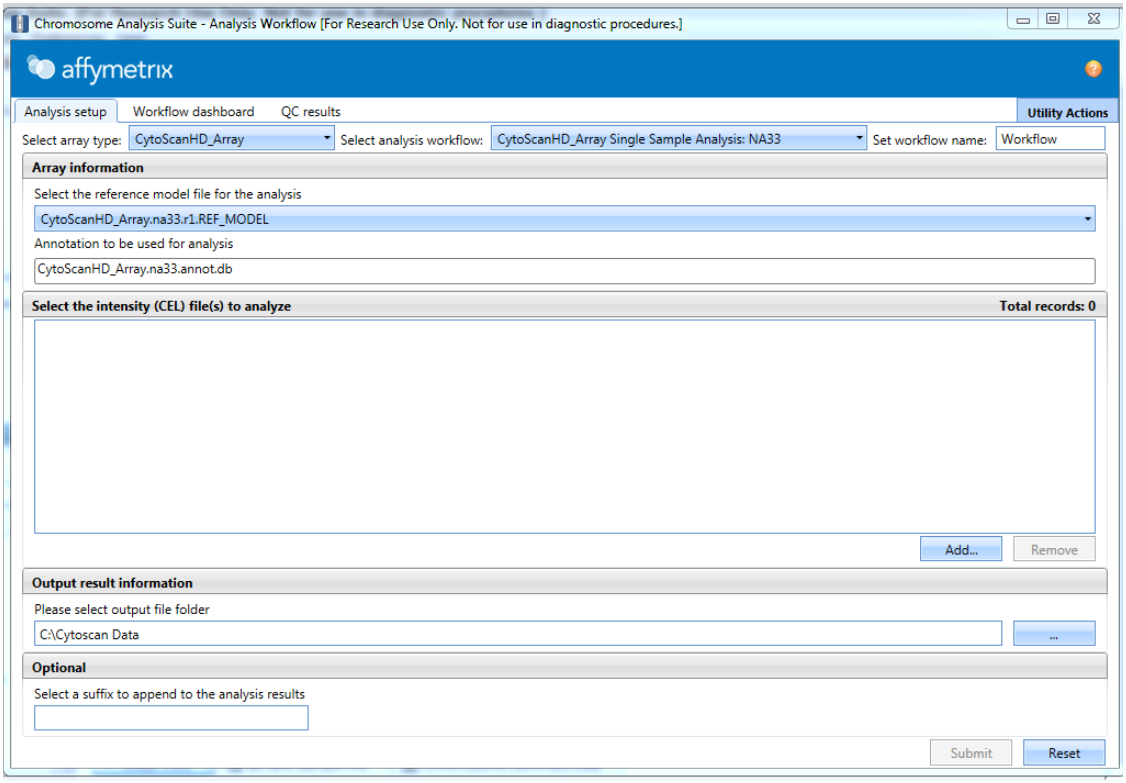
1. From the Analysis menu, select **Perform Analysis Setup.** (Figure 4.2)

Figure 4.2 Analysis drop-down menu



The Analysis Workflow window tab opens (Figure 4.3).

Figure 4.3 Analysis Workflow



2. From the *Select array type* drop-down list, click to select an array type (Example: CytoScanHD_Array)

NOTE: Once you select the array type, analysis workflow, and reference model file, then the annotation file will be auto selected for you based on your earlier selections.

NOTE: The *Select array type* drop-down list includes only the array types for which library (analysis) files have been downloaded from NetAffx or copied from the Library package provided with the installation.

3. From the *Select analysis workflow* drop-down list, click to select an analysis workflow. (Example CytoScanHD_Array Single Sample Analysis: NA33 or higher)
4. By default, the *Set workflow name* is "Workflow". Click Workflow (upper right) to enter a different workflow name.
5. From the *Select the reference model file for the analysis* drop-down list, click to select a reference model file for the analysis. (Example: CytoScanHD_Array.na33.r1.REF_MODEL).

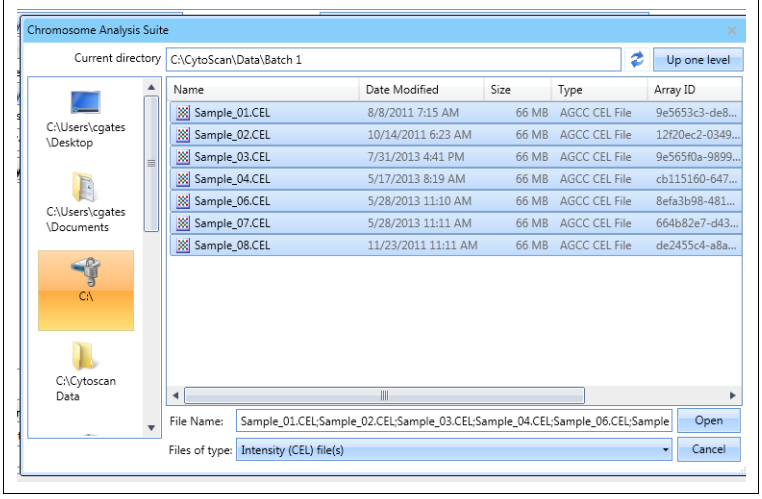
NOTE: For Single Sample Analysis, the *Annotation to be used for analysis* field is auto-populated and set with the annotation filename used when the reference model file was generated.

6. At the *Select the intensity (CEL) file(s) to analyze* pane, click Add....

NOTE: The Workflow Analysis window retains the drop-down selections used in your last submitted analysis. However, it does not display a previously set workflow name, CEL files to analyze list, or any suffix used to append your last analysis results. These 3 fields must be completed again.

The following window appears: (Figure 4.4)

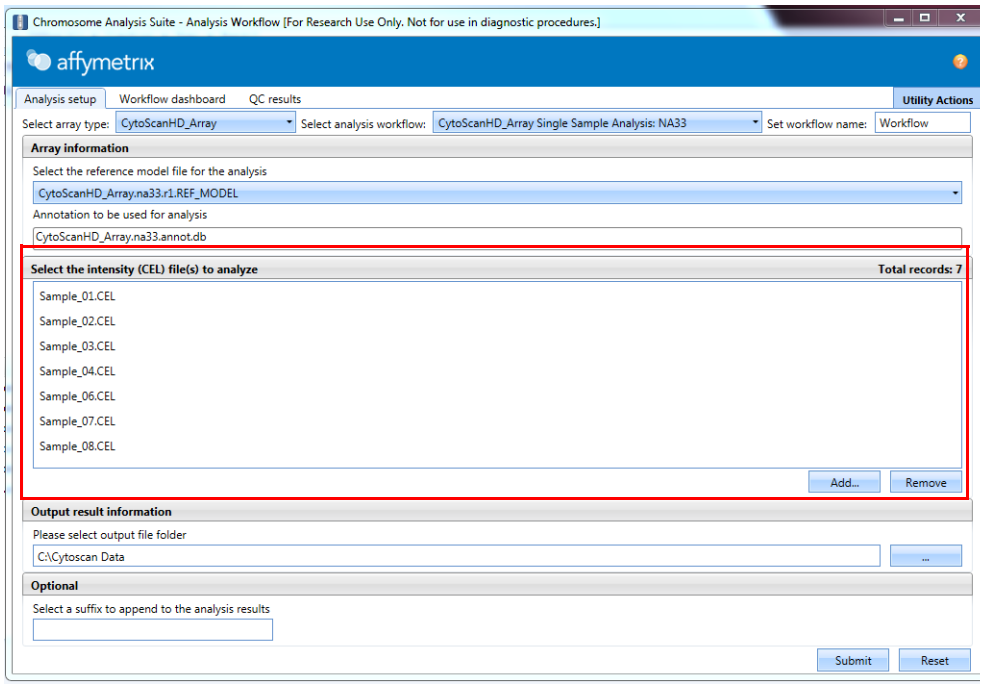
Figure 4.4 CEL file folder



7. If your CEL files are located somewhere other than your input path location, navigate to the desired folder. Single click, Ctrl click, or Shift click (to select multiple .CEL files).
   CEL files that are **in use** are locked and cannot be selected until their processing that is "*currently in progress*" is complete.
8. Click **Open**.

The *Select the intensity (CEL) file(s) to analyze* pane is now populated with your .CEL files. (Figure 4.5)

Figure 4.5 Select the intensity (CEL) file(s) to analyze pane



To remove a .CEL file from this list, click to highlight it, then click Remove.

9. At the *Output result information* pane, confirm the path shown for your output file folder. To change the current path/folder, click ... to select a different output path/folder.

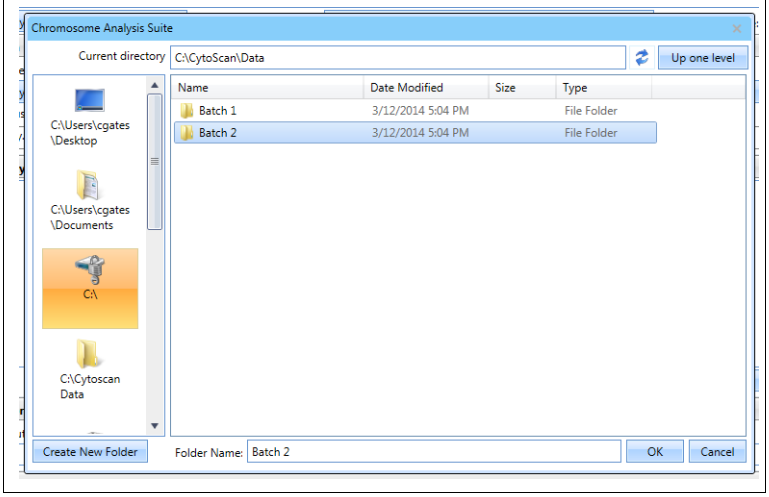
TIP: To better organize your output results, you can add sub-folders to your assigned output result path/folder.

To add sub-folders to your assigned result path/folder:

- Click ... to return to your assigned output path and/or folder.
- Click **Create New Folder**.
- Enter a sub-folder name.
- Click **OK**. Repeat the above steps to add more sub-folders.

The newly created sub-folders now appear in the output result information window. (Figure 4.6)

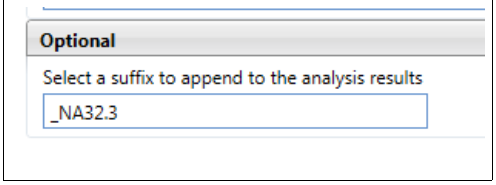
Figure 4.6 Output result information with sub-folders example



10. If you are using a previously analyzed CEL file(s) to verify new CHP data (against CHP data generated from previous versions of ChAS and Library files), you may want to use a suffix to append the new resulting CHP file(s). To do this, click inside the *Select a suffix to append to the analysis results* field to enter an appending file suffix. (Figure 4.7)

IMPORTANT: If you are saving the same .CYCHP file into the same output file folder that contains your originally run .CYCHP file, a "1" is automatically added into the filename (in addition to any suffix you may add) to differentiate the 2 runs of identical .CEL file names. Example: *na33(1).cyhd.cychp*

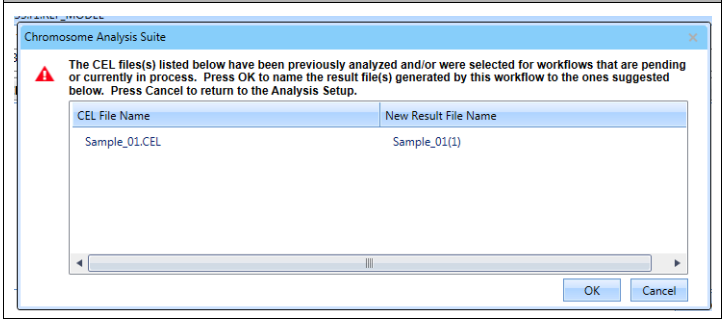
Figure 4.7 Adding a suffix



11. Click Submit.

If the following warning message appears (Figure 4.8), click **OK** to acknowledge it.

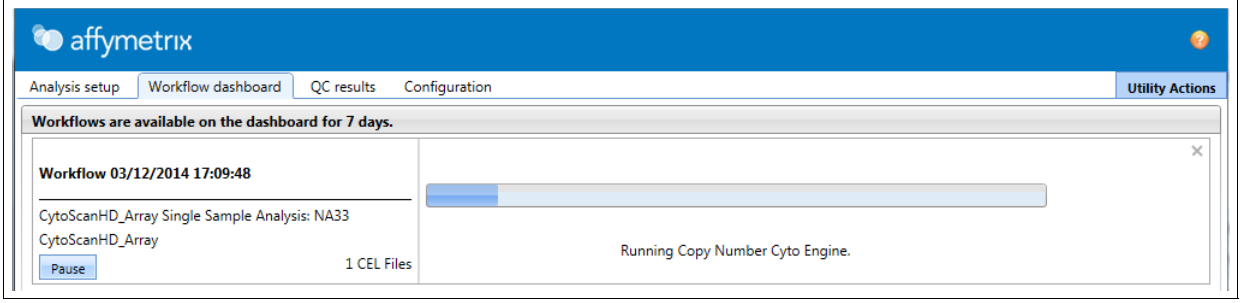
Figure 4.8 .CEL warning message



The Workflow dashboard window appears and your annotation files begin to load. (Figure 4.9).

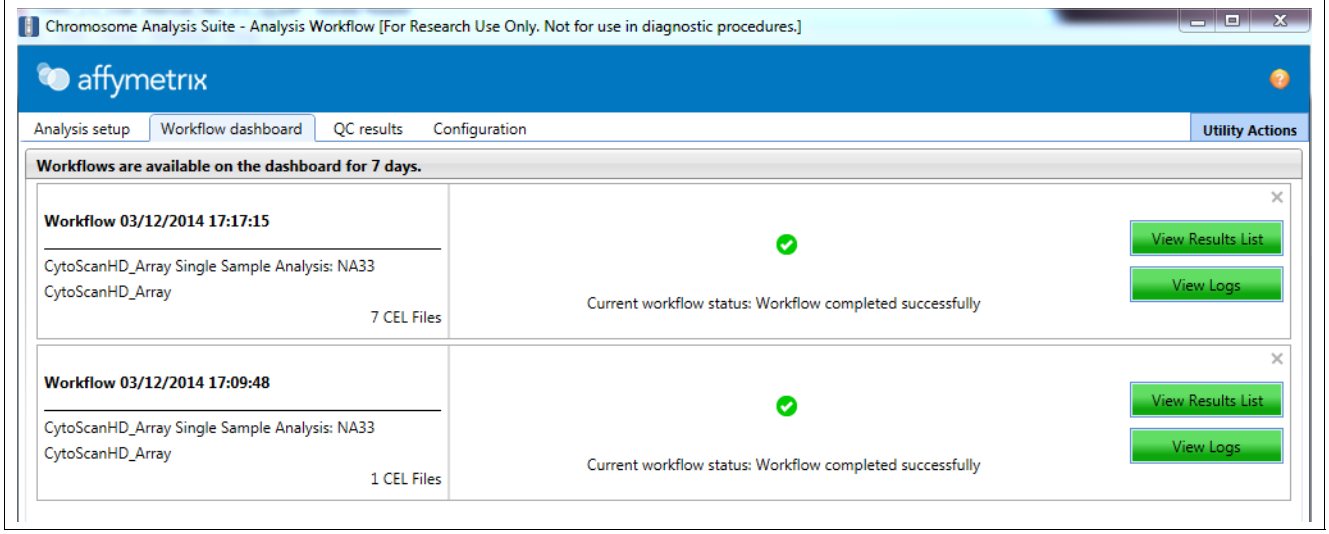
The Analysis Workflow Dashboard tracks ongoing analysis tasks for ChAS. It also delivers the results of analyses and restarts the Browser if it was shut down to free up memory for the analysis.

Figure 4.9 .CEL files loading inside the Workflow dashboard



After loading is complete, a Workflow completed successfully message appears. (Figure 4.10)

Figure 4.10 Workflow Dashboard example with multiple Single Samples loaded



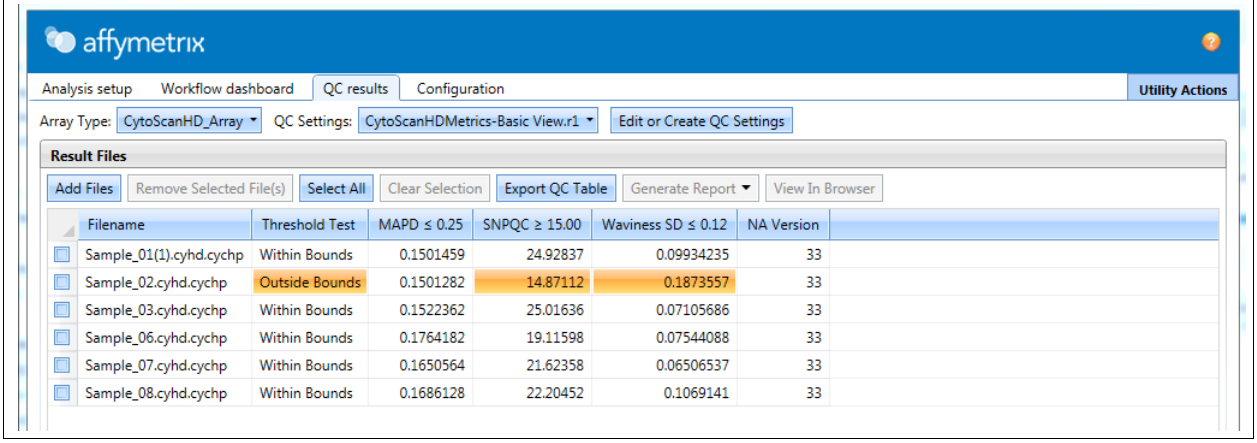
NOTE: The View Logs button will access the algorithm pipeline logs which may be useful if you have a Workflow that fails to complete.

12. Click to choose the analysis you want to view, then click View Results List.

The QC Results tab window appears showing the Basic View QC settings (Figure 4.11). A Detail View QC setting, which provides more columns of data, is also available in the QC settings drop down list. (Figure 4.11)

NOTE: QC parameters can also be viewed in the ChAS Browser see setting QC parameters in ChAS Browser.

Figure 4.11 QC Results window tab



13. Click each sample's checkbox or click the **Sample File** checkbox to select ALL samples.

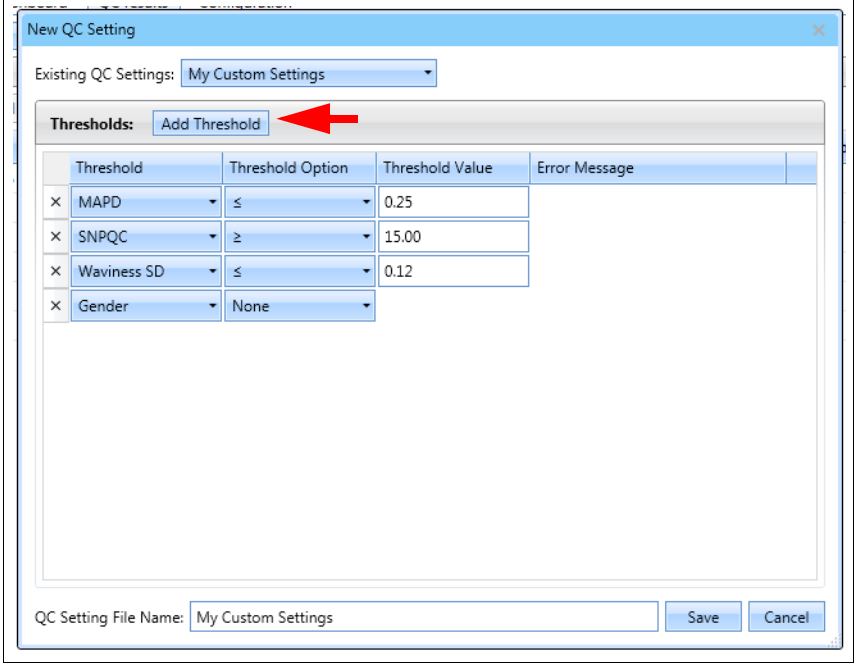
To create your own custom QC setting:

NOTE: Custom QC metrics can be viewed in both the Analysis Workflow and the Browser. However, any custom settings you wish to use, have to be entered separately in both the Analysis Workflow and the Browser. To create custom QC settings in the Analysis Workflow see below. To create custom QC settings in the Browser, see *Setting QC Parameters in the ChAS Browser* (on page 79).

1. Click **Edit or Create QC Settings**.

The **New QC Setting** window appears. (Figure 4.12)

Figure 4.12 New QC Setting window

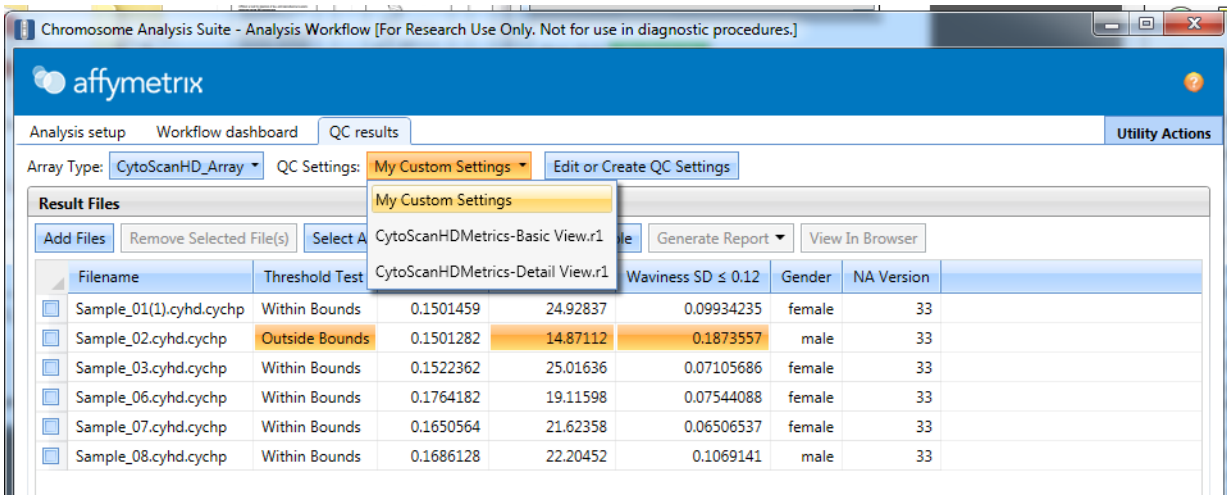


2. Click **Add Threshold** (Figure 4.12) to create a new row.

3. Select the threshold you want to view in your custom QC Setting. Note: The threshold metric you select is not required to have a threshold value.

4. Name your custom QC settings. (Example: *My Custom Settings*) (Figure 4.12), then click **Save**.
   Your custom QC Setting is now available from the QC Settings drop down menu. (Figure 4.13)

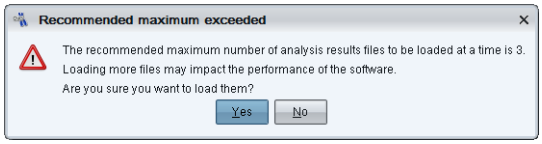
Figure 4.13 New QC Settings menu item



To view results in the Browser:

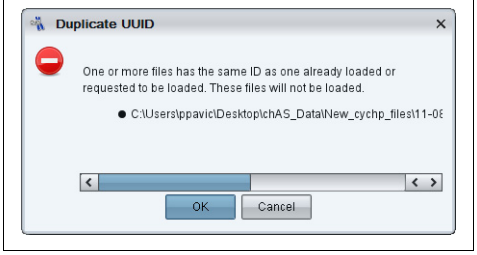
1. At the QC Results window, select the check box(es) for the sample(s) you want to view in the browser, then click **View In Browser**.
   If the following warning message appears (Figure 4.14), click **Yes** to acknowledge it.

Figure 4.14 Recommended maximum exceeded message



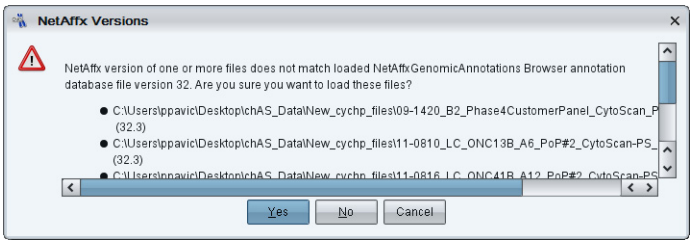
If the following warning message appears (Figure 4.15), click **OK** to acknowledge it.

Figure 4.15 Duplicated files message



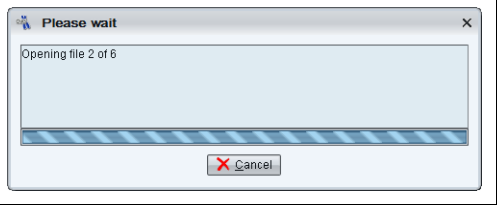
If the following warning message appears (Figure 4.16), click **Yes** to acknowledge it.

Figure 4.16 NetAffx versions message



A progress bar appears (Figure 4.17),.

Figure 4.17 Progress bar

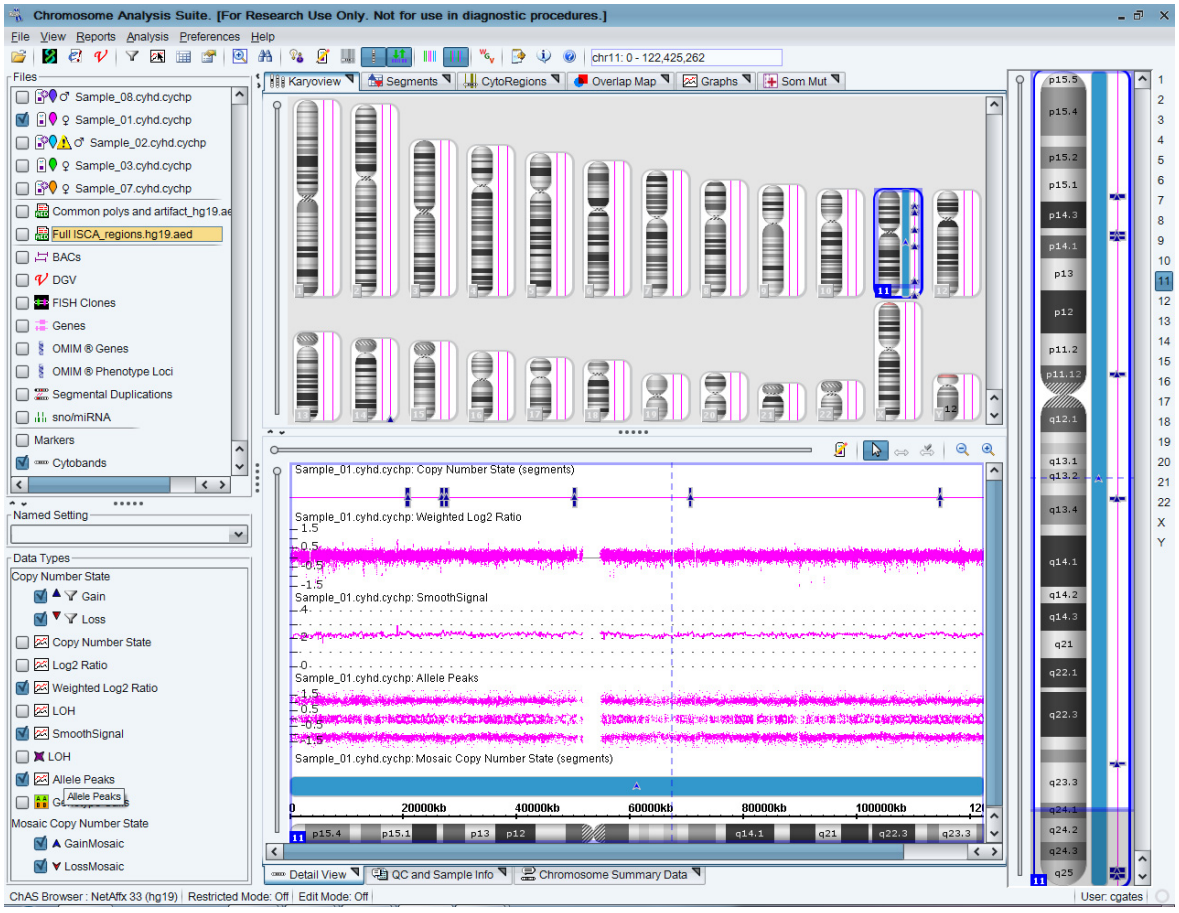


NOTE: The ChAS Browser allows loading of xxCHP files (analyzed by different NetAffx versions) at the same time, however the versions must be from the same genome build.

If NetAffx versions are from different genome builds (for example Hg18 and Hg19), the ChAS Browser will not load the files.

After a few moments, the ChAS browser featuring your selected samples appears. (Figure 4.17)

Figure 4.18 ChAS browser



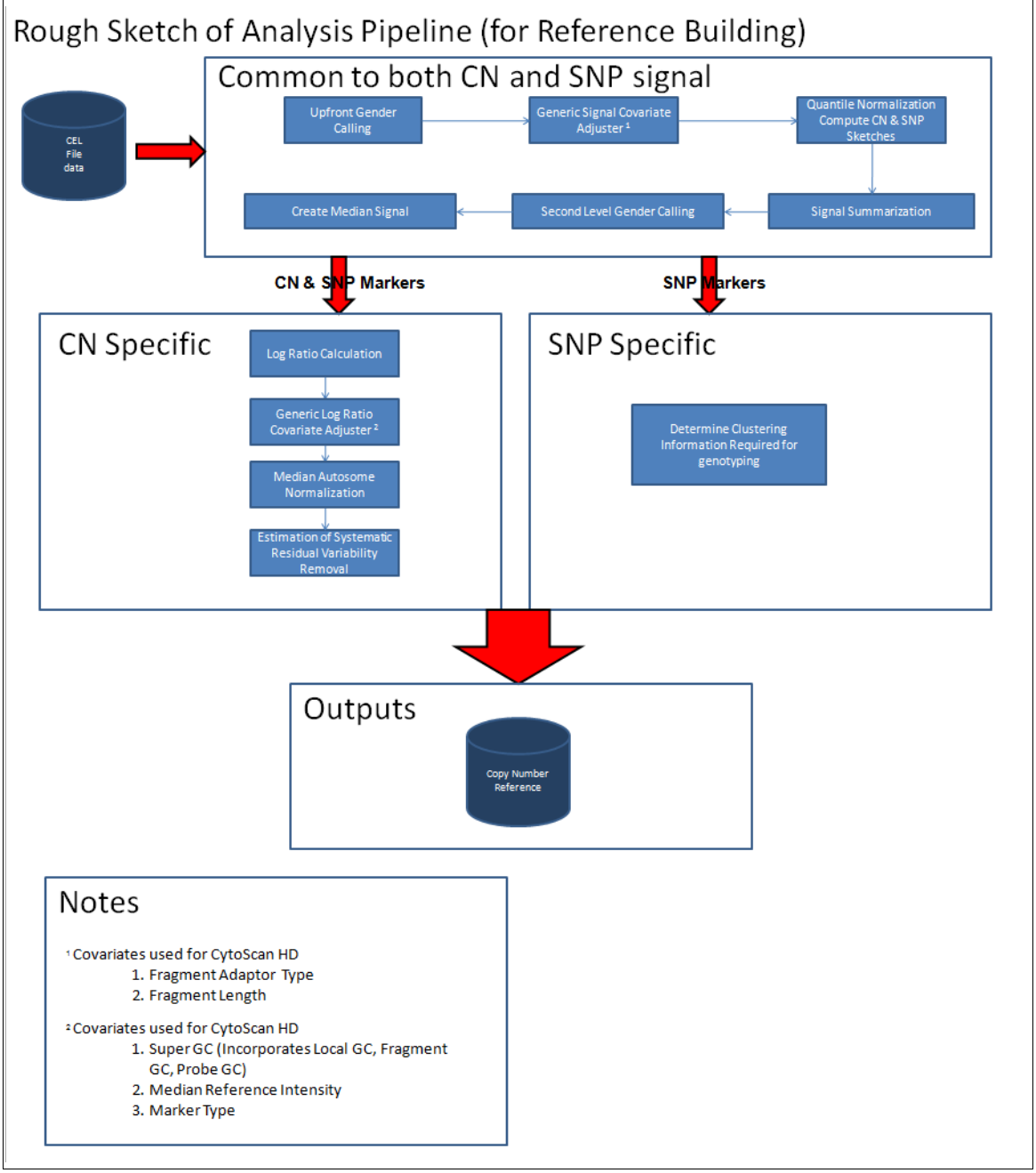
NOTE: You can load several CEL files at a time for a Single Sample Analysis. "Single Sample Analysis" refers to comparing one sample at a time against the selected reference model. CYCHP results are not dependent on whether one or many CEL files are selected for a Single Sample Analysis run.

Reference Files

This section explains how to create a reference file which is required to perform single sample analysis in ChAS. The software analyzes a sample file by comparing it to a reference file. You can use the reference file provided with ChAS, or you can create a reference file using your own sample data.

The figure below (Figure 4.19) shows an overview of the analyses involved in creating a reference file for the CytoScan® HD Array.

Figure 4.19 Overview of creating a reference file

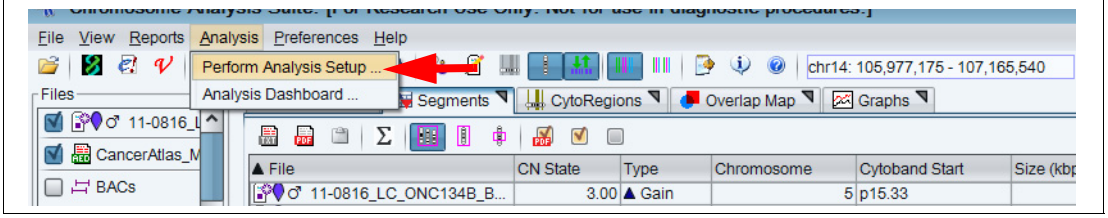


Creating Custom QC Views in the Analysis Workflow

IMPORTANT: When creating a CytoScan reference file, it is recommended that you use a minimum of 44 CEL files. These CEL files must include at least 20 male and 20 female samples.

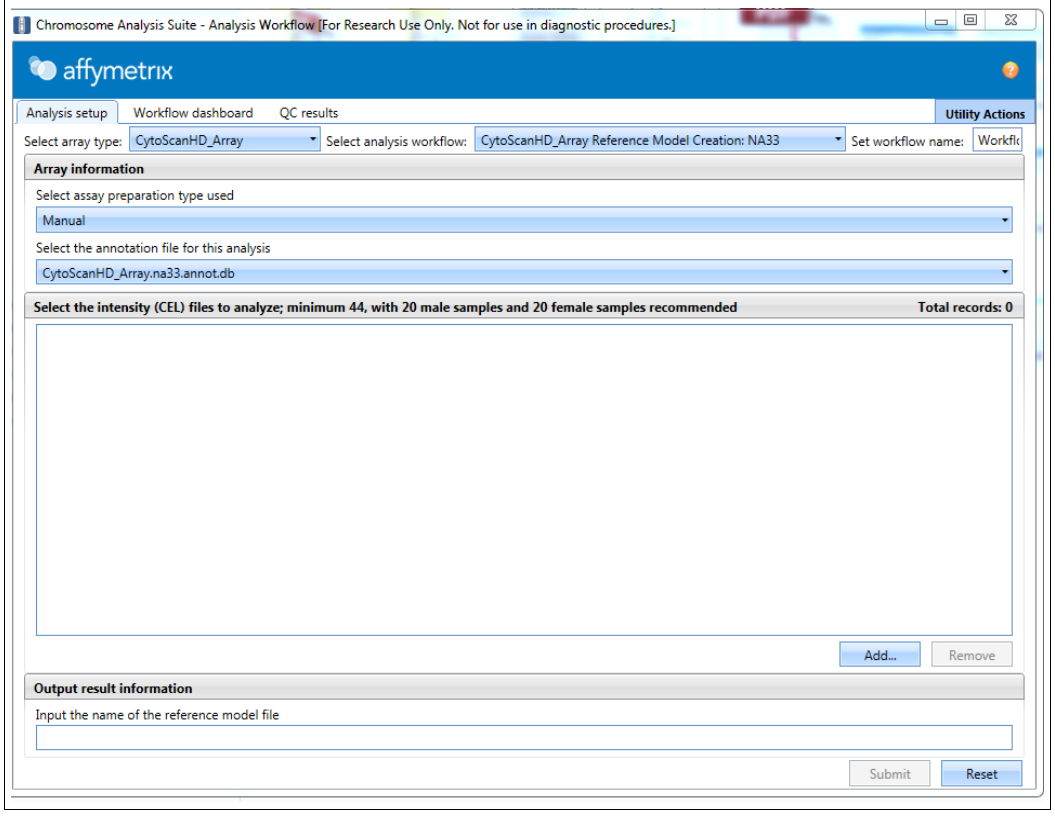
1. From the Analysis menu, select **Perform Analysis Setup.** (Figure 4.20)

Figure 4.20 Analysis drop-down menu



The Analysis Workflow window tab opens (Figure 4.21).

Figure 4.21 Analysis Workflow



2. From the *Select array type* drop-down list, click to select an array type (Example: CytoScanHD_Array)

NOTE: The *Select array type* drop-down list includes only the array types for which library (analysis) files have been downloaded from NetAffx or copied from the Library package provided with the installation.

3. From the *Select analysis workflow* drop-down list, click to select an analysis workflow. (Example: CytoScanHD_Array Single Sample Analysis: NA32.3 *or higher*)

4. By default, the *Set workflow name* is "Workflow". Click Workflow (upper right) to enter a different workflow name.
5. Click the *Assay Preparation type used* drop-down button, then select one of the following:
   - **Manual** (CEL files whose assays were performed by hand)
   - **Automation** (CEL files whose assays were assisted by a robot)
   - **Automation, Manual** (a mix of CEL files whose assays were performed either by hand or with the assistance of a robot

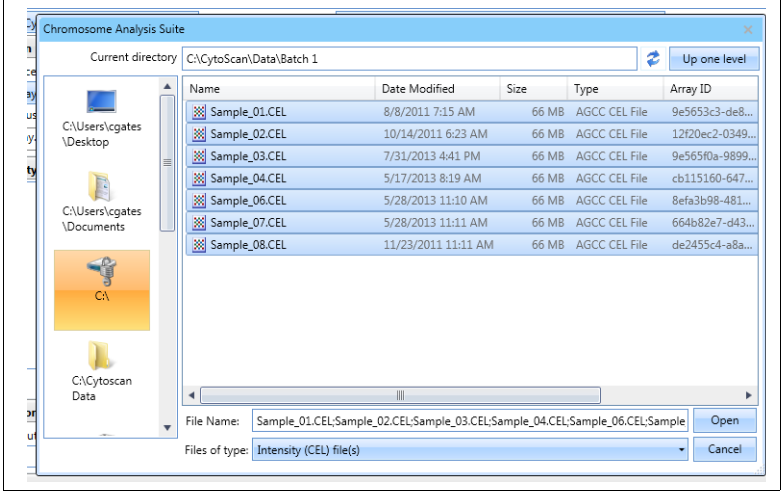
NOTE: The Annotation file for analysis is auto-selected once the array type, analysis workflow and assay preparation type used fields are selected.

6. From the *Select the annotation file for this analysis* drop-down list, verify the selection of the annotation file. (Example: CytoScanHD_Array.na.33.annot.db)
7. At the *Select the intensity (CEL) file(s) to analyze* pane, click Add....

IMPORTANT: The same annotation file you used to create a Reference Model File MUST also be used with future Single Sample Analyses runs that utilize your created Reference Model File.

The following window appears: (Figure 4.22)

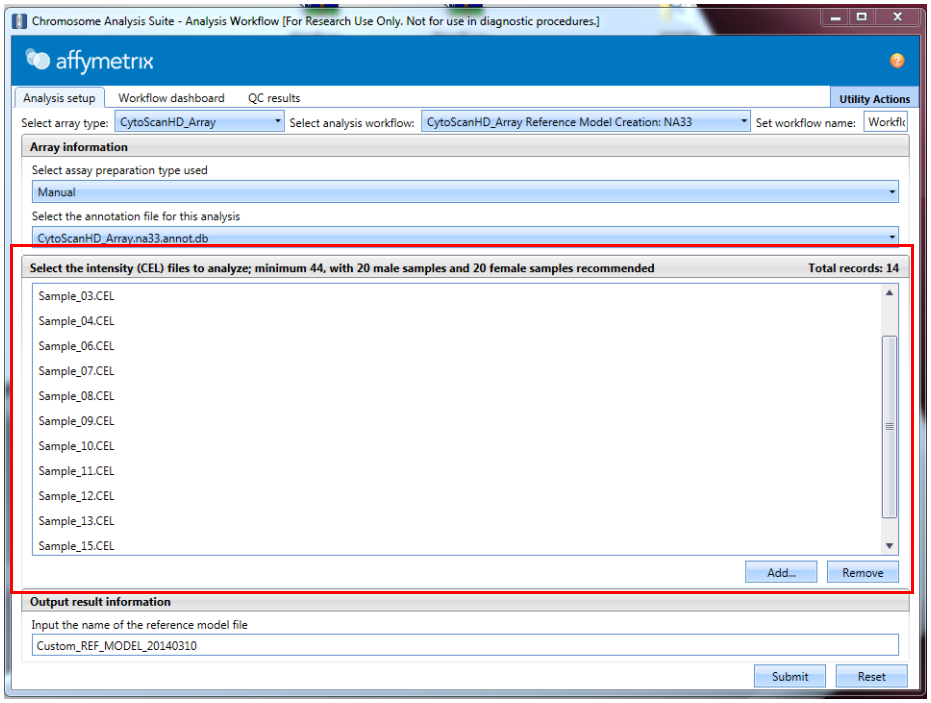
Figure 4.22 >CEL file folder



8. Single click, Ctrl click, or Shift click (to select multiple .CEL files), then click Open.

The *Select the intensity (CEL) file(s) to analyze* pane is now populated with your .CEL files. (Figure 4.23)

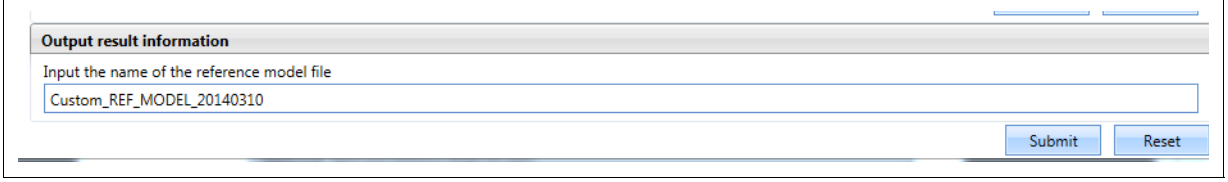
Figure 4.23 Select the intensity (CEL) file(s) to analyze pane



To remove a .CEL file from this list, click to highlight it, then click Remove.

9. At the *Output result information* pane, enter a name for your reference model file. (Figure 4.24)

Figure 4.24 Name your Reference Model File name

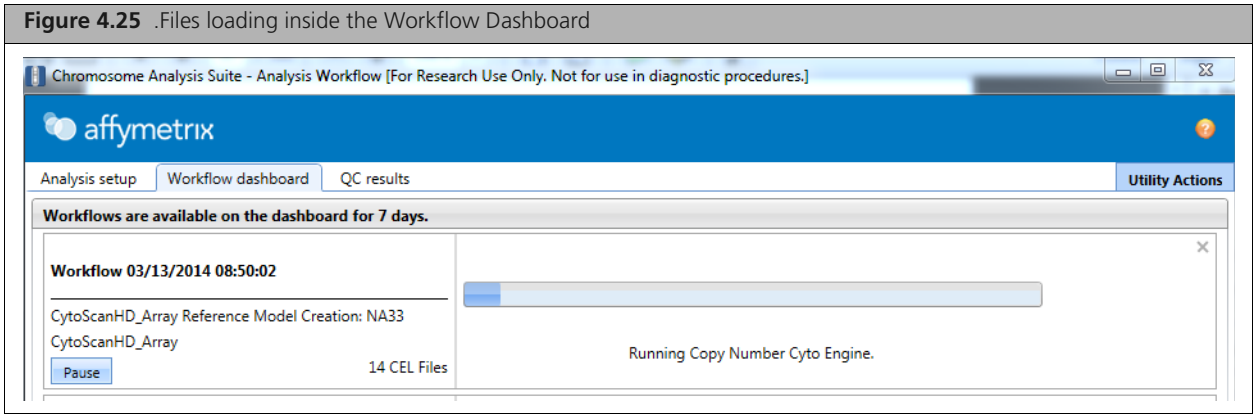


10. Click Submit.

The Workflow dashboard window appears and your annotation files begin to load. (Figure 4.25).

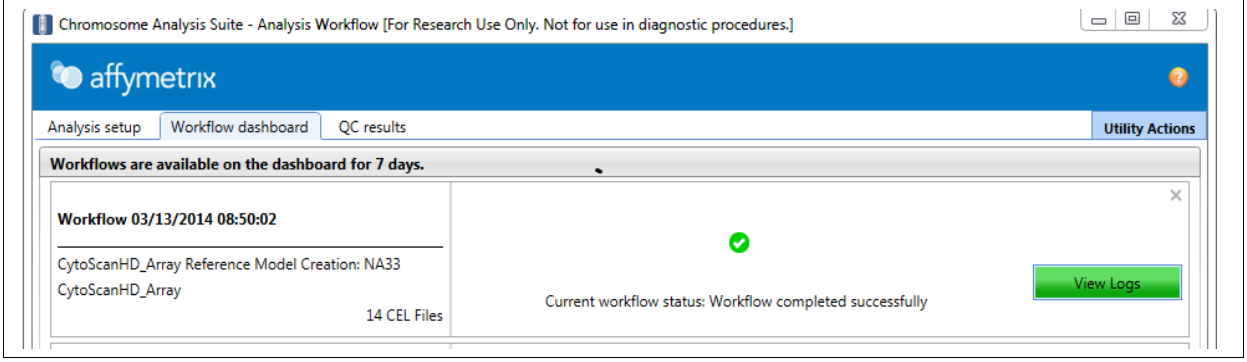
The Analysis Workflow Dashboard tracks ongoing analysis tasks for ChAS. It also delivers the results of analyses and restarts the Browser if it was shut down to free up memory for the analysis.

Figure 4.25 .Files loading inside the Workflow Dashboard



After loading is complete, a Workflow completed successfully message appears. (Figure 4.26)

Figure 4.26 Workflow Dashboard loaded



The successfully created Reference Model File has been placed in your existing ChAS Library folder and is ready to be used in your next Single Sample Analysis.

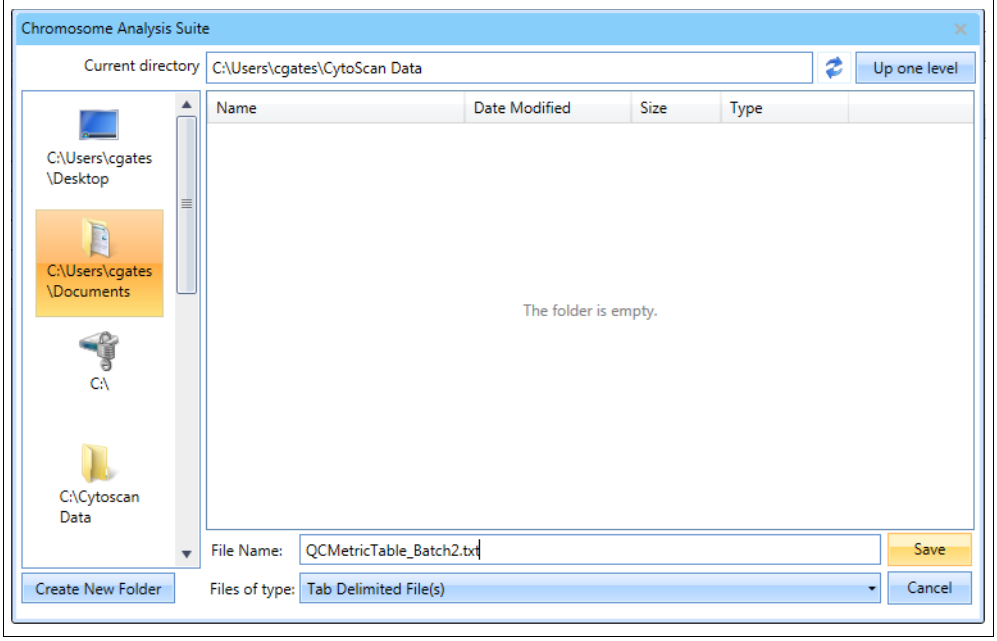
Exporting QC Tables and Probe-Level Data Using Generate Report

To Export QC Table Information:

1. Click on the **Export QC Table** button.

A window appears that allows saving of QC data and where you want to export it. (Figure 4.27)

Figure 4.27 Workflow Dashboard loaded



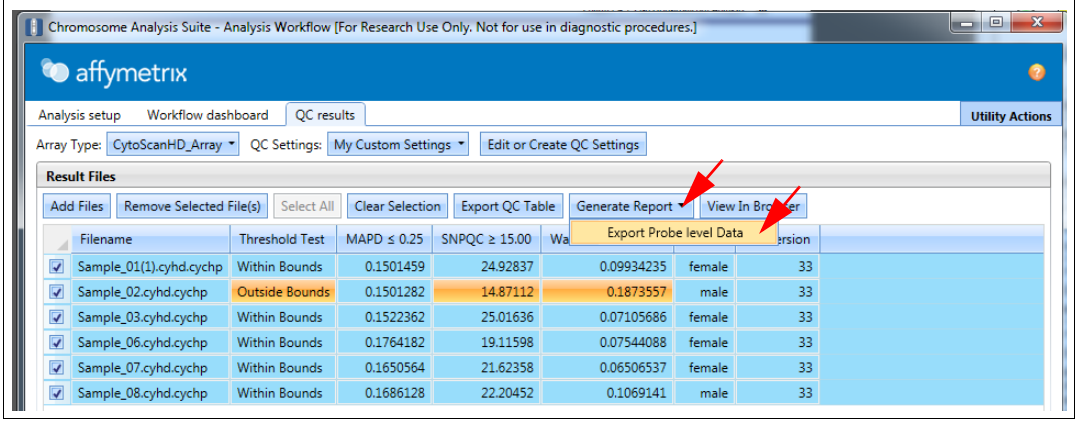
2. Click **Save**.

To export Probe-Level Data using Generate Report:

1. Click the checkbox next to the Results File(s) you want to generate a report for or click **Select All**.
2. To export probe-level data, click the **Generate Report** drop-down. (Figure 4.28)

   The following export reporting option appears:

Figure 4.28 Generate Report drop-down menu



3. Click to select the Export Probe Level Data option. (Figure 4.28)

Your previously assigned Output folder file window appears. At this point, you may navigate to another location to save the file.

NOTE: The default root filename is Result. Click inside the File Name field to enter a different root filename.

4. Enter a File Name for your text (tab-delimited) reporting file, then click Save.

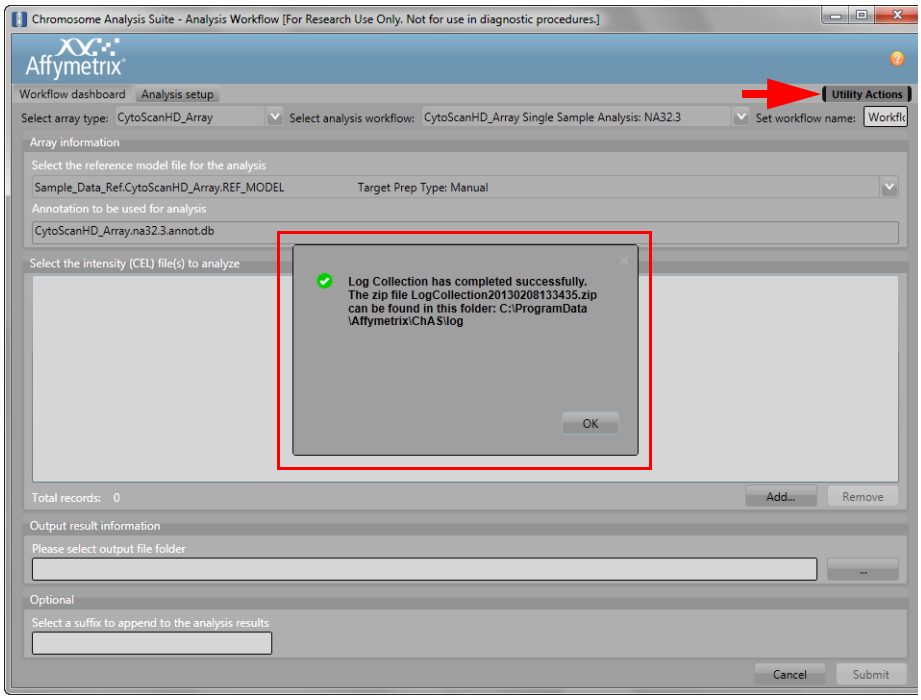
Analysis Workflow Troubleshooting

Do the following if you experienced any issues or failures with your analysis:

1. Click the **Utilities** button (top right of the Analysis Workflow window)
2. Click to select **Log Collection**.

   The following dialog window appears. (Figure 4.29)

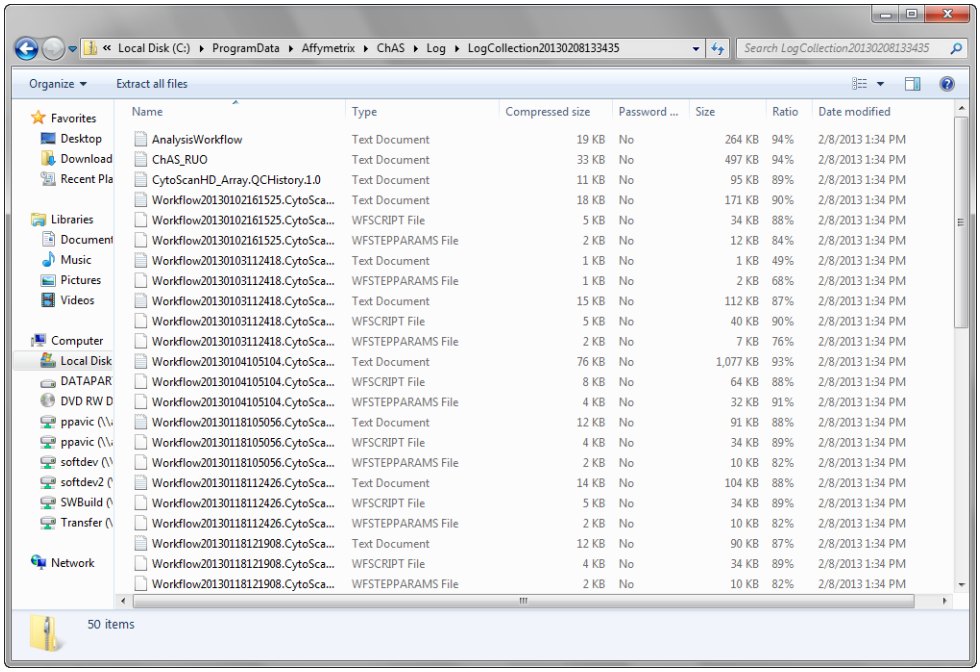
Figure 4.29 Log Collection dialog window



3. Make note of the assigned zip filename and its location.
4. Use Windows Explorer to navigate to the location. (Example: `C:\ProgramData\Affymetrix\ChAS\log`)
5. Locate the zip file you noted earlier, then double-click on it, or right-click to open it..

   (See example Figure 4.30 on page 65)

Figure 4.30 Example: Zip file contents of a Log Collection



Log Rollover

When the software determines that the log file for the Analysis Workflow (`C:\ProgramData\Affymetrix\ChAS\log\AnalysisWorkflow.log`) has reached a defined size (approximately 4MB), the following steps will be completed:

- A sub-folder will be created in `C:\ProgramData\Affymetrix\ChAS\log` called '**Log***' (the '*' denotes the current date and time).
- A zip file called **RolledLogFile*.zip** is created in that folder. The '*' is the same date and time used for the folder name. The files in the `C:\ProgramData\Affymetrix\ChAS\log` folder and all files found in the currently selected QC History Log folder will be included in this zip file.
- The Analysis Workflow files that are associated with analysis workflows that are no longer active on the Dashboard will be deleted from `C:\ProgramData\Affymetrix\ChAS\log`.
- A new AnalysisWorkflow.log file will be created in `C:\ProgramData\Affymetrix\ChAS\log`.

Log Collection

When the Log Collection option is selected from the Utility Actions menu, a file called **LogCollection*.zip** (the '*' denotes the current date and time) is created in the folder `C:\ProgramData\Affymetrix\ChAS\log`.

This created file contains the full contents of the folder `C:\ProgramData\Affymetrix\ChAS\log`, including the log file for the browser (ChAS_RUO.log).

If available, the sub-folders of `C:\ProgramData\Affymetrix\ChAS\log` and all files found in the currently selected QC History Log folder.

Loading Data

ChAS can display data from:

- *CytoScan® HD Array CYCHP files, generated in ChAS.*
- *Genome-Wide Human SNP Array 6.0 CNCHP files, generated in Genotyping Console (GTC).*
- *OncoScan FFPE Assay OSCHP files, generated in OncoScan Console.*

NOTE: **When referring to steps that apply to both CytoScan CYCHP and SNP 6.0 CNCHP data files, the CHP files are described as CxCHP files.**

When referring to steps that apply to CytoScan CYCHP, SNP 6.0 CNCHP data, and OncoScan FFPE Assay files, the resultant files are described as xxCHP files.

The following sections include:

- *Introduction to Loading Data (on page 66)*
- *Loading Files (on page 68)*
- *Segment Smoothing and Segment Joining (Optional) (on page 73)*
- *Setting QC Parameters in the ChAS Browser (on page 79)*

Introduction to Loading Data

The same steps are used to load results xxCHP files from a CytoScan HD, Genome-Wide Human SNP Array 6.0, or OncoScan FFPE Assay Array.

When loading CYCHP files into ChAS for viewing, the software:

1. Loads the run-length encoded segments in the CYCHP file to display as segments.
2. Applies any smoothing or joining that would alter the length and other properties of segments.

IMPORTANT: **In a new user profile, smoothing and joining are turned on by default. The smoothing and joining settings are specific for the array type CxCHP file (for more details on smoothing and joining, see *Segment Smoothing and Segment Joining (Optional)* (on page 73)). Smoothing and Joining is disabled for OSCHP files.**

3. Displays the segments and graph data:
   - Segment Data
     - Copy Number Gain/Loss
     - Mosaicism Copy Number Gain and Loss
     - Loss of Heterozygosity (LOH)
   - Graph Data
     - Copy Number State
     - Log2 Ratio
     - Weighted Log2 Ratio
     - LOH
     - Smooth Signal

- Allele Peaks or Allele Difference
- Genotype Calls

When loading SNP6 CNCHP files into ChAS for viewing, the software:

1. Performs segment detection by analyzing the CN and LOH graph data in the CNCHP file.

   NOTE: When running the Segment Reporting Tool in GTC on SNP 6 data, the software sets the end coordinate such that the segment ends at the base position of the last marker in the segment. When loading SNP 6 data into ChAS, the segment detection sets the end coordinate for a segment such that the segment ends one base after the last marker in the segment. This may result in a discrepancy between the end position for segments when comparing data analyzed in both GTC and ChAS.

   NOTE: ChAS uses the median of the aberrant markers' CNStates as the recalculated CNState of the smoothed (and/or joined) segment.

2. Applies any smoothing or joining that would alter the length and other properties of segments.

   Smoothing is similar to the process applied when running the Segment Reporting Tool in GTC.

   IMPORTANT: For CNCHP files from the SNP 6.0 Array, smoothing but not joining is turned on by default in a new user profile.

3. Displays the segments and graph data:
   - Segment data
     - Copy Number Gain/Loss
     - Loss of Heterozygosity (LOH)

       NOTE: The expected Copy Number State on the X chromosome of normal males is not constant over its entire length. This is due to the structure of the sex chromosomes. for more information see *Copy Number Segments on the X and Y Chromosomes* (on page 47).

   - Graph Data
     - Copy Number State
     - Log2 Ratio
     - Allele Difference
     - SmoothSignal
     - LOH

When loading OncoScan FFPE Assay OSCHP files for viewing, the software:

1. Displays segments in the OSCHP created by the TuScan Copy number algorithm. For details on this algorithm, please refer to the *OncoScan Console User Guide* (P/N 703195).

   IMPORTANT: Smoothing and Joining is disabled for OSCHP files.

2. Displays the segments and graph data:
   - Segment Data
     - Copy Number Gain/Loss
     - Loss of Heterozygosity

- Graph Data
  - Copy Number State
  - Log2 Ratio
  - Weighted Log2 Ratio
  - Allele Difference
  - B-allele Frequency
  - LOH
  - Smooth Signal
  - Somatic Mutation

Loading Files

Loading xxCHP data for viewing in ChAS involves the following steps:

1. **Optional:** Before loading: Select Segment Smoothing and Segment Joining parameters for processing the CN Gain and Loss Segment data for CxCHP files.

   IMPORTANT: Smoothing is turned on by default in a new user profile. Joining is also on by default for the CytoScan® HD Array, but not for the Genome-Wide Human SNP Array 6.0.

   Smoothing and Joining is disabled for OSCHP files.

2. Select xxCHP files from:
   - CytoScan® HD Arrays (generated by analyzing CEL files in ChAS)
   - Genome-Wide Human SNP 6.0 Arrays (generated by analyzing CEL files in GTC).
   - OncoScan FFPE Assay arrays (generated by analyzing in OncoScan Console).

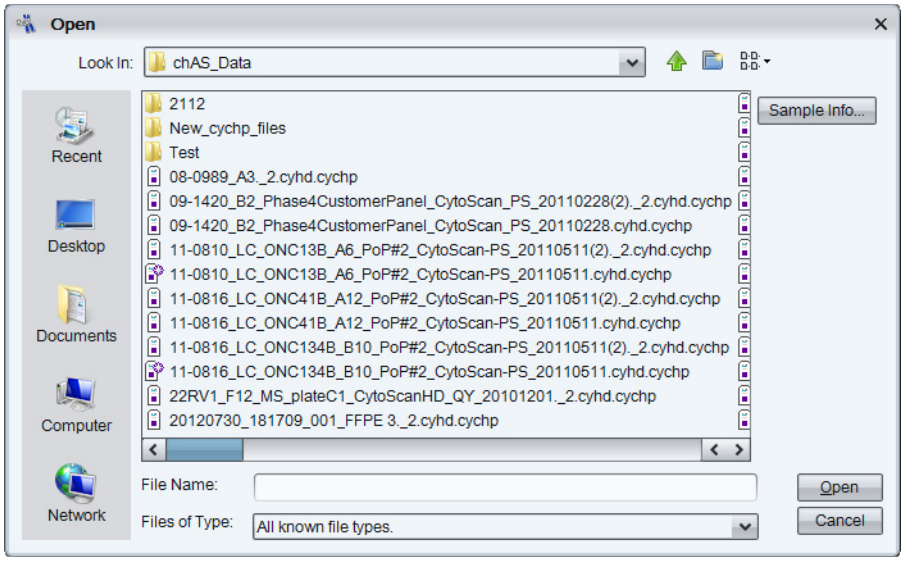
   You can also select region information files in AED and BED format for loading. Use the Open window (click the button) to load xxCHP data files, Affymetrix Extensible Data (AED), or Browser Extensible Data (BED) annotation files. The AED and BED files that you open will be automatically loaded when a new session is started with the same user profile.

   NOTE: You may want to edit smoothing and joining parameters. This can be done before or after loading the CxCHP data. See *Segment Smoothing and Segment Joining (Optional)* (on page 73) for more information.

3. Select **File** → **Open** on the menu bar. Alternatively, click the **File Open** button.

   The Open window appears (Figure 5.1).

Figure 5.1 Open window

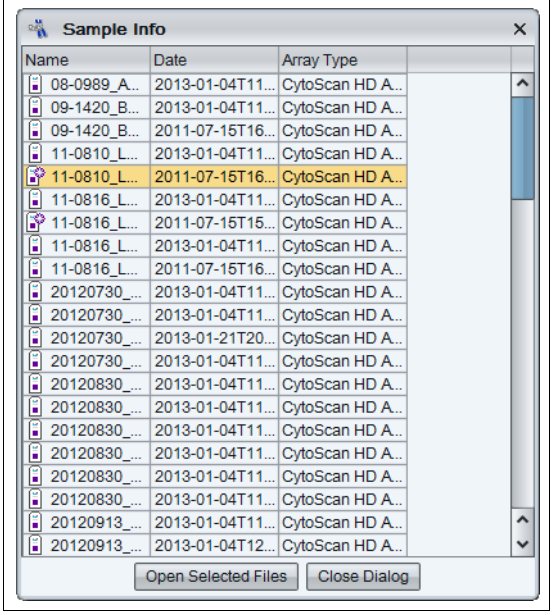


4. To view information about results, select one or more files, then click **Sample Info**.

   The Sample Info window opens (Figure 5.2 on page 69).

> **NOTE: If the xxCHP and ARR files are located in the same folder, the Sample Info window shows information about both the sample and the results.**

Figure 5.2 Sample Info dialog box



5. To load files from the Sample Info dialog box, select the files, then click **Open Selected Files**.

Loading Files using the Search Feature

1. Click **Close Dialog** (Figure 5.2) to close Close the Sample Info dialog box.
   The Open window appears.
2. Navigate to the folder with the files that you want to search for and load.
3. Enter a text string with an asterisk (*) ***before*** and ***after*** the search term in the ***File Name*** field. See example in Figure 5.3 on page 70.
4. Select a file type from the drop-down list. See example in Figure 5.3 on page 70.
5. Click **Open**.
   Files with names that include the search term are displayed in the Open dialog box. See example in Figure 5.4 on page 71.

Figure 5.3 Enter a file name search term (*text string*) and select a file type

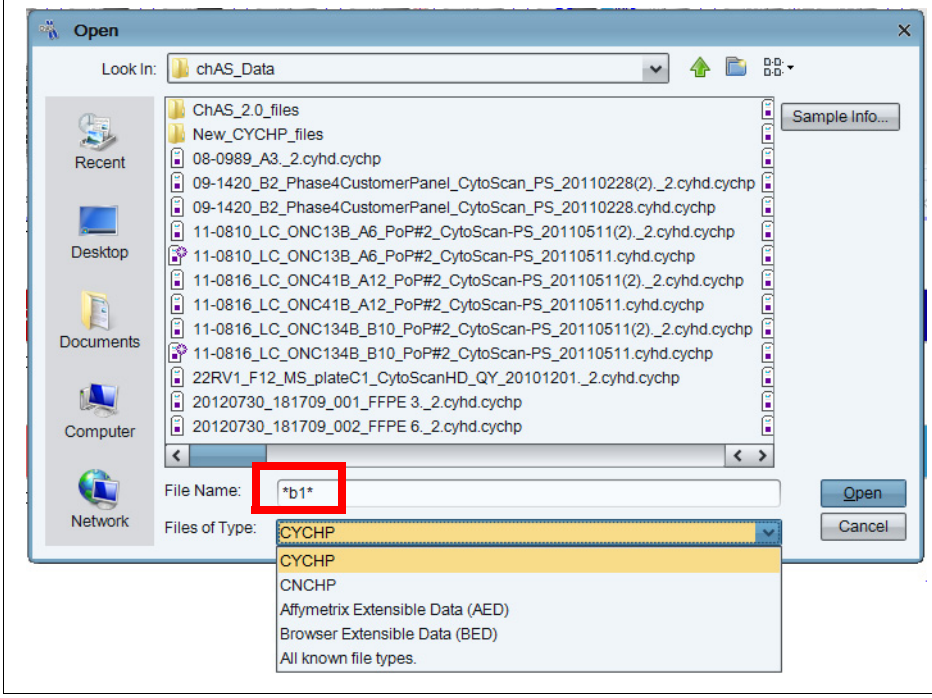
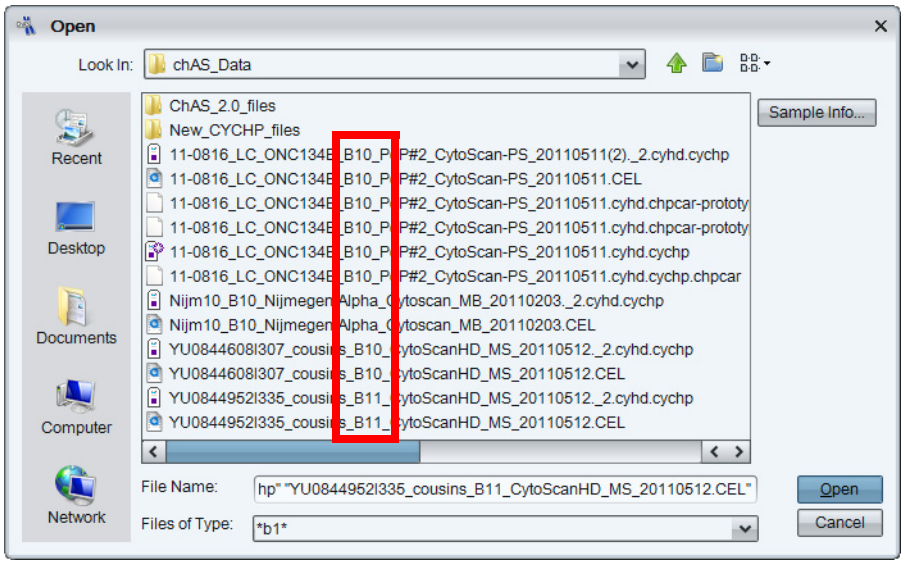


Figure 5.4 Open dialog box shows files with names that include the search term

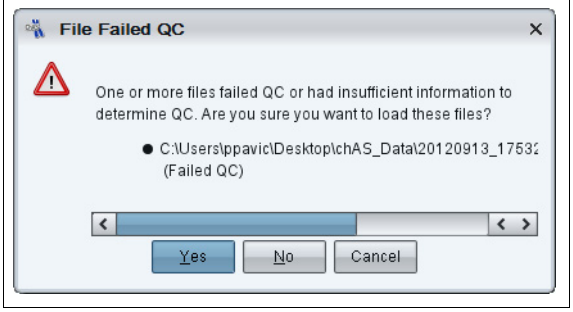


6. Select the files (you can use Shift click or CTRL click to select multiple files)
7. Click **Open**.

   If any of the files fail the QC checks, a warning notice appears (Figure 5.5).

   You can click **Yes** to continue to load the files.

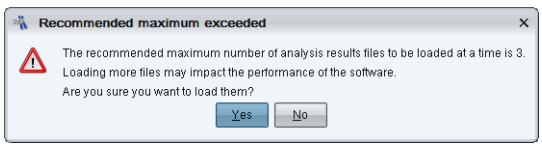
Figure 5.5 Warning Notice for QC failure



It is recommended you load no more than 3 files at one time.

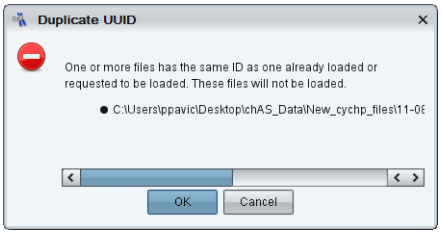
If the following warning message appears (Figure 5.6), click **OK** to acknowledge it or click **Cancel** to reselect a maximum of 3 files.

Figure 5.6 Recommended maximum exceeded message



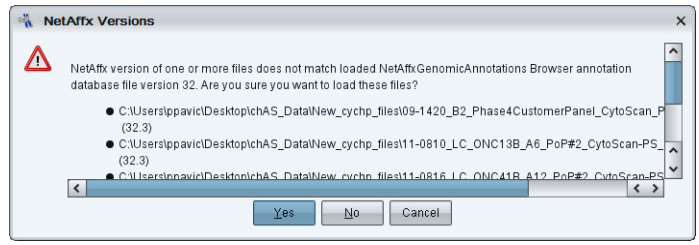
If the following warning message appears (Figure 5.7), click **OK** to acknowledge it.

Figure 5.7 Duplicated files message



If the following warning message appears (Figure 5.8), click **Yes** to acknowledge it.

Figure 5.8 NetAffx versions message

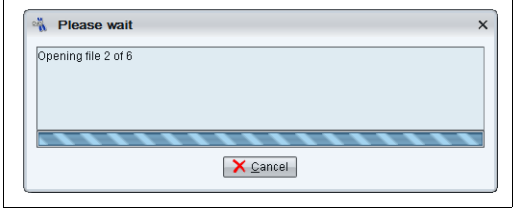


NOTE: The ChAS Browser allows loading of xxCHP files (analyzed by different NetAffx versions) at the same time, however the versions must be from the same genome build.

If NetAffx versions are from different genome builds (for example Hg18 and Hg19), the ChAS Browser will not load the files.

A progress bar appears (Figure 5.9)

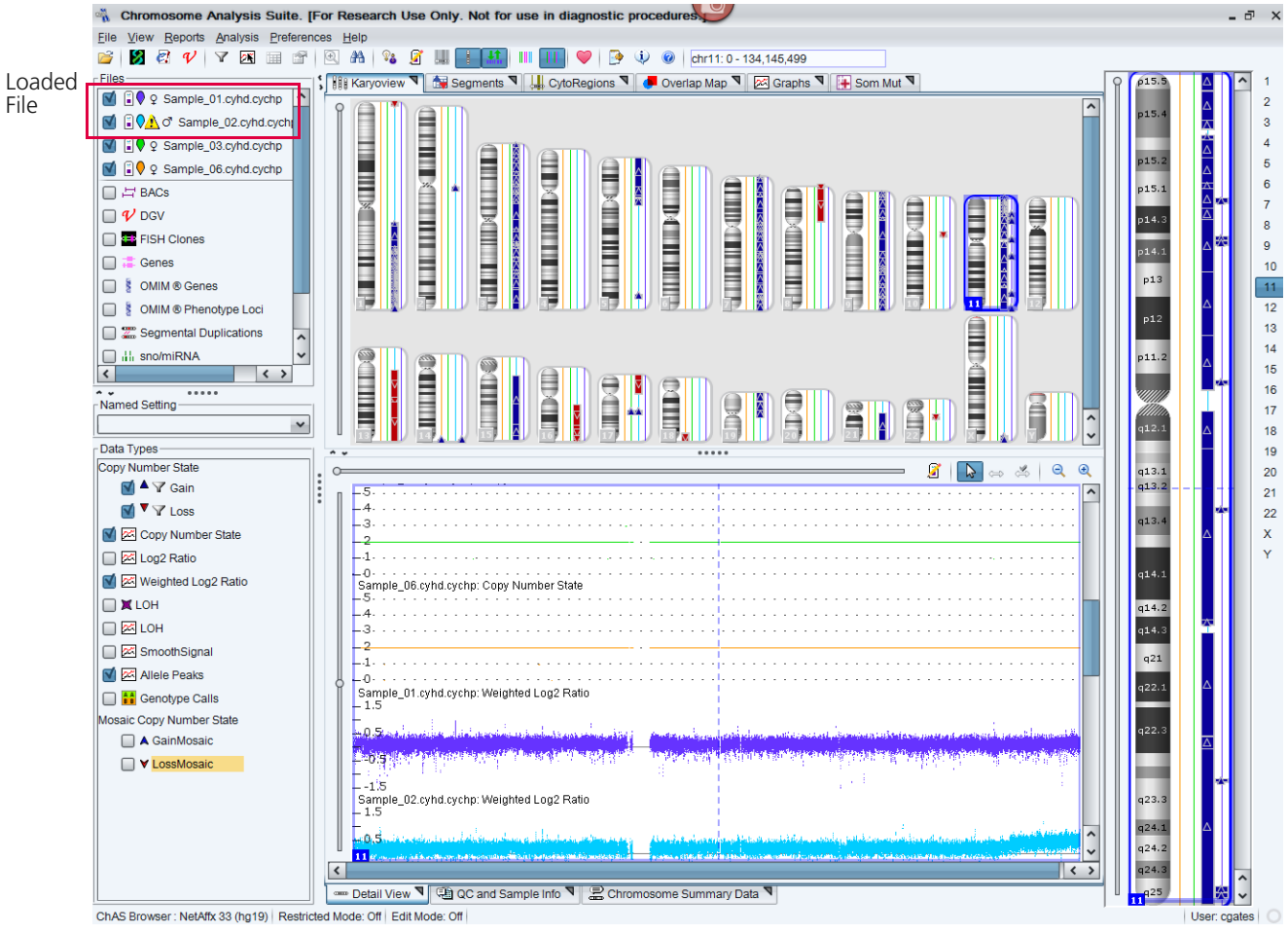
Figure 5.9 Progress bar



After a few moments, the ChAS browser featuring your selected samples appears.

The loaded files appear in the Files list pane. (Figure 5.10).

Figure 5.10 Files list shows the loaded data



Segment Smoothing and Segment Joining (Optional)

"Smoothing" and "joining" are non-destructive processes that affect the display of Copy Number segments. Smoothing and joining are performed on the Copy Number State data during the loading process, based on settings that are specified before loading. Any data filtering is applied after smoothing and joining.

IMPORTANT: Smoothing and joining are turned on by default in a new user profile.

Smoothing and joining are specified per array type. The processes do not affect the marker data in the CNCHP or CYCHP file. If these settings are turned off, the Copy Number segment data is displayed without smoothing or joining.

IMPORTANT: Smoothing and joining affect only data loaded from CNCHP and CYCHP files. This ONLY applies to copy number data, NOT LOH or Mosaic types. Smoothing and Joining is disabled for OSCHP files.

Segments which have been smoothed and/or joined are indicated by a blue check mark in the Smoothed/Joined column of the Segments table (Figure 5.11). The segment ID name indicates whether smoothing and/or joining has occurred. A red "X" indicates no smoothing or joining has been applied.

To view what segments were smoothed/joined:

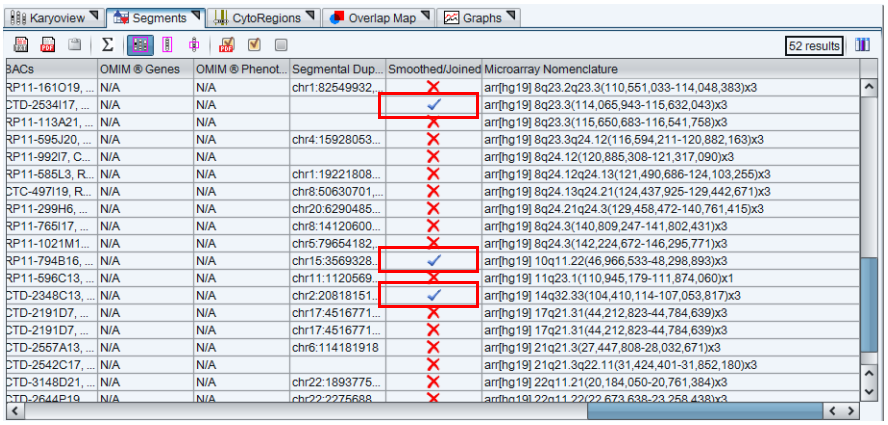
1. Click the Segments tab.
2. Use the horizontal scroll bar to move the window to the far right.

   The Smoothed/Joined column appears. (Figure 5.11)

For more information:

- *See "About Smoothing" on page 77.*
- *See "About Joining" on page 78.*

Figure 5.11 Example Segments table with smooth/joined segments

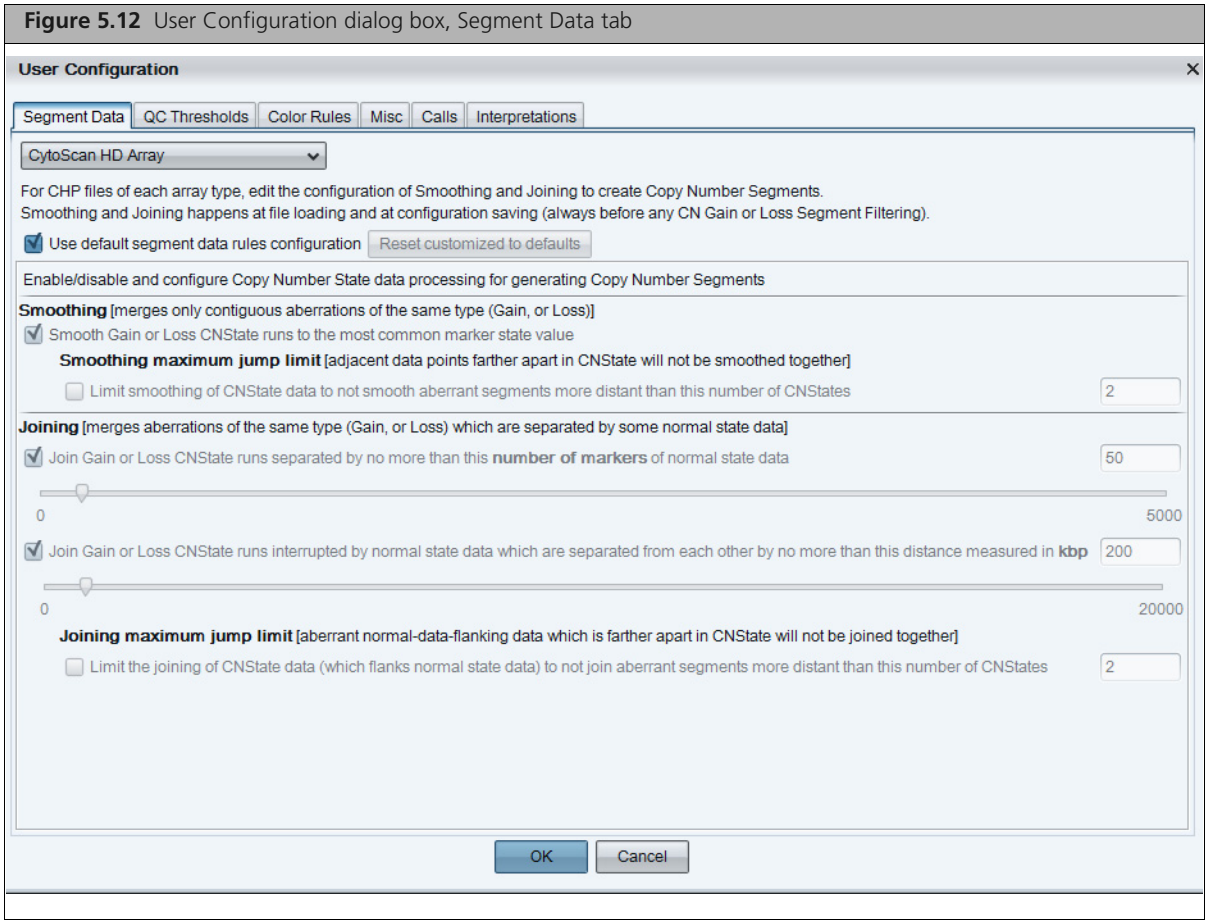


To set Smoothing and joining parameters:

1. Click **Preferences → Edit User Configuration**.

   The User Configuration window appears.
2. Click the **Segment Data** tab.
3. Click the **Choose Array Type** drop-down menu to select the array type to view or edit its Smoothing and Joining settings. (Figure 5.12 on page 75)

Figure 5.12 User Configuration dialog box, Segment Data tab



Segment Tab Options

Use default segment data rules configuration	For the CytoScan® HD Array, the default smoothing and joining rules are: Smooth Gain or Loss CNState runs to the most common marker value, then generate segments. Join any "split" CNState runs separated by no more than 50 normal-state markers. Join Gain or Loss CNState runs interrupted by normal state data which are separated from each other by no more than 200 kbp For SNP 6 arrays, the default smoothing rule: Smooth Gain or Loss CNState runs to the most common marker value, then generate segments.
Smooth Gain or Loss CNState runs to the most common marker value	Smoothing to the most common marker state value is only applied to contiguous CNState runs of the same type (gain or loss).
Limit smoothing of CNState data to not smooth aberrant segments more distant than this number of CNStates	If this option is chosen, CNState runs which are farther apart than the "smoothing maximum jump limit" will not be smoothed. For example, if the smoothing maximum jump limit is set at 1, then adjacent segments with CNState 3 and 5 will not be smoothed.
Join Gain or Loss CNState runs separated by no more than this number of markers of normal state data	If this option is chosen, only Gain or Loss CNState Runs which are separated by less than a threshold number of markers of normal state data will be joined. For example, if the marker threshold is set at 50, then CNState runs separated by more than 50 markers of normal state data will not be joined.
Join Gain or Loss CNState runs interrupted by normal state data which are separated from each other by no more than this distance measured in kbp	Choose this option to limit the distance between two CNState runs separated by normal state data that can be joined. For example, if the distance set at 20000 kbp, CNState runs separated by normal state data spanning more than 20000 kbp from the end of the first aberrant segment to the start of the next aberrant segment will not be joined.
Limit the joining of CNState data (which flanks normal state data) to not join aberrant segments more distant than this number of CNStates	If this option is chosen, CNState runs that flank normal state data and are father apart from each other than a user-specified CN threshold will not be joined. For example, if the CNState joining maximum jump limit is set at 1, then CNState runs of 3 and 5 that flank normal state data will not be joined.

IMPORTANT: If multiple smoothing and/or joining check boxes are selected, all criteria must be met to smooth and/or join the segments.

After smoothing and joining, the marker count, mean marker distance and confidence values get recalculated.

For more information:

- *See "About Smoothing" on page 77.*
- *See "About Joining" on page 78.*

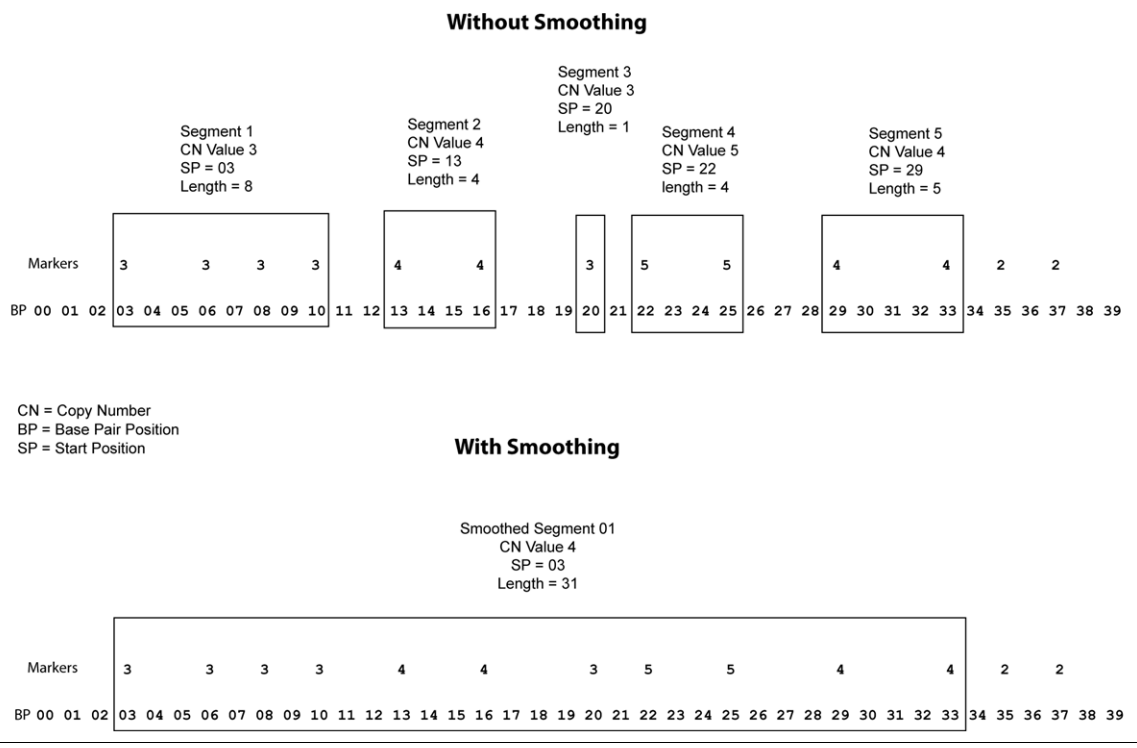
About Smoothing

NOTE: The examples provided are for a case where the expected copy number is 2. Similar calculations take place for the X and Y chromosomes where the expected copy number may be 0, 1 or 2, depending on gender and whether the segment is located within or outside of the PAR region.

If you have a contiguous set of segments with gain values (for instance, of CN State values of three and four), with no markers of copy number 2 or lower, **without smoothing** they will be treated as a series of individual gain segments. The same rules apply to a set of segments with loss values of 0 or 1.

If you have a contiguous set of markers with gain values of three and four, with no intervening markers of copy number 2 or lower, **with smoothing** they will be consolidated into a single gain segment. (Figure 5.13)

Figure 5.13 Simplified schematic representation of smoothing



If you have a contiguous set of markers with loss values of zero and one, with no intervening markers of copy number 2 or higher, after smoothing, they will be consolidated into a single loss segment.

The smoothing process is the same as the process automatically performed by the Segment Reporting Tool in GTC. Different methods are used to assign the CN state value and perform the confidence calculations, as described below. See the *GTC User Manual* for more information.

Copy Number State for Smoothed Segments

The median CNState of all the markers in the Segment is assigned as the Copy Number State value for the new smoothed segment. The median will thus always be either an integer or a half integer (like 3.5).

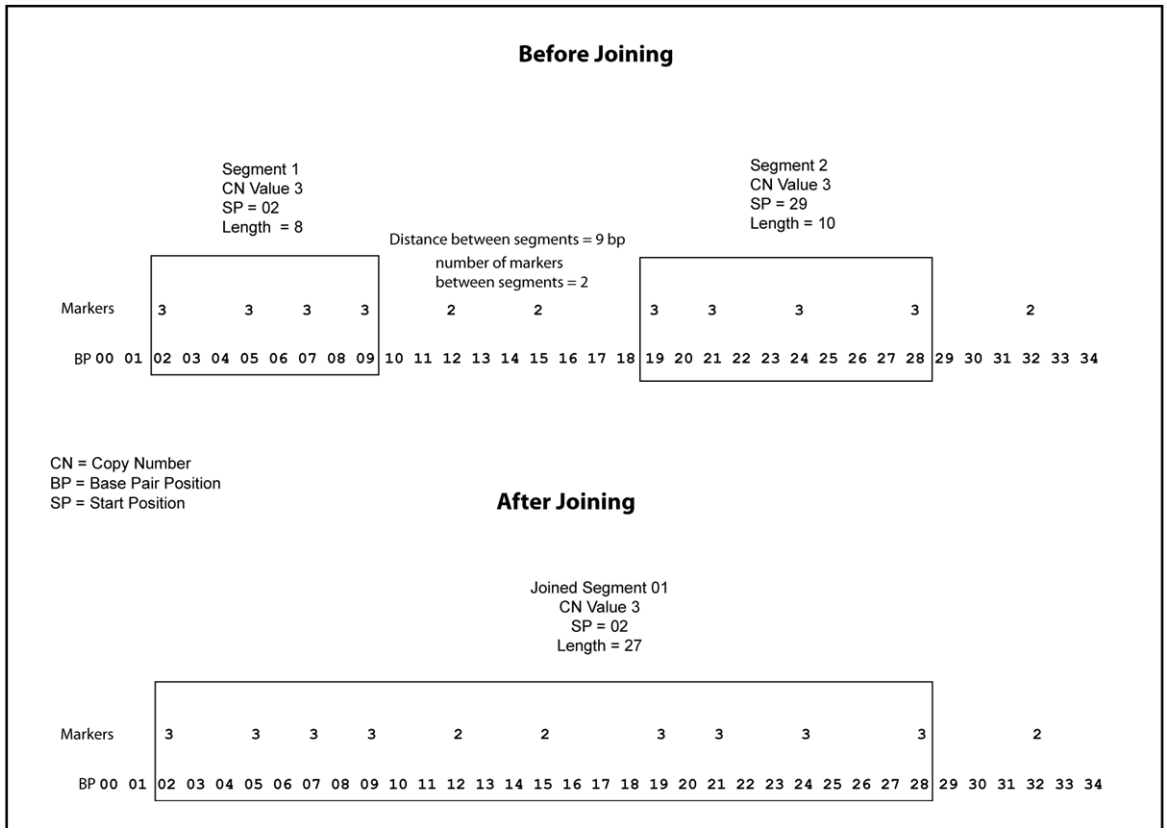
For all the half-integer cases:

- *Gains are rounded up to the next full integer (3.5 goes to 4)*
- *Losses are rounded down to the next full integer (0.5 goes to 0, 1.5 goes to 1).*

About Joining

The joining options allow you to join segments with the same type (gain or loss) aberrant CNState that are separated by no more than a specified number of normal-state markers or by no more than a specified distance of normal-state data (Figure 5.14).

Figure 5.14 Simplified schematic representation of joining



The equivalence of CNState of the segments to be joined could have happened as a result of smoothing, or been from “raw” unsmoothed segments with the same CNState.

Setting QC Parameters in the ChAS Browser

ChAS checks the xxCHP files for certain QC values. The software notifies you if the QC parameters do not meet the thresholds.

NOTE: Custom QC metrics can be viewed in both the Analysis Workflow and the Browser. However, any custom settings you wish to use, have to be entered separately in both the Analysis Workflow and the Browser.

To create custom QC settings in the Browser see below. To create custom QC settings in the Analysis Workflow, see *To create your own custom QC setting:* (on page 54).

You can adjust the QC threshold values or select different QC metrics.

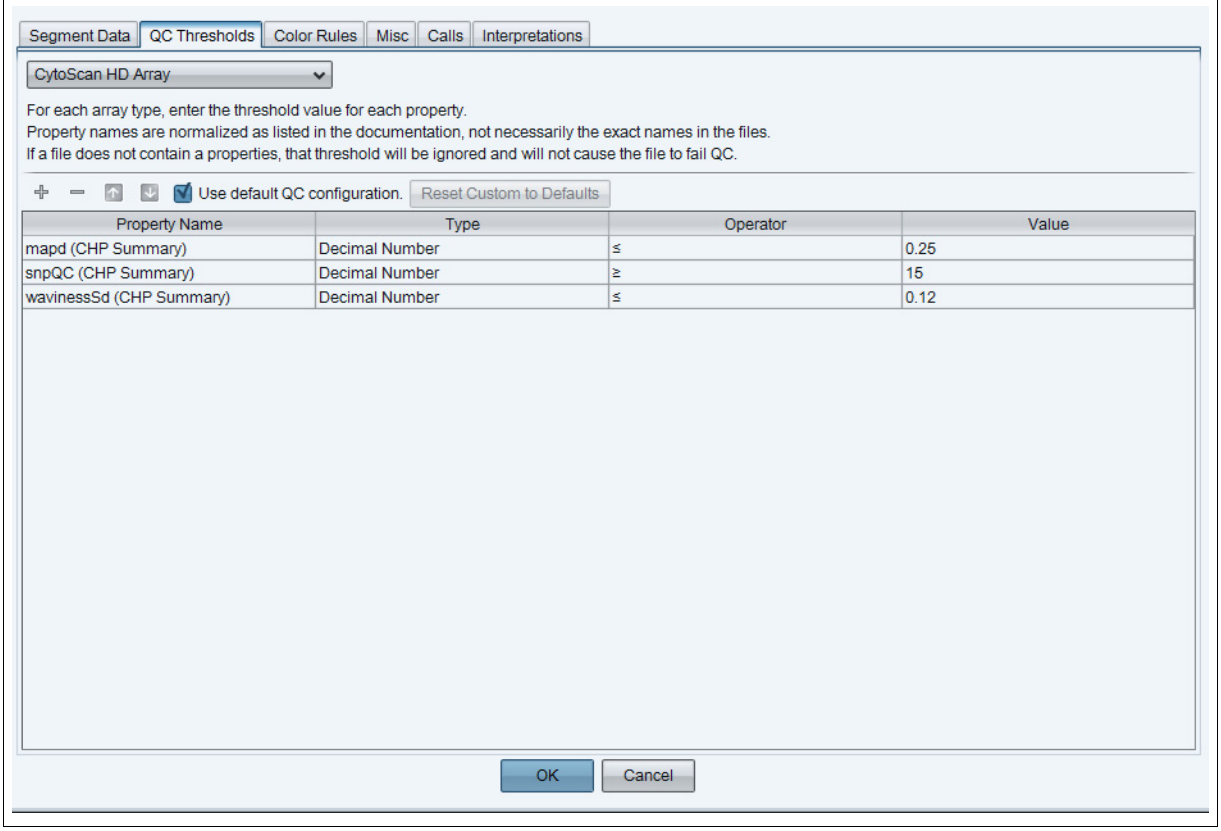
NOTE: Affymetrix does not recommend selecting different QC metrics.

To view the QC thresholds:

1. Click **Preferences** → **Edit User Configuration**.
   The User Configuration window appears.
2. Click the **QC Thresholds** tab (Figure 5.15).
3. Select an array type from the drop-down list.

NOTE: QC parameters are specified per array type.

Figure 5.15 CytoScan HD array QC Thresholds default settings.



QC Thresholds Tab Options

Property Name

snpQC is a QC metric for SNP probes that is derived from polymorphic (SNP) probes

mapd – A QC metric for all probes used to determine copy number that is derived from both polymorphic (SNP) and non-polymorphic (CN) probes

Waviness SD - A global measure of variation of microarray probes that is insensitive to short-range variation and focuses on long-range variation.

ndSNPQC is a QC metric for SNP probes that is derived from polymorphic SNP probes in normal diploid regions. (OSCHP only).

CelPairCheck is a test that inspects each pair of intensity (*.cel) files to determine whether the files have been properly paired and assigned to the correct channel. (OSCHP only).

Note: The property names are from the header information of the xxCHP file.

Type Value or algorithm used for that type of QC.

Operator The type of comparison performed.

Value Value assigned to the threshold.

The default QC thresholds are:

Table 5.1 Default QC thresholds

Array Type	QC Parameter		
	mapd	**snpQC**	**Waviness SD**
CytoScan® HD Array	≤ 0.25	≥ 15.0	≤ 0.12
Genome-Wide Human Array SNP 6.0	≤ 0.35	-	-

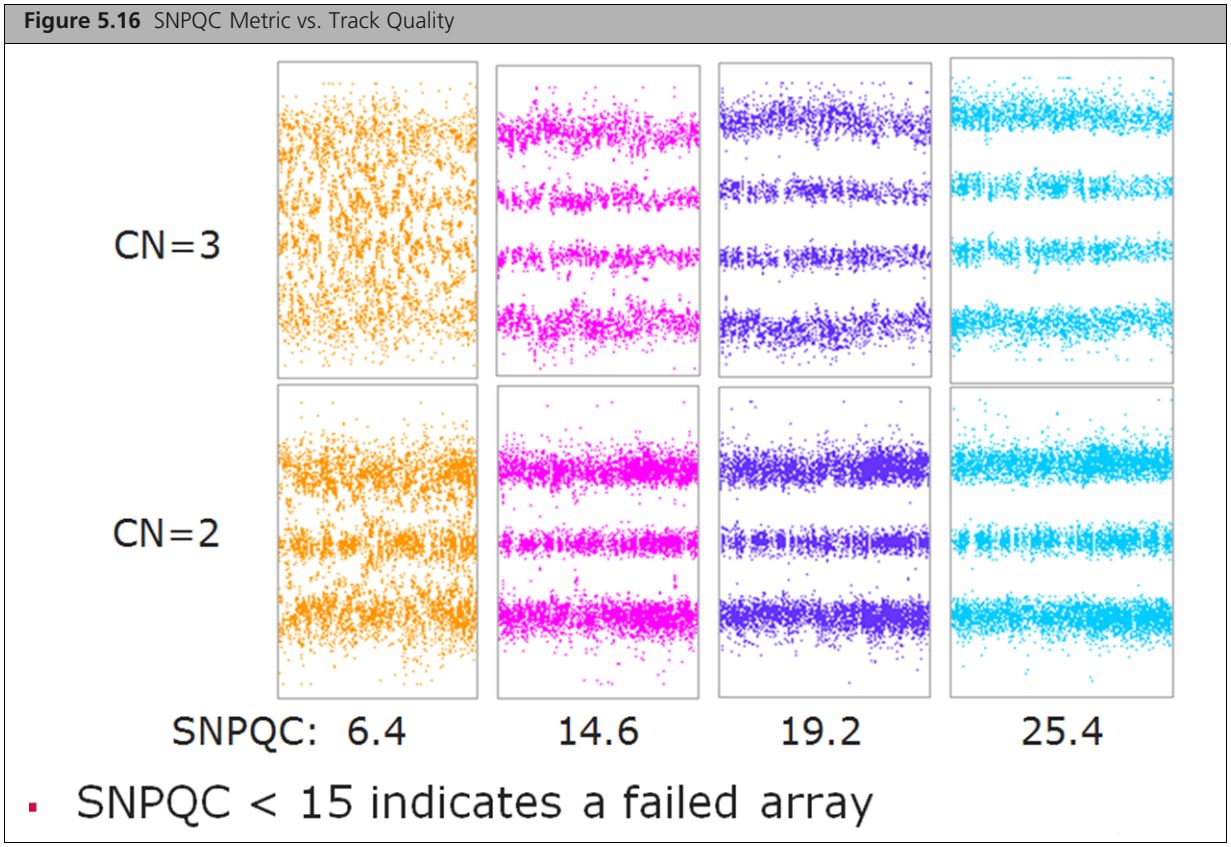
Table 5.2 Default QC metrics for OncoScan Assay

Array Type	QC Parameter		
	mapd	**ndSNPQC**	**CellPairCheck Status**
OncoScan FFPE Assay	≤ 0.3	≥ 26	Pass

NOTE: The waviness SD metric is applicable to blood and cell line data. The waviness SD metric is not intended for alternative sample types such as solid tumor or FFPE samples in which the results may vary as a result of the biological complexity.

Effect of SNPQC value on the Allele Peaks Track. (Figure 5.16)

Figure 5.16 SNPQC Metric vs. Track Quality



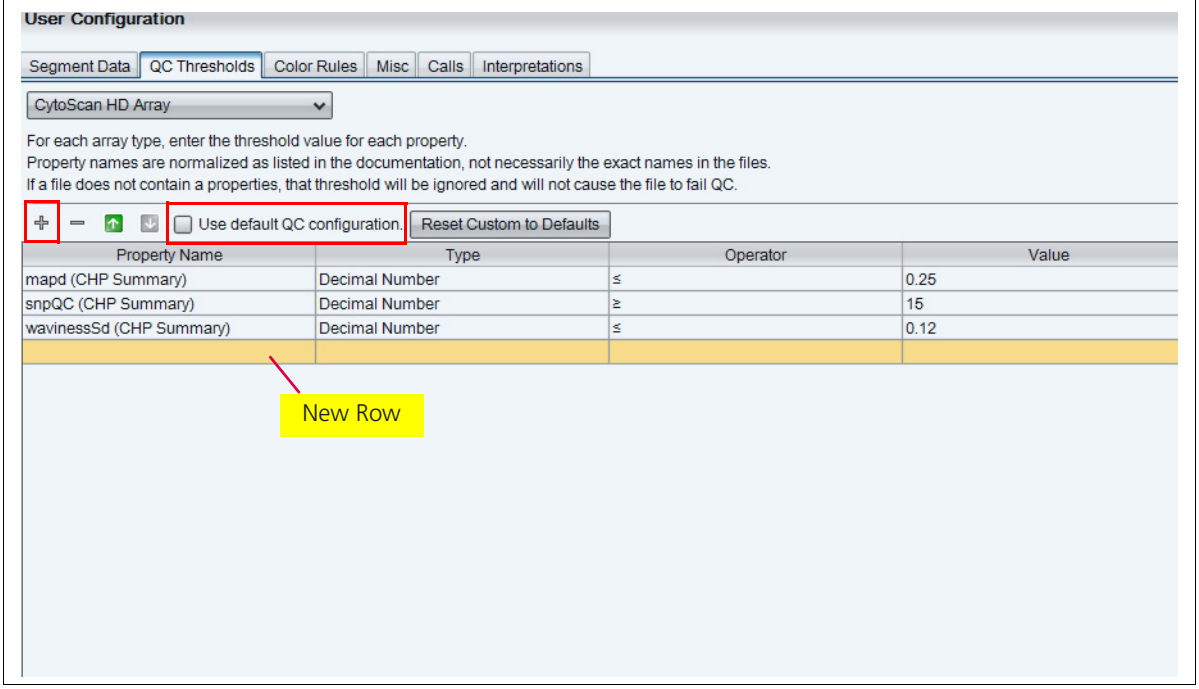
SNPQC is one of the CytoScan® within-array QC metrics which provides insight into the overall level of data quality from a SNP perspective. When evaluating the SNPQC values, the key consideration is to ensure that the threshold is exceeded. The quality of the SNP allele data is compromised, from an interpretation perspective, when the SNPQC values are below the recommended acceptance threshold as illustrated by the two left most graphs representing the two and three copy allelic state. When the SNPQC value is below 15, as illustrated by the data represented in the two graphs above, the noise within the array is higher than expected which compromises the overall data quality and clarity of results. However, when the SNPQC value is above 15, the consideration is whether the SNPQC value is above or below the threshold value and not the absolute magnitude.

As long as the SNPQC value exceeds the threshold, there is a retention in the data quality as illustrated by the graphs which demonstrate clear allelic data across a broad range of SNPQC values that exceed the recommended threshold. The threshold was determined from thousands of arrays processed across multiple reagent lots, operators, and sample aberration types. SNPQC is one of the metrics used to assess array quality and should be helpful in determining which experimental data sets are of satisfactory quality to continue with subsequent interpretation.

Adding a QC Property

1. Click **Preferences → Edit User Configuration**.
   The User Configuration window appears.
2. Click the **QC Thresholds** tab (Figure 5.17).

Figure 5.17 Adding a new row to the QC Thresholds table



3. Select an array type from the drop-down list.
4. In the QC Thresholds tab, *uncheck* the **Use default QC configuration** checkbox.
5. Click the **Add** button .

   A new row appears in the table.

   To delete a property row, select the row, then click the **Remove** button.
6. Click the **Property Name** field, then enter a new QC property name.
7. Click the **Type** field to select **Decimal Number** or **Whole Number**.
8. Click the **Operator** field to select an operator from the drop-down list.
9. Double-click the **Value** field to enter the threshold value for the newly added QC property.
10. To add another QC property, repeat steps 5 - 9.
11. Click **OK** to apply the newly added QC threshold(s).

Editing an Existing QC Threshold

1. Click on an existing **Property Name**, then edit its QC property name.
2. Click the **Type** field to edit its **Decimal Number** or **Whole Number**.
3. Click the **Operator** field to choose a different operator from the drop-down list.
4. Double-click the **Value** field to edit the current threshold value.
5. To edit another existing QC property, repeat steps 1 - 4.
6. Click **OK** to apply the newly edited QC threshold(s).

Viewing Data in ChAS

The loaded xxCHP data can be displayed in:

- Graphic Displays

  See Displaying Data in Graphic Views (on page 91)

- Tables

  See Displaying Data in Table Views (on page 228)

After the data is loaded, you can:

- Filter the segments by Segment Parameters to hide segments that do not meet your requirements for significance.

  See Filtering Segments (on page 142)

- Select a region information file for use as a CytoRegion file and:
  - Perform differential filtering for segments in CytoRegions and in the rest of the genome.
  - Display only segments that appear in CytoRegions using Restricted Mode.

  See Using CytoRegions (on page 178)

- Select a region information file for use as an Overlap Map and use the Overlap filter to conceal segments that overlap with the Overlap Map items. This functionality may be helpful for tracking or filtering out benign copy number change regions.

  See Using the Overlap Map (on page 188)

- Add selected features of the genome to new or existing Region (AED) files, and edit annotation data on existing annotations.

  See Creating and Editing AED Files (on page 196)

- Prepare reports on your findings by exporting graphics and table data in PDF and other formats.

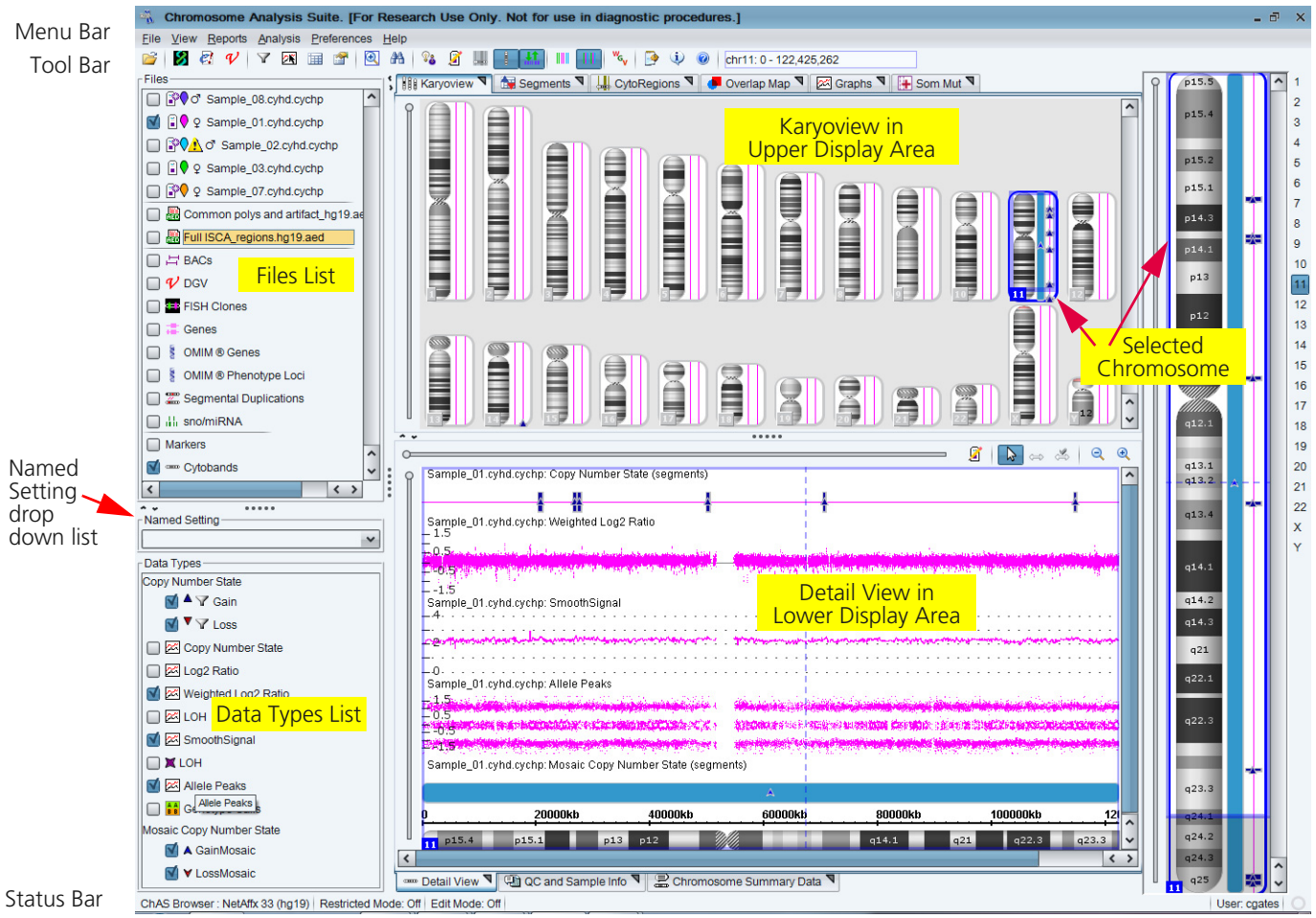
  See Reporting Results (on page 269)

- Save setups of ChAS for different tasks in user profiles and named settings.

  See User Profiles and Named Settings (on page 289)

Overview of ChAS Window Components

Figure 6.1 ChAS window



The ChAS window components (Figure 6.1) include:

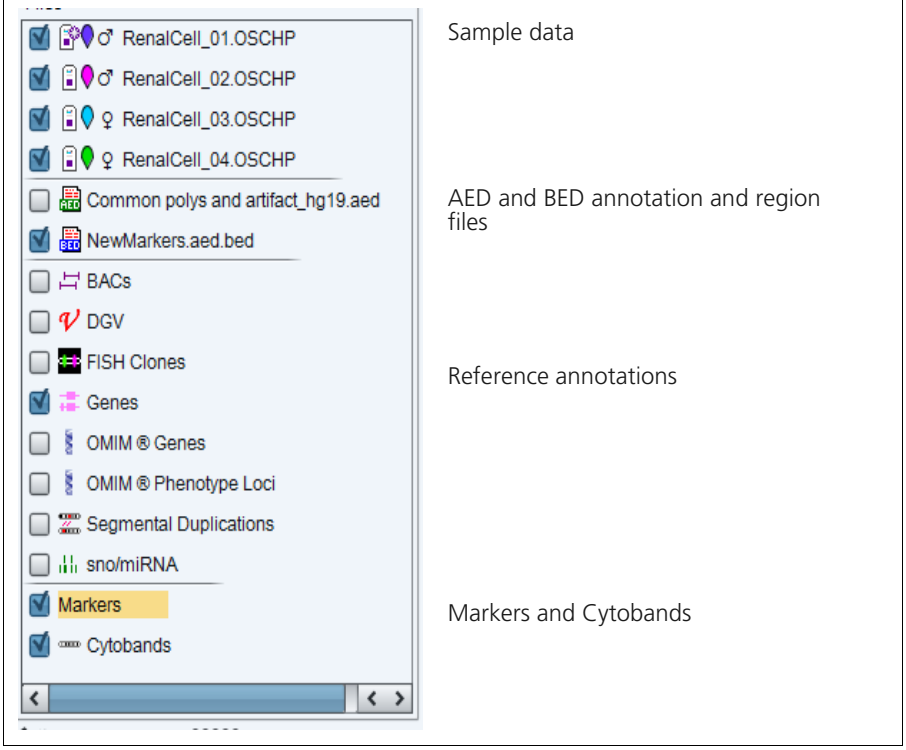
- **Menu Bar**: Access to the functions of the software.
- **Tool Bar**: Quick access to commonly used functions.
- **Files List**: Displays data and annotation files that can be displayed. (page 85)
- **Data Types List**: Displays the type of data available in the files. (page 86)
- **Named Settings**: Displays a list of the previously saved display settings for ChAS. (page 87)
- **Status Bar** (page 87): Displays information on:
  - The status of the software.
  - The ChAS Browser NetAffx Genomic Annotation file version.
  - The hg version.
  - The annotation or probe that the mouse pointer is nearest to in the Detail View.
- **Display Area** (page 88): Displays the following data in graphical and table formats:
  - xxCHP graph data.
  - Detected Segments.

- Regions
- Reference Annotations.

Files List

The Files list (Figure 6.2) displays the different sources of data and annotations that are loaded in the Chromosome Analysis Suite. Files are grouped by type in the Files list.

Figure 6.2 Files list



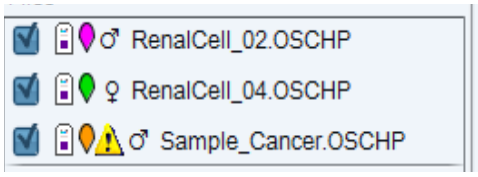
Displays the files grouped by type:

- Sample Data

  Colored nibs display the color used for the data lanes for that sample in the Karyoview, Selected Chromosome View, and Detail View. The appropriate gender symbol is displayed to the right of the colored nib.

  If a loaded file has a QC parameter that is out of range, an alert symbol appears next to the file name (Figure 6.3).

Figure 6.3 QC alert for Sample Data files



- Region Information Files

Icons indicate the file type (AED or BED) and whether the loaded files have been selected as a CytoRegion file or Overlap Map file.

- Reference Annotations (loaded during software start-up)

  Icons indicate the annotation type.

- Cytobands (Separated from other reference annotations because cytobands cannot be moved in the displays)

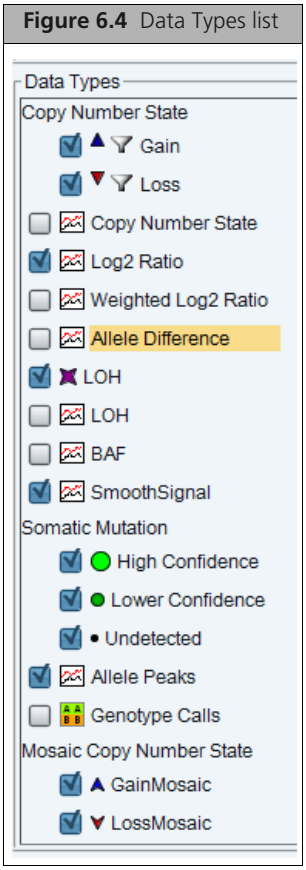
See Selecting Data for Display (on page 116) for information on using the Files list to select loaded data and region information files, and reference annotation for display.

You can export the feature information in these files to a new AED region information file. *See Exporting Information in AED or BED Format (on page 224)*.

Data Types List

The Data Types list (Figure 6.4) shows the types of data available for display in the Karyoview, Selected Chromosome View, and Detail View. The available data types may vary, depending upon the type of sample data available. *See Introduction to Loading Data (on page 66)*.

Figure 6.4 Data Types list



The Data Types list allows you to select from Segments data and Graph data.

The Segments data is displayed graphically in:

- Karyoview
- Selected Chromosome View
- Detail View

If filtering is applied to a segment type, a funnel icon appears next to the segment symbol in the list.

Graph data, indicated by the icon, is displayed only in the Detail View. See *Detail View* (on page 104) for more information.

Unselected data is also concealed from the different tables and graphs.

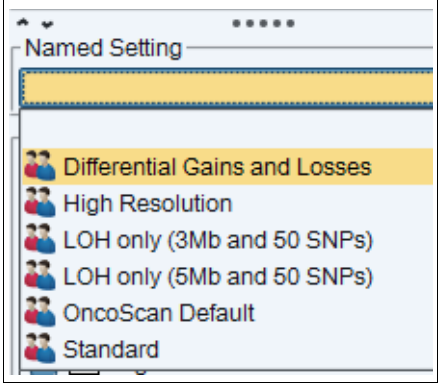
See Selecting Data Types for Display (on page 118) for information on using the Data Types list to select different data types for display.

Named Settings

The Named Settings drop-down list (Figure 6.5) enables you to apply a previously created setting for ChAS. The settings may include things like:

- Segment Filter and Overlap Map Filter settings
- Types of data to be displayed
- Restricted Mode ON/OFF status. *See Using Restricted Mode (on page 186)*

Figure 6.5 Named Settings



NOTE: Affymetrix-supplied Named Settings (indicated by the icon) should not be deleted from the system because they are shared by all user profiles.

See Chapter 15, *User Profiles and Named Settings* on page 289 for information on creating and using Named Settings.

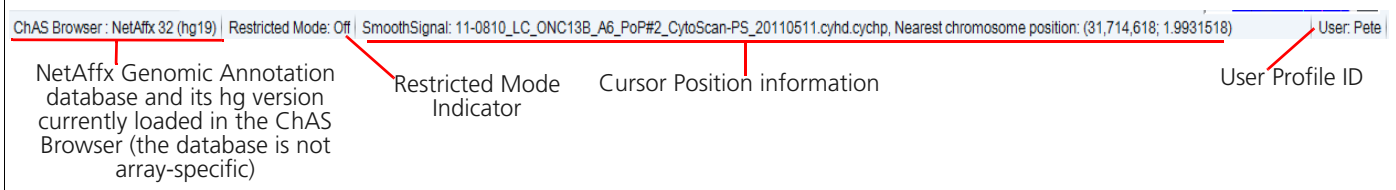
NOTE: The OncoScan Default Named Setting has no filters applied so all segments called by the TuScan algorithm can be viewed. Users can create their own appropriate custom filter Named Setting for their data, see *User Profiles and Named Settings* (on page 289).

Status Bar

The status bar (very bottom of browser) (Figure 6.6) displays information on:

- NetAffx Genomic Annotation database and its hg version
- Restricted Mode status (*See Using Restricted Mode (on page 186)*)
- Edit Mode status (*See Using Edit Mode (on page 146)*)
- Cursor (Mouse Over) Position
- User Profile ID

Figure 6.6 Status Bar

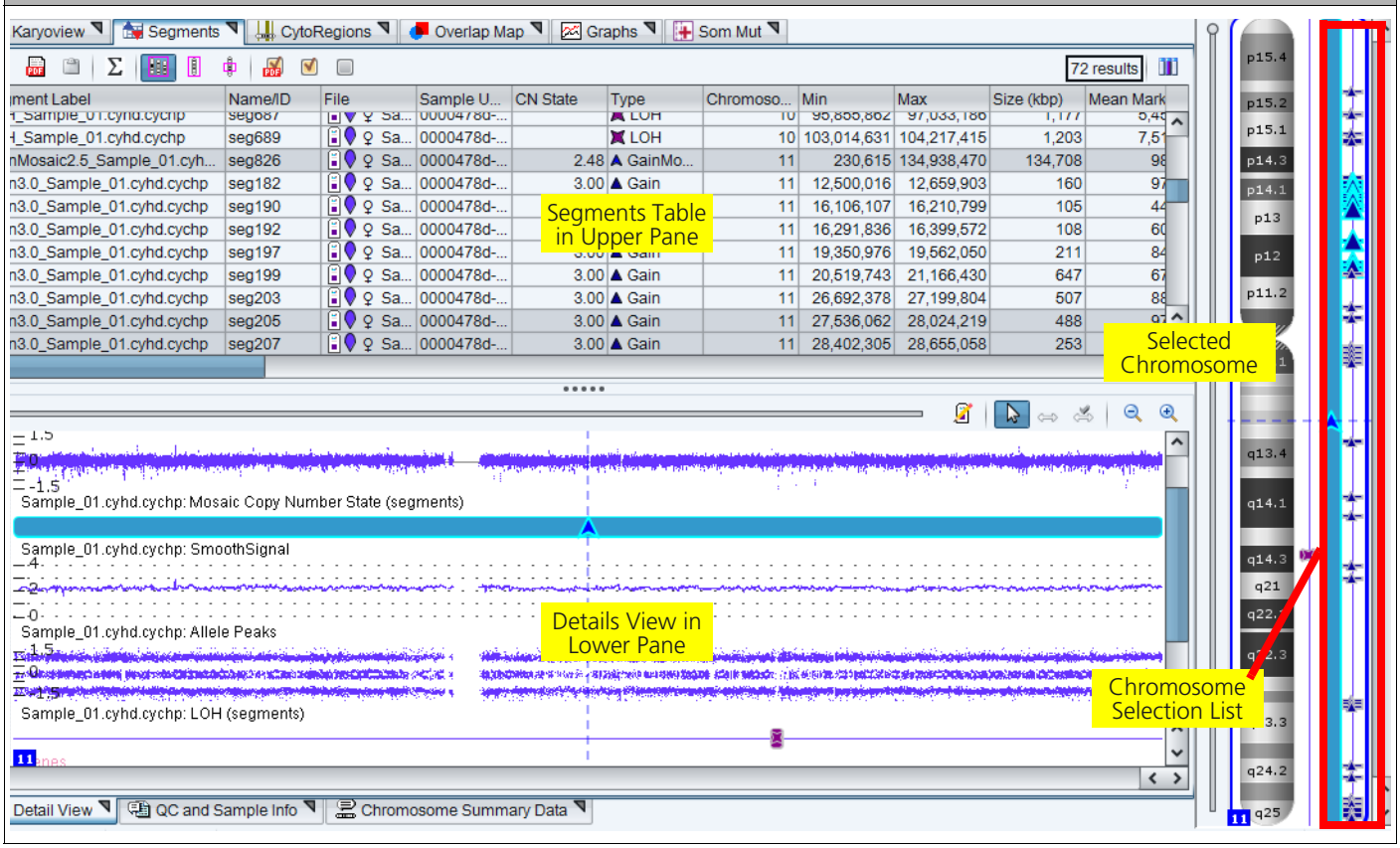


Display Area

The Display Area (Figure 6.7) is divided into three panes:

- *Upper Panes* (on page 89)
- *Selected Chromosome View* (on page 89)
- *Lower Panes* (on page 90)

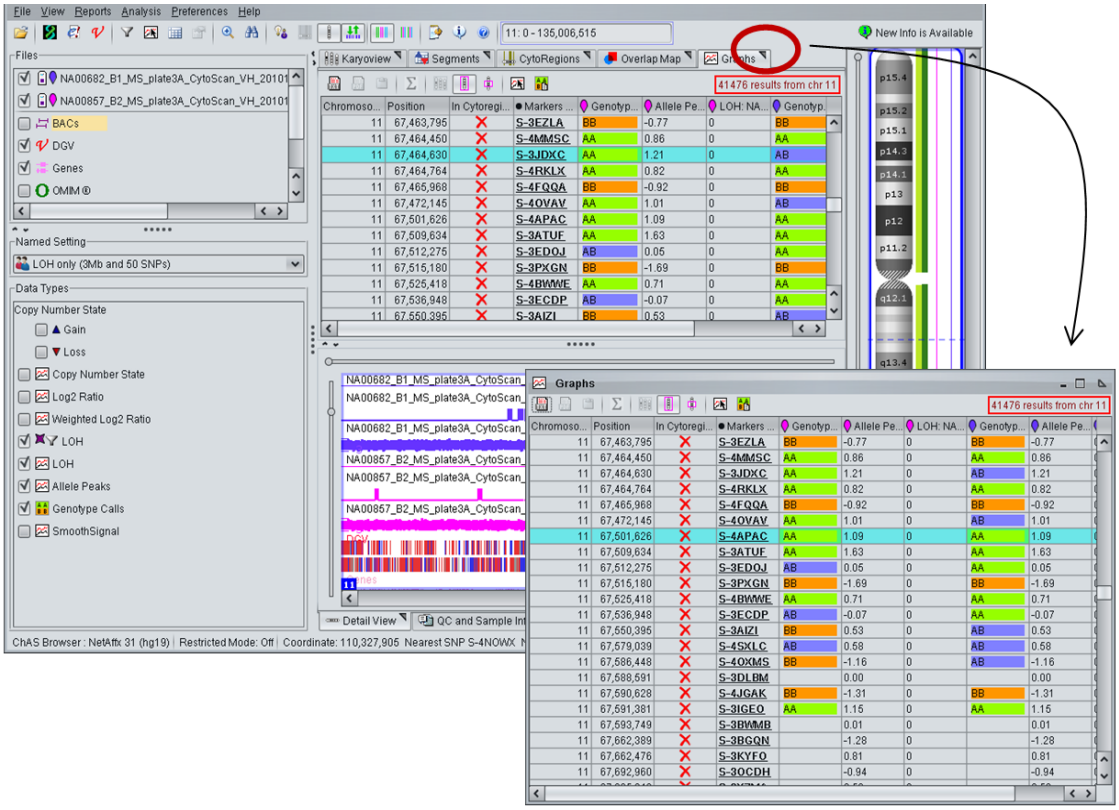
Figure 6.7 Display Area showing Segments Table and Detail View



The tabs in the upper and lower panes display different types of data, in both graphical and table formats. Data from the same sample files is displayed in all three panes.

You can display a pane in a separate window by clicking the icon on the tab (Figure 6.8). To close the window and return the information to the tab panel, click the icon in the window.

Figure 6.8 Graphs pane opened in new window



Upper Panes

The Upper pane displays data in graphics and table formats:

- *Karyoview* (on page 93): Displays selected segment types for selected sample files for all chromosomes.
- *Segments Table* (on page 236): Displays a list of the detected segments in the selected sample files.
- *CytoRegions Table* (on page 182): Displays a list of the Regions in the AED/BED file selected as the CytoRegion file. Includes information on detected segments which lie in CytoRegions.
- *Viewing the Overlap Map Table* (on page 192): Displays a list of the Regions in the AED/BED/ Reference Annotation file selected as the Overlap Map. Includes information on detected segments that are overlapped by Overlap Map Items.
- *Graphs Table* (on page 248): Displays marker data for the loaded and selected xxCHP files.

Selected Chromosome View

The *Chromosome View* displays detected segments in selected sample files for the chromosome selected in the Karyoview, while the *Chromosome Selection List* (far right column) displays its number.

See *Selected Chromosome View* (on page 101) for more information.

Lower Panes

The Lower Pane displays:

- *Detail View* (on page 104): Displays the selected section of the chromosome displayed in Selected Chromosome View, with
  - Detected segments and graph data in selected xxCHP files
  - AED/BED file regions and annotations.
  - Reference annotation files
- *QC and Sample Info Tab* (on page 255): Displays QC metric information as well as information about the loaded Sample and Region (AED) files.
- *Chromosome Summary Data* (on page 258):
  - Summarizes particular data across each chromosome in the loaded sample data files (for example, proportion of each chromosome found to be in the state of LOH).
  - Calculated Properties: calculates the Percent Autosome LOH in the sample based on the LOH filter setting.

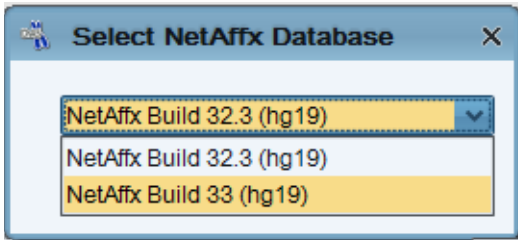
Changing the ChAS Browser NetAffx Genomics Annotation File Version

You can change the NetAffx Genomic Annotations file loaded in the ChAS browser.

1. Select **File → ChAS browser NetAffx Genomic Annotation file version** on the menu bar
2. In the dialog box that appears (Figure 6.9), select a NetAffx Build from the drop-down list and click **OK**.

NOTE: If there are loaded xxCHP files with hg version different from the selected NetAffx database, a message informs you that these data files will be closed before the NetAffx database is loaded.

Figure 6.9 Select a NetAffx Genomics Annotation file version



For more information on what types of data is in the database files and how this information varies in different versions, see *ChAS Browser NetAffxGenomic Annotations* (on page 311).

Displaying Data in Graphic Views

ChAS provides different graphic views for the detected segment and other data. These views enable you to:

- Get an overview of the detected segments
- Compare segments between samples
- Drill down to examine areas of interest in more detail
- View the graph and marker information used to generate the detected segments
- Take advantage of reference annotations and external web sites
- Create your own Affymetrix Extensible Data (AED) files with regions of interest and annotations

This chapter discusses:

- *The Graphic Display Views* (on page 92)
- *Controlling the Display of Data* (on page 115)
- *Edit Configurations in the Miscellaneous (Misc) Tab* (on page 111)
- *Learning More About Features* (on page 132)

The Graphic Display Views

The data can be displayed in the following graphic views (Figure 7.1):

- *Karyoview* (on page 93)
- *Whole Genome View* (on page 97)
- *Selected Chromosome View* (on page 101)
- *Detail View* (on page 104)

Figure 7.1 Graphic display views

Data from the same sample files is displayed in all the views at different scales.

If an item in any of the views is selected, the icon for that item is enlarged or highlighted in the views.

Karyoview

The Karyoview (Figure 7.2) displays a genome-wide view of the detected segments and other data. In the Karyoview:

- Click a chromosome in the Karyoview to select it.
- Press Ctrl + Left/Right Arrow keys to move between chromosomes.
- To jump to chromosome 1, press Ctrl+Home
- To jump to chromosome Y (last chromosome in the Karyoview), press the Ctrl+End.

Figure 7.2 Karyoview Controls

Using the mouse, click and drag on a selected chromosome to select an area for display in the Detail View. This area is highlighted in the Selected Chromosome View (Figure 7.3).

Figure 7.3 Karyoview, Selected Chromosome View, and Detail View

You can use the Stretch Slider and Vertical Scroll Bar to zoom in on a section of the Karyoview. You can also use the mouse wheel as shown below:

- Alt + mouse wheel stretches the display
- Mouse wheel scrolls up and down

The following information is displayed for each chromosome (Figure 7.4):

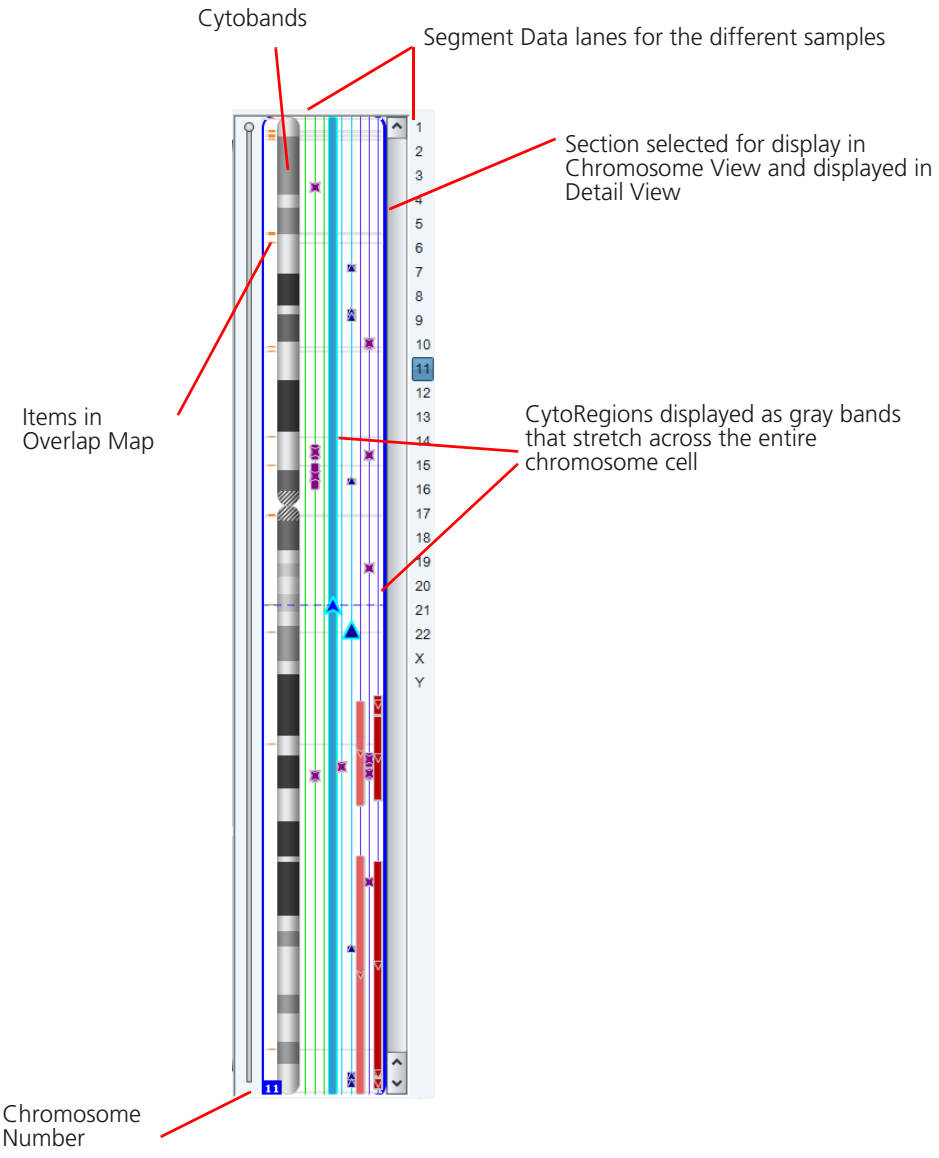
- Chromosome number
- Cytobands
- Segment Data, with separate lanes for each sample file and each displayed segment type.

  Each sample is assigned a unique color in the display that is used for the lane.

  You can:

  - Change the grouping of samples and segment types (see *Changing the Grouping of Samples and Data Types* (on page 119)).
  - Change the position of samples (see *Lanes Grouped by Sample* (on page 119)) and data types (*Tracks Grouped by Data Type* (on page 122)) in a group.
- Somatic Mutation Data (Only available for OSCHP file data)
- CytoRegions in selected CytoRegion File
- Items in selected Overlap Map file

Figure 7.4 Detail of Selected Chromosome



Click on a chromosome in the Karyoview to select it for display in the Selected Chromosome View and the Detail View. The selected Chromosome is highlighted in the Karyoview.

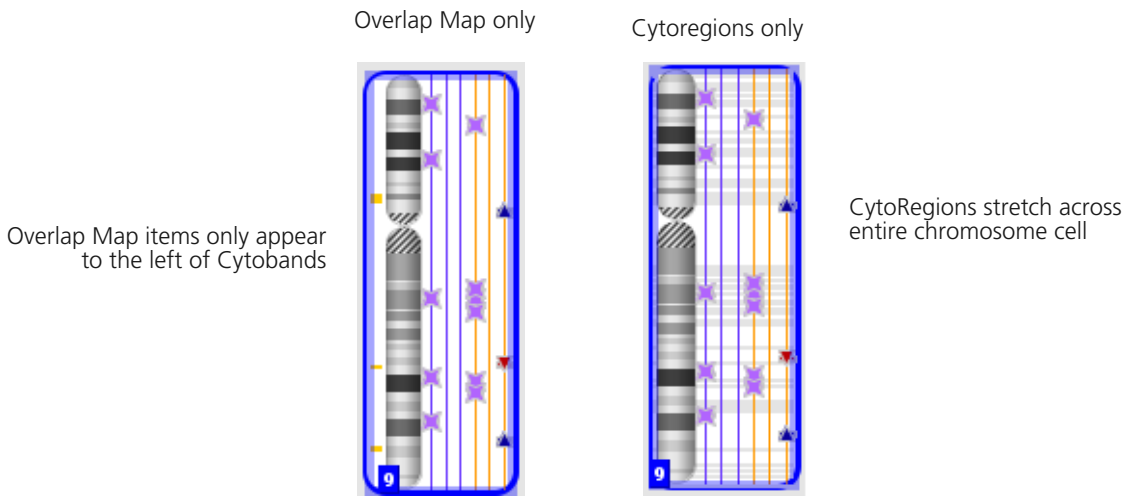
The Stretch Slider and the vertical scroll bar controls the vertical stretch of the Display area.

At higher magnifications, more details of the cytobands are displayed in the Karyoview. You can see cytoband labels if the display has room.

The portion of the chromosome selected in the Selected Chromosome View and displayed in the Detail View is highlighted in the Karyoview.

If you have selected a CytoRegions file, the cytoregions are displayed as gray bands that stretch across the entire chromosome cell, from right to left of the cytobands (Figure 7.5).

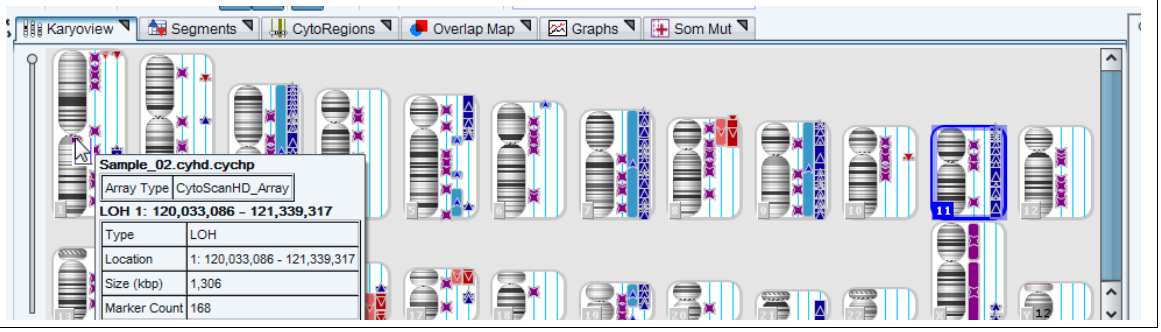
Figure 7.5 Display of Overlap Map items and CytoRegions in Karyoview



If you selected an Overlap Map file, the overlap map items are displayed as small rectangles to the left of the cytobands. Its color is the same color used in the Details View.

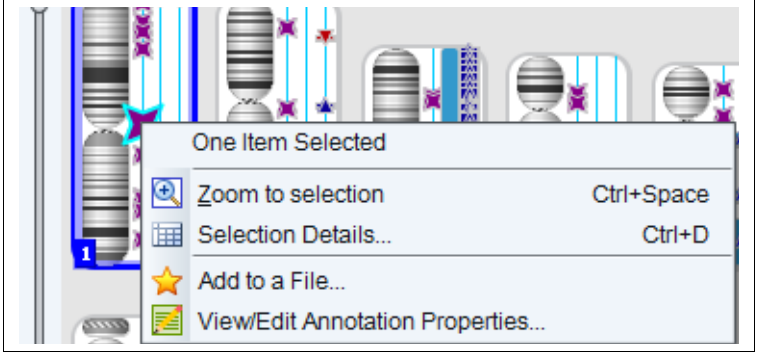
You can mouse over a feature in the Karyoview, Selected Chromosome View, or Detail View to display a pop-up (Figure 7.6) with information about the feature.

Figure 7.6 Karyoview pop-up



You can right-click on a feature in the Karyoview or Selected Chromosome View to open a shortcut menu of options (Figure 7.7).

Figure 7.7 Karyoview shortcut menu



See *Learning More About Features* (on page 132) for more information.

Whole Genome View

The Whole Genome View (Figure 7.8) displays the following data in a single display from Chromosome 1 to Y for:

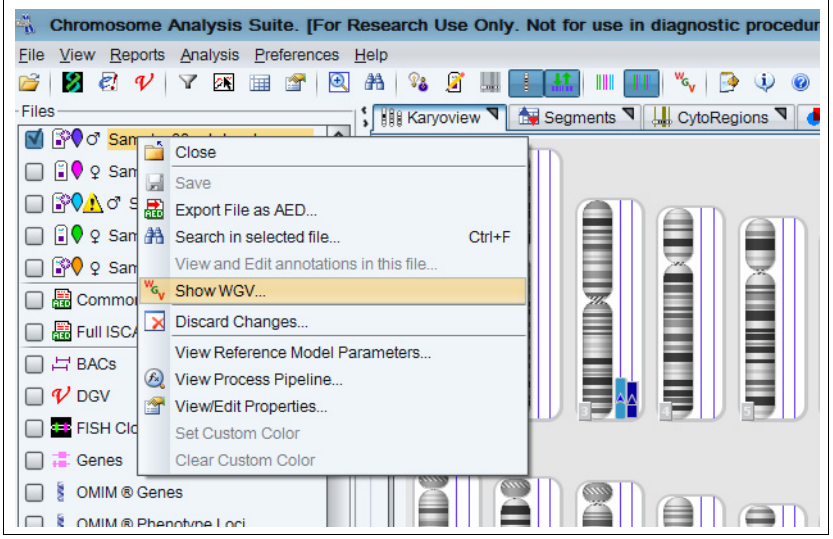
- CYCHP: Log2 Ratio, Weighted Log2 Ratio, Copy Number State, Smoothed Signal, Allele Peaks,.
- CNCHP: Log2 Ratio, Copy Number State, Smooth Signal, Allele Difference.
- OSCHP: Log2 Ratio, Weighted Log2 Ratio, Copy Number State, Smoothed Signal, Allele Difference, B-allele Frequency. Note: For OSCHP files it is recommended to use the Log2 Ratio and Copy Number State for the top graph.

The Y axis on the left represents the data points in that graph (Log2/Weighted Log2 Ratio). The Y axis on the right represents the line graph data (Smooth Signal/Copy Number).

To display the Whole Genome View for a sample:

1. Right-click on the same name in the **File Tree** (Figure 7.8), then select **Show WGV**.

Figure 7.8 Right-click menu - File Tree

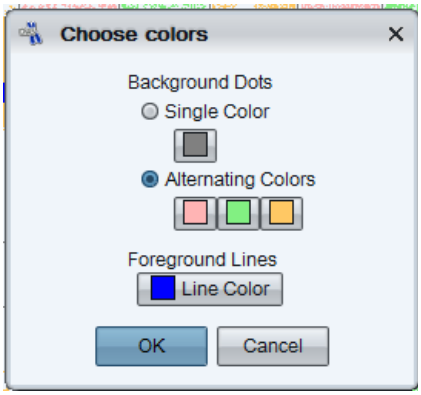


2. The Whole Genome View for the selected file opens. (Figure 7.9)

Figure 7.9 Whole Genome View

3. To select a different sample to display:
   - Select **File** → **Choose File**.
   - Click on a different file name, then click **OK**.
4. Graphs may be viewed in a single color or alternating colors every 3 chromosomes. To change the color of the data points or line data in the WGV, click on **View** → **Colors**. Select the radio button to change to the desired color. You may select a different color for the Data Points and Foreground Lines, by clicking on the colored square. A color palette appears allowing you to select the color of choice. (Figure 7.10)

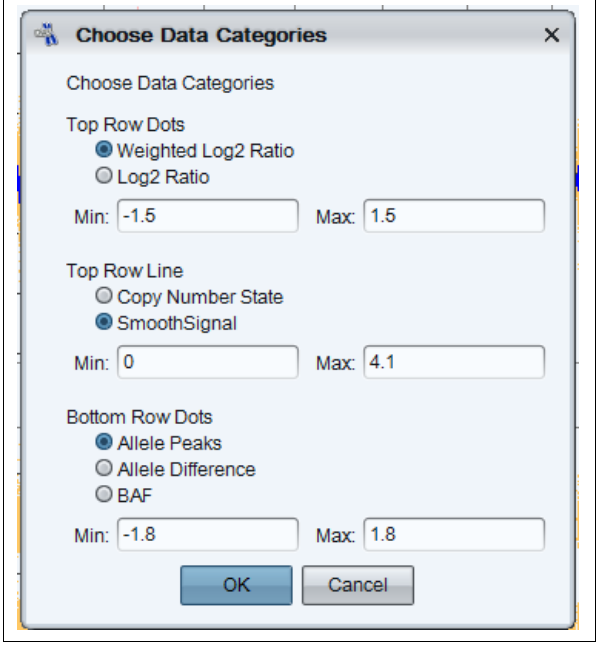
Figure 7.10 Whole Genome View - Choose colors window



5. To change which graph is currently in view, select **View** → **Choose Data**. A Choose Data Category Window appears allowing you to select the radio button for:
   - Top Graph: Log2 Ratio or Weighted Log2 Ratio for the data points and either Smooth Signal or Copy Number State for the line data.

- Bottom Graph: Allele Peaks, Allele Difference or B-allele Frequency (BAF). (Figure 7.11)

Figure 7.11 Choose Data Categories



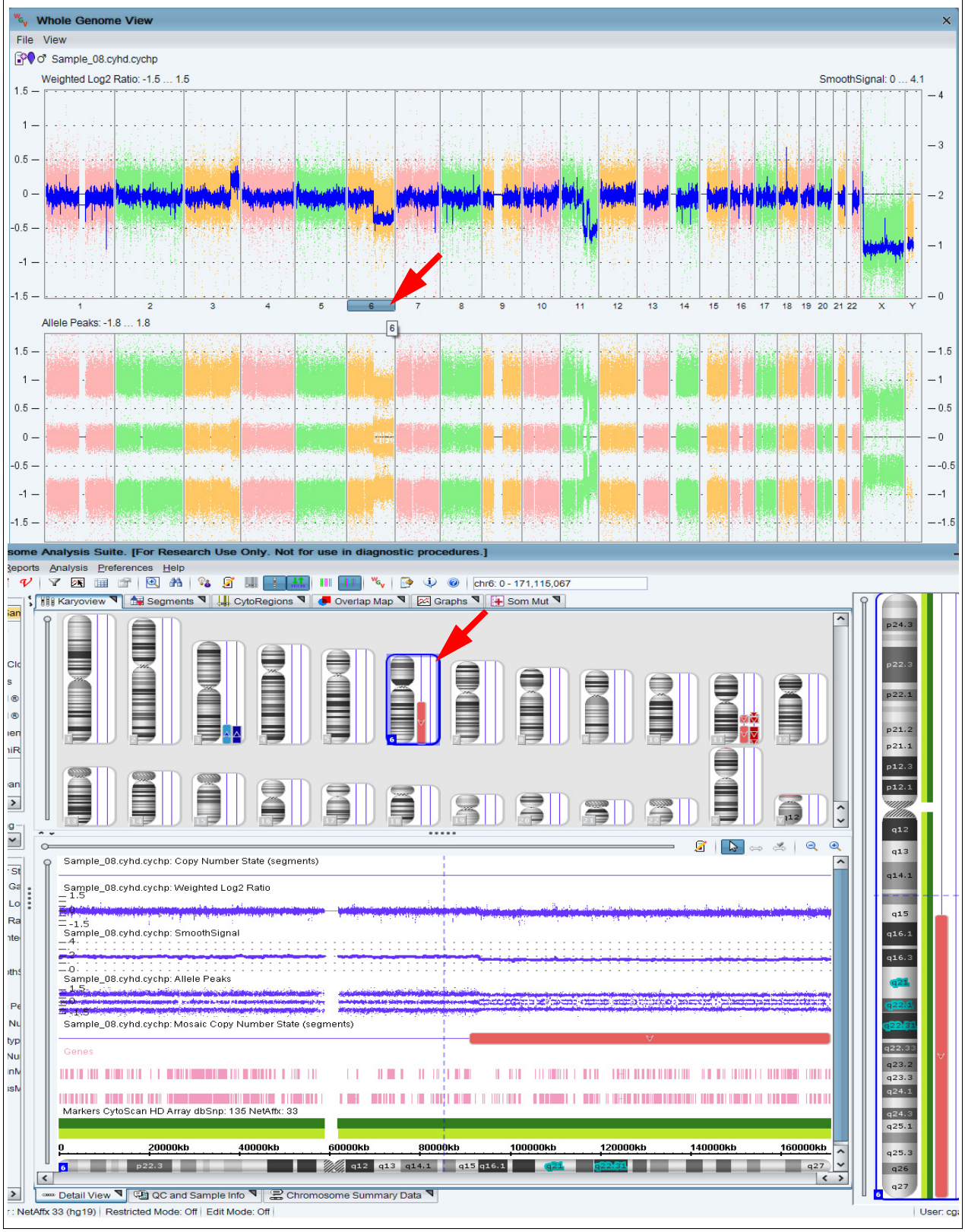
NOTE: The Y axis settings are initially determined from the Y axis settings set in the Detail View Graph Settings.

Not all graphs are available for a given xxCHP. If a graph type is selected in which data is not available, the graph will appear with no data points.

TIP: To return all settings back to the default options, click *View* → *Reset to Defaults*.

- You can export the Whole Genome View as a PNG file. To do this, click on **File** → **Export window PNG**.
- You can jump to a chromosome of interest from the Whole Genome View to the Browser's Graphic Views by clicking on your chromosome number of interest, as shown in Figure 7.12 on page 100.

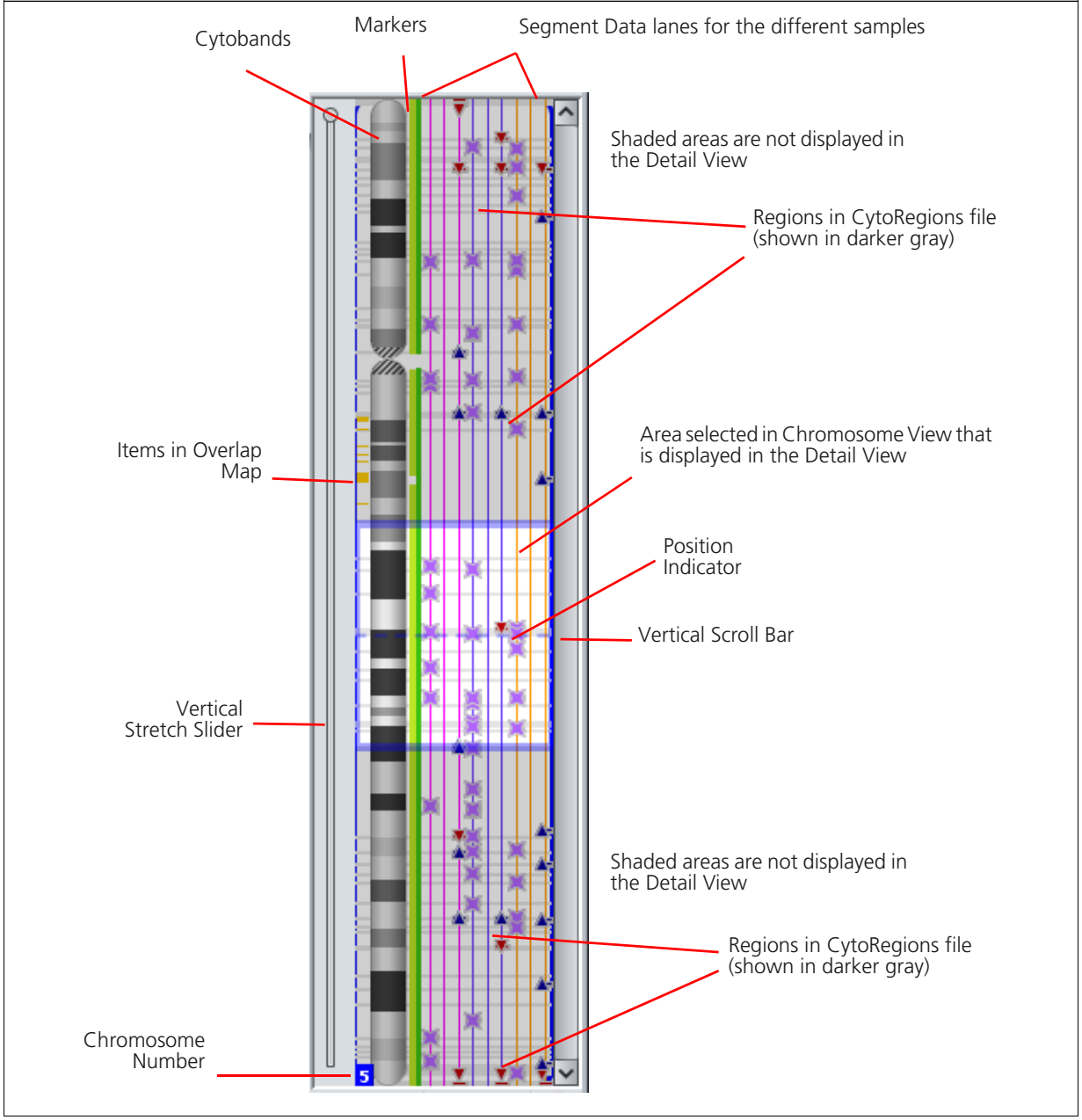
Figure 7.12 Whole Genome View to Browser's Graphic Views



Selected Chromosome View

The Selected Chromosome View (Figure 7.13) is similar to the Karyoview, but it displays a single selected chromosome at higher magnification. Click and drag in the Chromosome View to select an area for display in the Detail View.

Figure 7.13 Selected Chromosome View

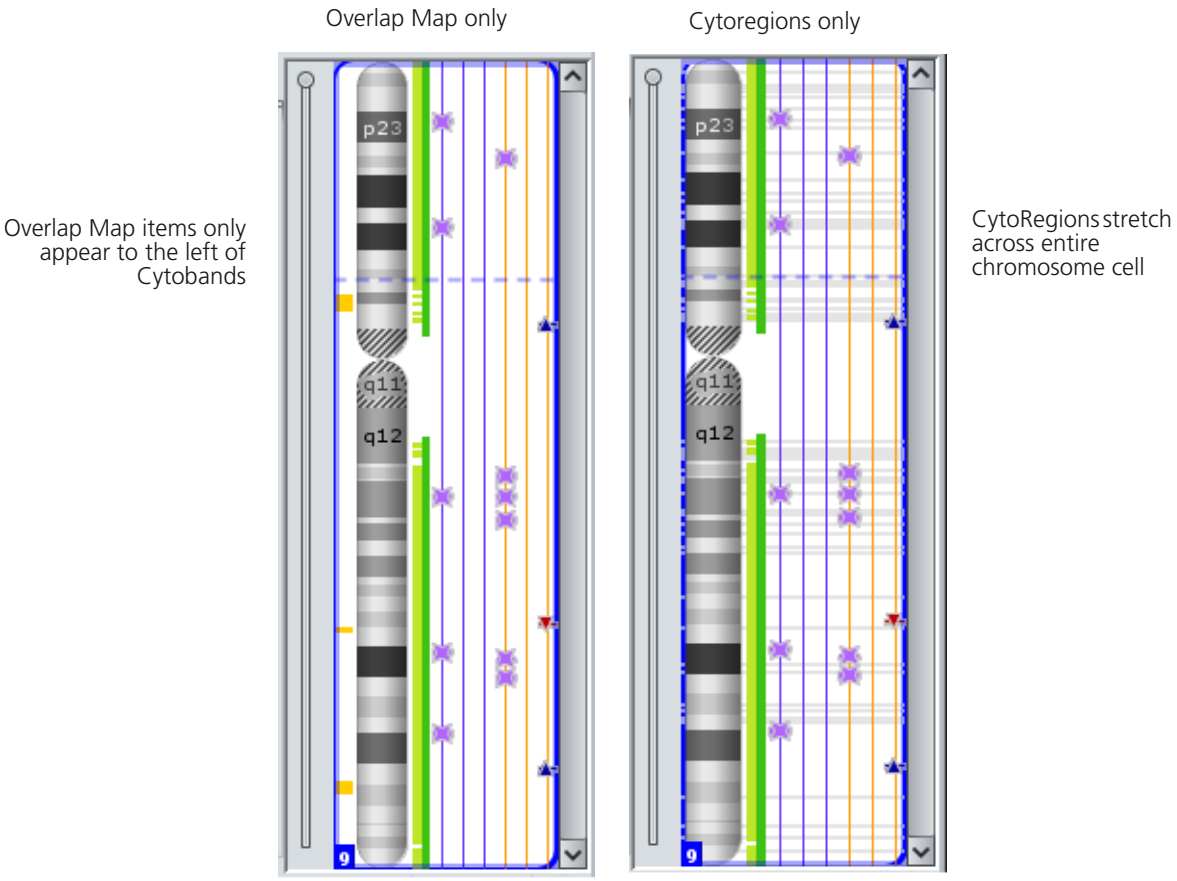


You can use the Stretch Slider and Vertical Scroll Bar to zoom in on a section of the Selected Chromosome View. You can also use the mouse wheel as shown below:

- Alt + mouse wheel stretches the display

If you have selected a CytoRegions file, the cytoregions are displayed as gray bands that stretch across the entire chromosome cell, from right to left of the cytobands.

Figure 7.14 Display of Overlap Map items and CytoRegions in Karyoview

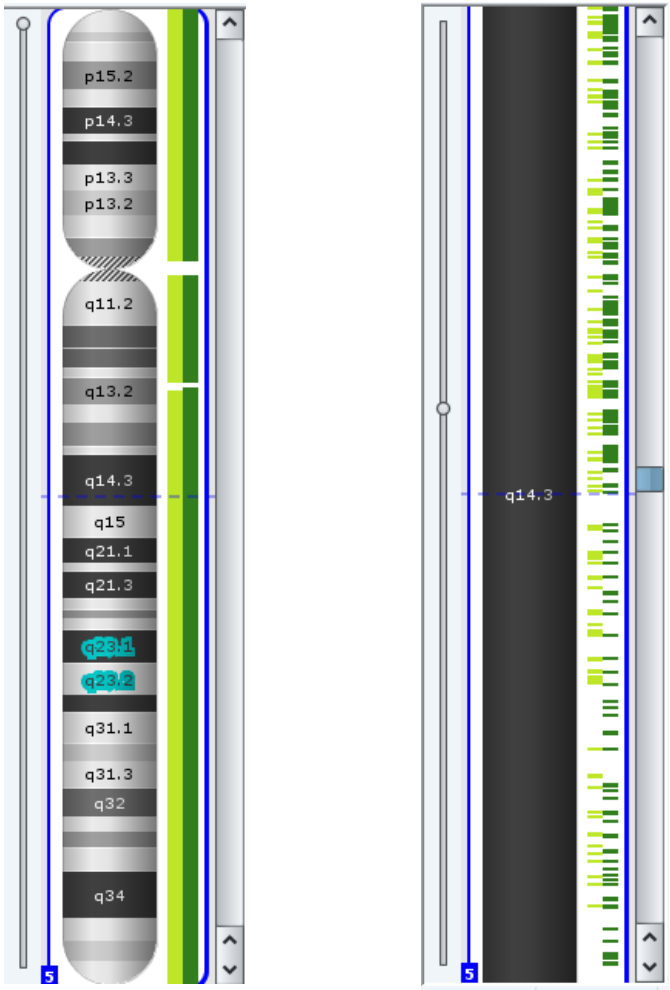


If you have selected an Overlap Map file, the overlap map items are displayed as small rectangles to the left of the cytobands. Its color is the same color used in the Details View.

The Position Indicator is a dashed horizontal blue line. Click in the Selected Chromosome View to set the position of the indicator. The position is highlighted in the graphs table and used as the center point when zooming.

The marker positions are displayed to the right of the cytobands. When zoomed out, they appear as green ribbons. When zoomed in, the markers and their positions can be seen (Figure 7.15).

Figure 7.15 Selected Chromosome View, markers in zoomed-out view (left) and zoomed-in view (right)



- SNP markers are displayed in the light green band nearest the cytobands. The SNP marker/probe names in the CytoScan® HD start with the letter 'S'.
- There is one marker track for every distinct array type that is loaded.
- Copy Number markers are displayed in the dark green band nearest the detected segments. The non-polymorphic copy number probe names on the CytoScan® HD start with the letter 'C'.
- You can mouse over a marker to learn more about it.
- Segments selected in any view are highlighted in the Selected Chromosome View.
- For information on the other features of the Selected Chromosome View, see *Selected Chromosome View* (on page 101).

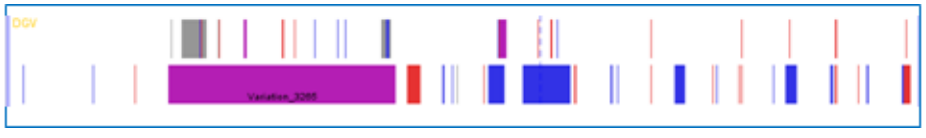
Detail View

The Detail View (Figure 7.16) enables you to look in detail at the detected segments, marker data, regions, and reference annotations for the loaded files.

Figure 7.16 Detail View

Annotation Color Codes

In the Detail View, DGV annotations are color-coded to indicate association with gain or loss.



- Purple – Gain and loss are associated with the region
- Red – Only loss is associated with the region
- Blue – Only gain is associated with the region

- Gray – Copy number variation is associated with the region, but information regarding the number of times gains or losses were observed is not present in the annotation record in the DGV database

Annotation OMIM Color Codes

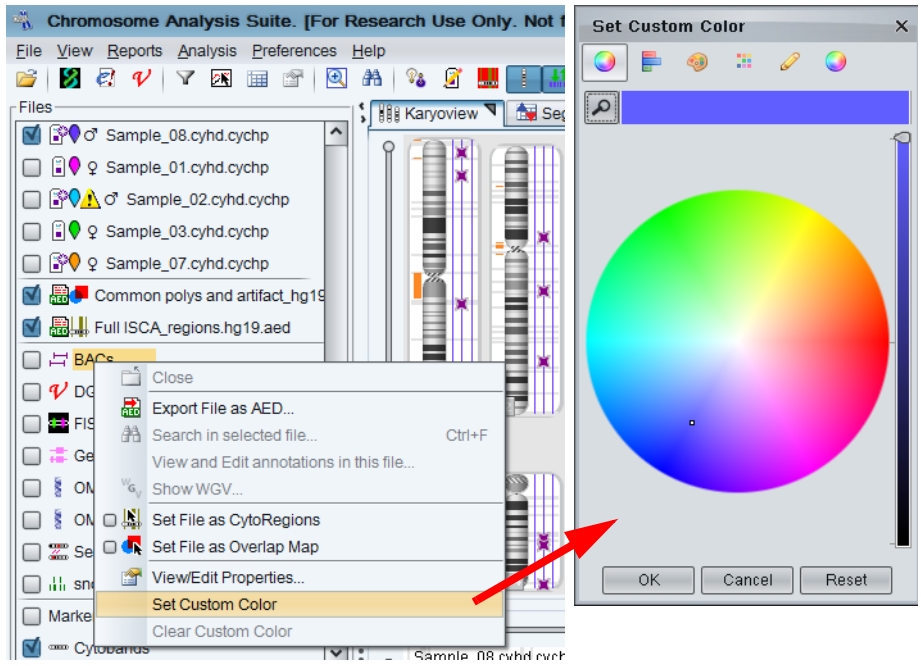
In Browser annotation files version NA32.3 and higher, the following OMIM colored gene entries were generated by genome.ucsc.edu and are based on the associated OMIM phenotype map key. For more information on OMIM display conventions, go to: **www.genome.ucsc.edu**

- Lighter Green for phenotype map key 1 OMIM records - the disorder has been placed on the map based on its association with a gene, but the underlying defect is not known.
- Light Green for phenotype map key 2 OMIM records - the disorder has been placed on the map by linkage; no mutation has been found.
- Dark Green for phenotype map key 3 OMIM records - the molecular basis for the disorder is known; a mutation has been found in the gene.
- Purple for phenotype map key 4 OMIM records - a contiguous gene deletion or duplication syndrome; multiple genes are deleted or duplicated causing the phenotype.
- Light Gray for Others - no associated OMIM phenotype map key info available.

To change an annotation color:

1. Right-click an annotation type in the Files windowpane and select **Set Custom Color** on the shortcut menu (Figure 7.17).

Figure 7.17 Open the color palette



Current Color of BACs annotations in the Detail View

2. Specify a color for the selected annotation type using the color controls in the color palette (Figure 7.17), and click **OK**.

   The new color is applied to the annotations in the Details View.

3. To return to the default annotation color, right-click the annotation in the Files windowpane, and select **Clear Custom Color** on the shortcut menu.

Data in the Detail View

The Detail View displays the following kind of data for CytoScan® HD (CYCHP):

Table 7.1 Data for CytoScan® HD Array

Data Types	Definition
Detected Segment Types	
GainMosaic	Non-integer amplifications or duplications
LossMosaic	Non-integer hemizygous or homozygous deletions
Gain	Amplifications or duplications
Loss	Hemizygous or homozygous deletions
LOH	Loss of Heterozygosity
Probe array data (displayed as graph data)	
Copy Number State	HMM-derived integer Copy Number State
Log2 Ratio	Per marker Log2 Ratio of normalized intensity with respect to a reference, with further correction for sample specific variation.
Weighted Log2 Ratio	Contains the Log2 Ratios processed through a Bayes wavelet shrinkage estimator. These processed values are input to the CNState algorithm HMM.
LOH	Loss of Heterozygosity
Allele Peaks	Filtered and smoothed values for individual markers. Nonparameteric estimation is used to understand possible regional peak structure towards which the data is smoothed. The amount of filtration and smoothing is dynamically adapted based on sample quality. Allele difference is computed based on differencing A signal and B signal, then standardizing based on reference file information.
Genotype Calls	SNP genotype calls (single sample, BRLMM-P-plus algorithm)
Smooth Signal	Gaussian Smoothed Calibrated Copy Number Estimate

NOTE: There is a subset of ~55,000 SNP probes which are used for allelic information analysis but which are not used for Copy Number analysis (on the CytoScan HD Array).

For these SNP probes, LOH and Allele Peaks data will be displayed, but these SNP probes will not have Log2 Ratio, Weighted Log2 Ratio, SmoothSignal, or Copy Number State data displayed, nor will they be used for ascertainment of Mosaicism.

The calculation of Segment data for all the various Segment types takes this into account. All non-polymorphic (copy number) and the vast majority of SNP probes are NOT affected by this change, and will continue to display all graphs and their data points from the CytoScan HD Array CYCHP files.

The Detail View displays the following kind of data for Genome-Wide SNP Array 6.0 Array data (CNCHP):

Table 7.2 Data for Genome-Wide SNP Array 6.0 (CNCHP)

Data Types	Definition
Detected Segment Types	
Gain	Amplifications or duplication
Loss	Hemizygous or homozygous deletions
LOH	Loss of Heterozygosity
Probe array data (displayed as graph data)	
Copy Number State	HMM-derived integer Copy Number State
Log2 Ratio	Per marker Log2 Ratio of normalized intensity with respect to a reference, with further correction for sample specific variation.
LOH	Loss of Heterozygosity
Allele Difference	Difference of A signal and B signal, each standardized with respect to their median values in the reference.
Smooth Signal	Smoothed Calibrated Copy Number Estimate

Table 7.3 Data for OncoScan FFPE Assay Array (OSCHP)

Data Types	Definition
Detected Segment Types	
Gain	Amplifications or duplication
Loss	Hemizygous or homozygous deletions
LOH	Loss of Heterozygosity
Probe array data (displayed as graph data)	
Copy Number State	TuScan derived Copy Number State.
Log2 Ratio	Per marker Log2 Ratio of normalized intensity with respect to a reference, with further correction for sample specific variation.
Weighted Log2 Ratio	Contains the Log2 Ratios processed through a Bayes wavelet shrinkage estimator. These processed values are input to the CNState algorithm HMM.
LOH	Loss of Heterozygosity.
Allele Difference	Difference of A signal and B signal, each standardized with respect to their median values in the reference.
Smooth Signal	Smoothed Calibrated Copy Number Estimate.
Somatic Mutations	Location and detection of Somatic Mutation.
B-allele Frequency	Number of B alleles/ number of A+B alleles, used to show allelic imbalances.

See *Changing Graph Appearance* (on page 125) for more information about controlling the display of graph data.

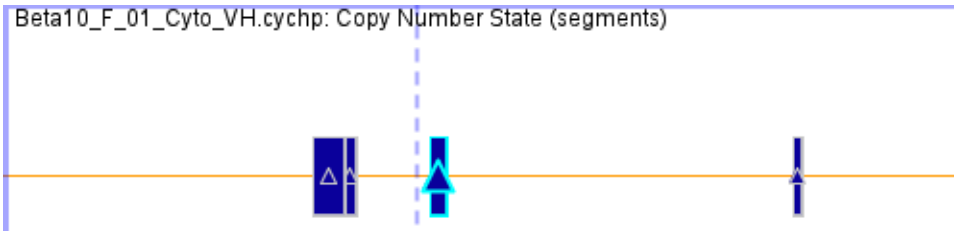
In addition, the Detail View displays:

- Regions: Features in the various region files loaded into ChAS, including CytoRegions and Overlap Map items.
- Annotations: Indicate the known or suspected locations of features, such as mRNAs, exons, FISH/BAC clones, structural variants, and so forth.

  You can expand or contract the annotations. See *Expanding and Contracting Annotations* (on page 124).
- Chromosome info, with:
  - Coordinate scale
  - Marker position information
  - Chromosome number
  - Cytoband information

Selected segments are displayed with enlarged icons; selected regions or annotations are outlined and highlighted (Figure 7.18).

Figure 7.18 Selected Segment



Detail View Navigation Controls (Figure 7.19)

Figure 7.19 Detail View navigation controls

The Detail View has the following navigation controls:

Detail View's Available Navigation Controls

Chromosome Coordinates scale	Shows the position along the genome.
Zoom Slider	Controls the horizontal zoom and the area of the chromosome displayed.
3x Zoom In	Press Ctrl + Plus to view up to 3 preset Zoom In settings.
3x Zoom Out	Press Ctrl + Minus to view up to 3 preset Zoom Out settings.
Zoom in Using Click and Drag	Place your mouse cursor over a point of interest. Press Shift while holding down the left mouse button. Drag the mouse cursor to frame/zoom in on your point of interest.
Stretch Slider	Controls the vertical stretch of the Display area.
Scroll bars	Used to select the area displayed after zooming or stretching the vertical or horizontal scale.
Position Indicator	Dashed vertical blue line. Click in the view to set the position of the indicator The position that is highlighted in the graphs table. The position that is used as the center point when zooming.

You can use the Stretch Slider and Vertical Scroll Bar to zoom in on a section of the Detail. You can also use the mouse wheel as shown below:

- Alt + mouse wheel stretches the display
- Ctrl + mouse wheel zooms in on the horizontal scale
- Mouse wheel scrolls up and down

Selecting a Chromosome Section for Display

Data from the same sample files is displayed in all 3 views, at different scales.

You can select a particular chromosome, or a section of the chromosome, for detailed study using:

- *Karyoview and Selected Chromosome View*
- *Coordinate Range Box* (on page 110)
- *Zoom to a Selected Item* (on page 111)
- *Detail View Navigation Controls (Figure 7.19)* (on page 108)

You can also double-click on an item in a table to zoom to the region of the chromosome where that item is located.

Karyoview and Selected Chromosome View

To select a chromosome for detailed examination:

- Click a chromosome in the Karyoview.

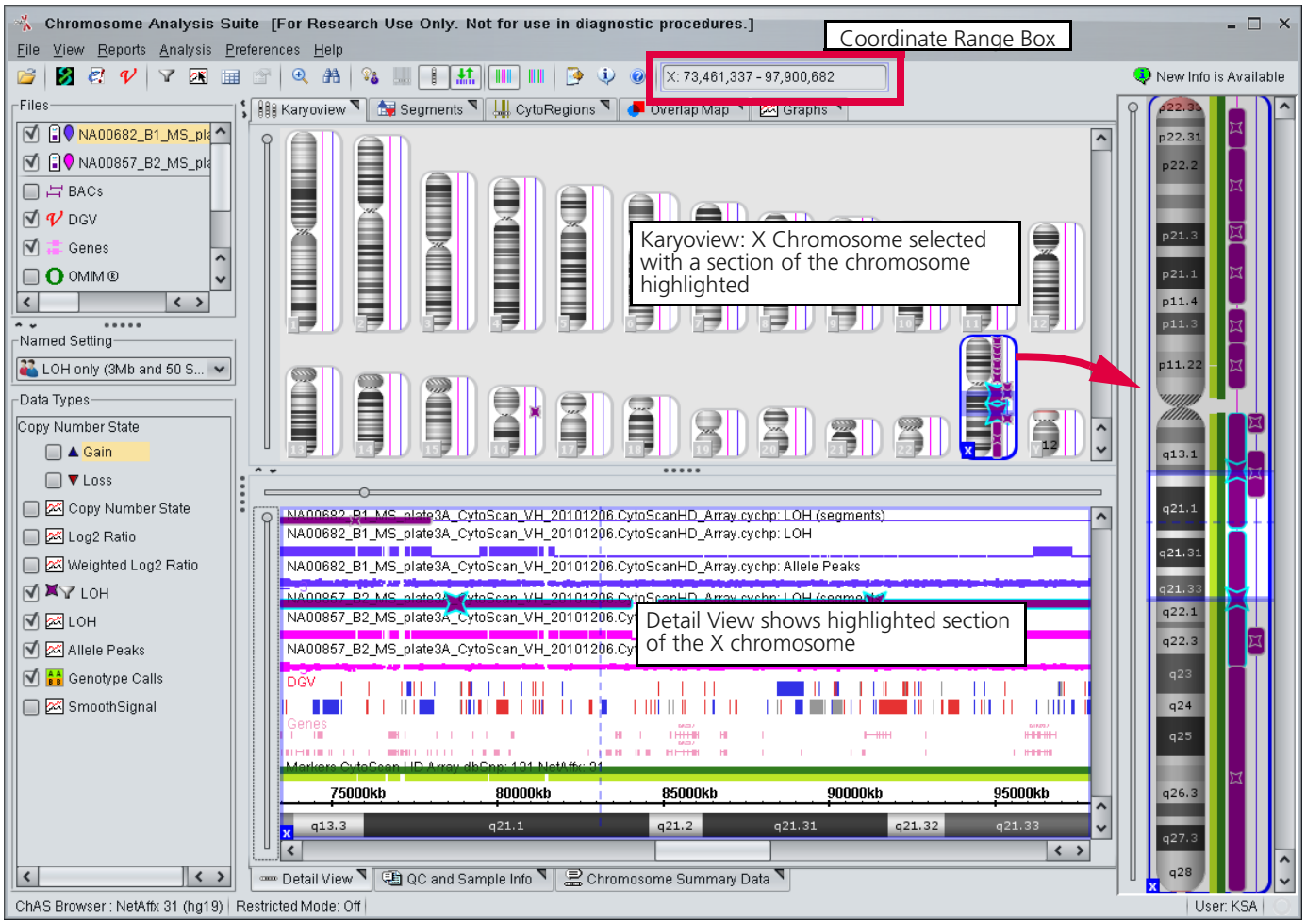
  The chromosome is displayed in the Selected Chromosome View and the Detail view.

To examine a section of the chromosome:

- Click and drag on the section in the Karyoview or the Selected Chromosome View.

  The selected section is displayed in the Detail View (Figure 7.20).

Figure 7.20 Areas displayed in Karyoview, Selected Chromosome View, and Detail View



Selected Chromosome View shows the X chromosome and the highlighted section

Coordinate Range Box

The Coordinate Range box (Figure 7.21) is located in the ChAS main tool bar. It shows the selected chromosome and the start and stop positions displayed in the Detail View. You can enter coordinates in the box to update the Detail View.

Figure 7.21 Coordinate Range Box in Main tool bar

Coordinate Range Box

To go to a specific coordinate or coordinate range:

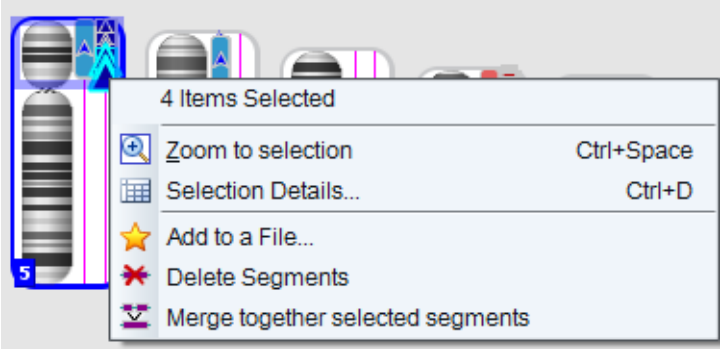
- Enter the desired location in any one of these formats then press the <Enter> key:
  - "**chromosome number : start - end**": sets the view to the given start and end coordinates on the given chromosome.
  - "**start : end**" or "**start - end**": sets the view to the given start and end coordinates of the current chromosome.

Zoom to a Selected Item

There are several ways to zoom in on a feature.

- Click a segment in the Segment table, CytoRegions table, or Overlap Map. Clicking on items auto-zooms to your configured zoom buffer (15% of item size is default), To edit the zoom buffer, see *Zoom Buffer* (on page 113).
- Double-click a feature in the **Details View**, **Karyoview** or **Selected Chromosome View**.
- Press **Ctrl+Space** to zoom *all the way in* to the selected item's start and end coordinates, with no buffer.
- In the Karyoview, Selected Chromosome View, or Details View, you can use the **Zoom to selection** option in the feature right-click menu (Figure 7.22) to go to the start and stop coordinates of the selected feature.

Figure 7.22 Right-click menu options



Edit Configurations in the Miscellaneous (Misc) Tab

The Misc tab contains:

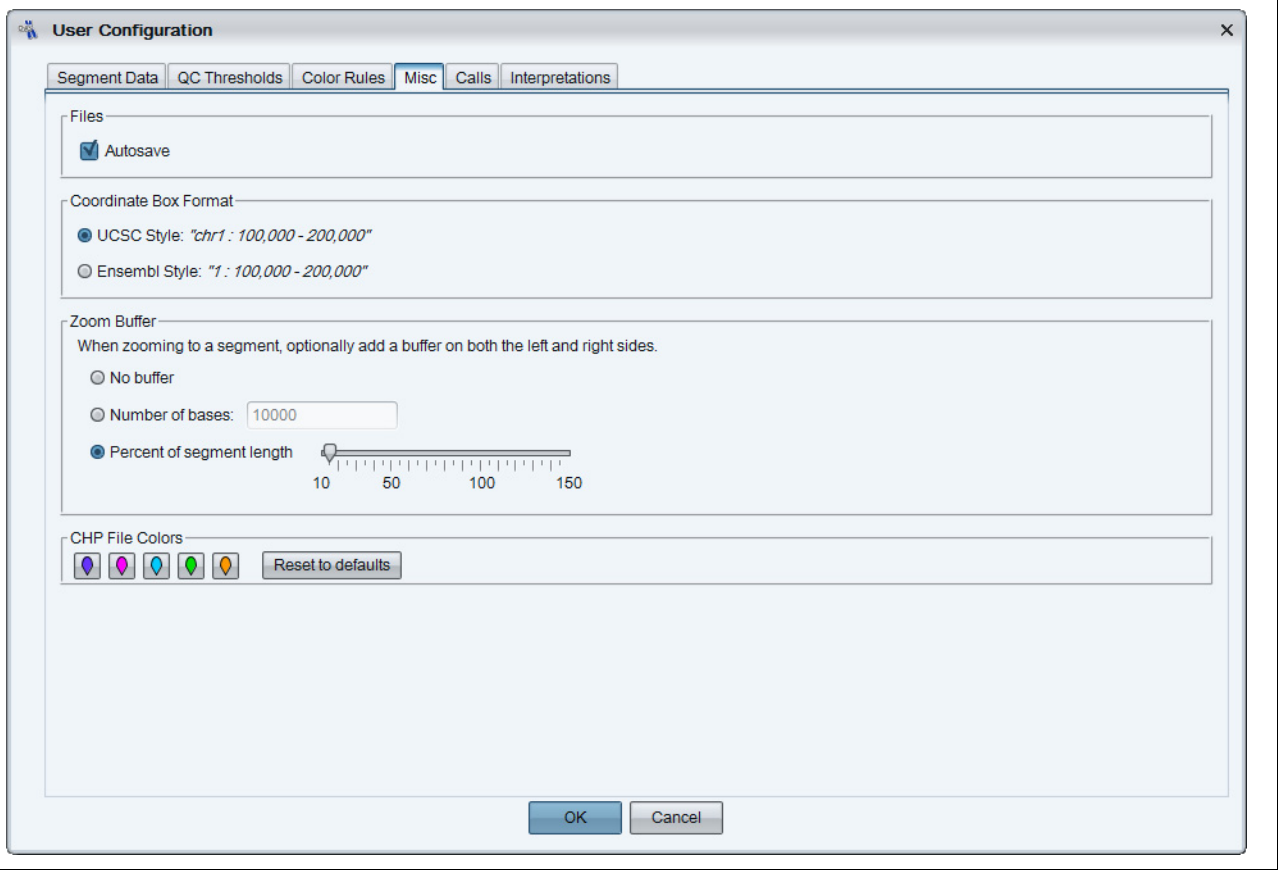
- Autosave
- Coordinate Box Format
- Zoom Buffer
- CHP File Colors

To access the Misc tab:

1. Click **Preferences** → **Edit User Configurations** or click on the upper tool bar.
   The User Configuration window appears.

2. Click the **Misc** tab. (Figure 7.23)

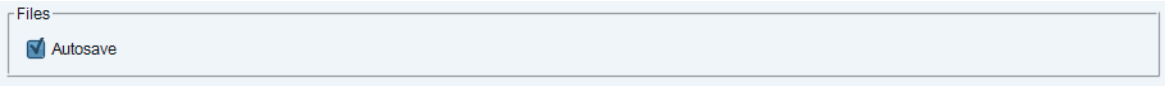
Figure 7.23 User Configuration window - Interpretation tab window



Autosave

Click to check the **Autosave** checkbox to automatically save your files as they are edited. Uncheck to disable autosave. (Figure 7.24).

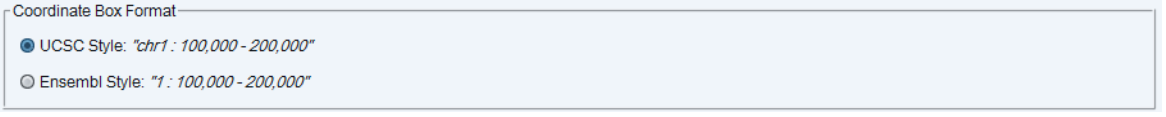
Figure 7.24 Misc tab window - Autosave



Coordinated Box Format

Click the appropriate radio button to choose the format of your displayed data. (Figure 7.25)

Figure 7.25 Misc tab window - Coordinate Box Format



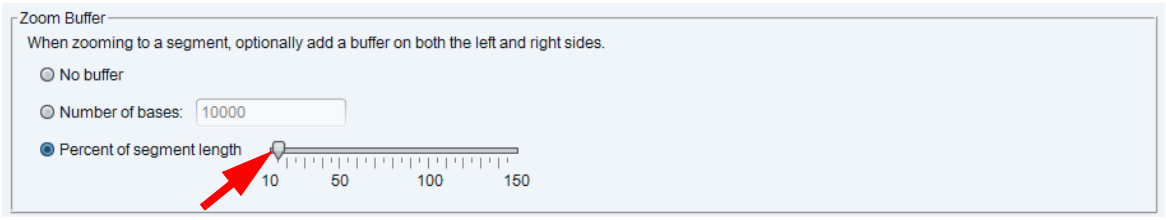
Zoom Buffer

By default, the zoom percentage is set to 15% in new user profiles.

The Zoom Buffer feature offers 3 settings:

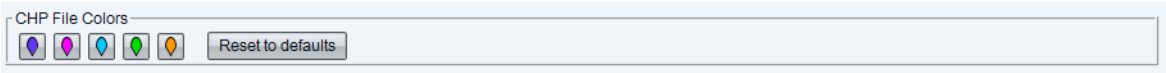
- **No Buffer**: Click this radio button to turn off the Zoom Buffer feature.
- **Number of bases**: Click this radio button, then manually enter the number of bases you want.
- **Percent of segment length**: Click, then drag the slider bar (Figure 7.26) to the zoom percentage you want.

Figure 7.26 Misc tab window - Zoom Buffer



CHP File Colors

Figure 7.27 Misc tab window - CHP File Colors



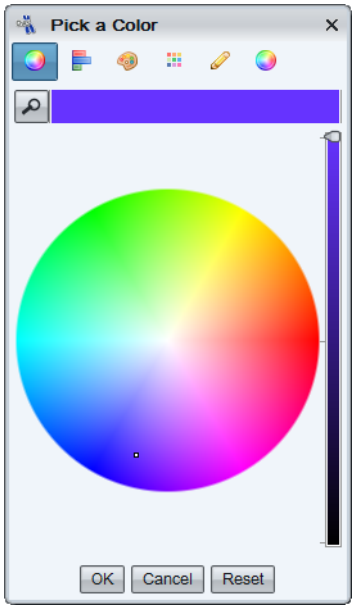
There are 5 preset CHP file colors assigned to your CHP data (Figure 7.27), but each default color, can be changed to a different color.

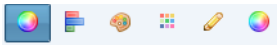
To change a CHP file color:

1. Click on the colored icon you want to change.

A color wheel appears. (Figure 7.28)

Figure 7.28 Misc tab window - CHP File Colors - Color Wheel



2. Use the color wheel to locate the specific color you want or click on one of the several coloring options  (including a website-safe color pallet).
3. Click **OK**.
4. Repeat steps 1-3 to change additional default colors.

   At anytime, click **Reset to Default** to return the 5 CHP file colors back to their default colors. (Figure 7.29)

Figure 7.29 Misc tab window - CHP File Colors

To return a single CHP file color back to its default color, click on the CHP file color, then click the Color Wheel's **Reset** button.

5. Click the User Configuration window's **OK** button to save your changes and exit.

NOTE: Samples that are currently loaded while a color change is made, may not reflect your new color scheme. To remedy this, close, then re-open ChAS 2.1 to guarantee your new color choices are reflected throughout all your samples.

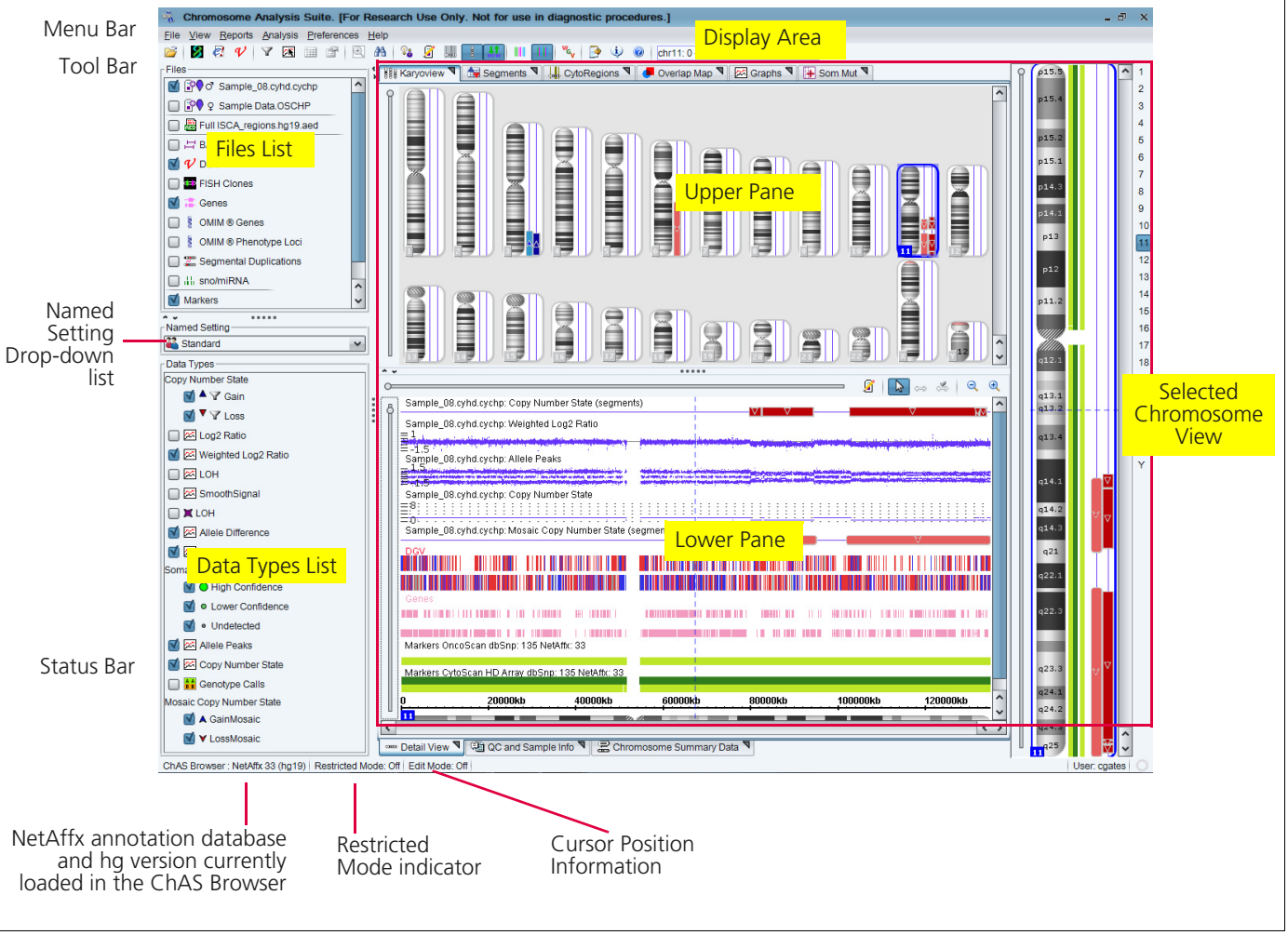
Controlling the Display of Data

Chromosome Analysis Suite (Figure 7.30) provides controls for:

- *Selecting Data for Display* (on page 116)
- *Selecting Data Types for Display* (on page 118)
- *Changing the Grouping of Samples and Data Types* (on page 119)
- *Selecting Light or Dark Schemes for Display* (on page 124)
- *Expanding and Contracting Annotations* (on page 124)
- *Changing Graph Appearance* (on page 125)

Later chapters explain other options for filtering data, how to specify certain regions for extra attention or ignoring, and how to create Region files with region information and annotations.

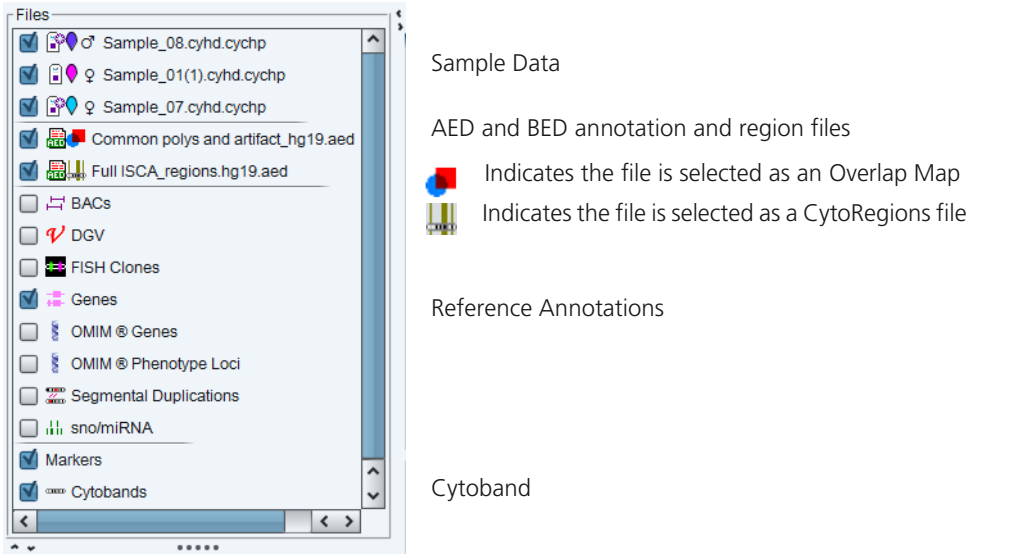
Figure 7.30 ChAS with data loaded



Selecting Data for Display

The Files list (Figure 7.31) allows you to select sample data, region files, and reference annotations for display in the graphic view.

Figure 7.31 Files list



The Files list displays the files grouped by:

- Sample Data
  - Colored nibs display the color used for the data lanes for that sample in the Karyoview, Selected Chromosome View, and Detail View.
  - The appropriate gender symbol ♂ ♀ is displayed to the right of the colored nib. GenomeWide SNP 6.0 CNCHP file data may also contain a "**?**" if the gender was determined to be "unknown".
  - If a loaded file has a QC parameter that is out of range, an alert symbol appears next to the file name.
- Region Data Files

  Icons indicate the file type (AED or BED) and whether the loaded files have been selected as a CytoRegions file or Overlap Map file.
- Reference Annotations: Loaded during software installation and startup. Only displayed in Detail View.
- Cytobands: Separated from other reference annotations because they cannot be moved in the displays, and because they are also displayed in the Karyoview and Selected Chromosome View.

To select and deselect files for display:

- Click in the checkbox next to the file name.

The order in the Files list determines the order of display of the lanes in the Karyoview, Chromosome view, and Detail View.

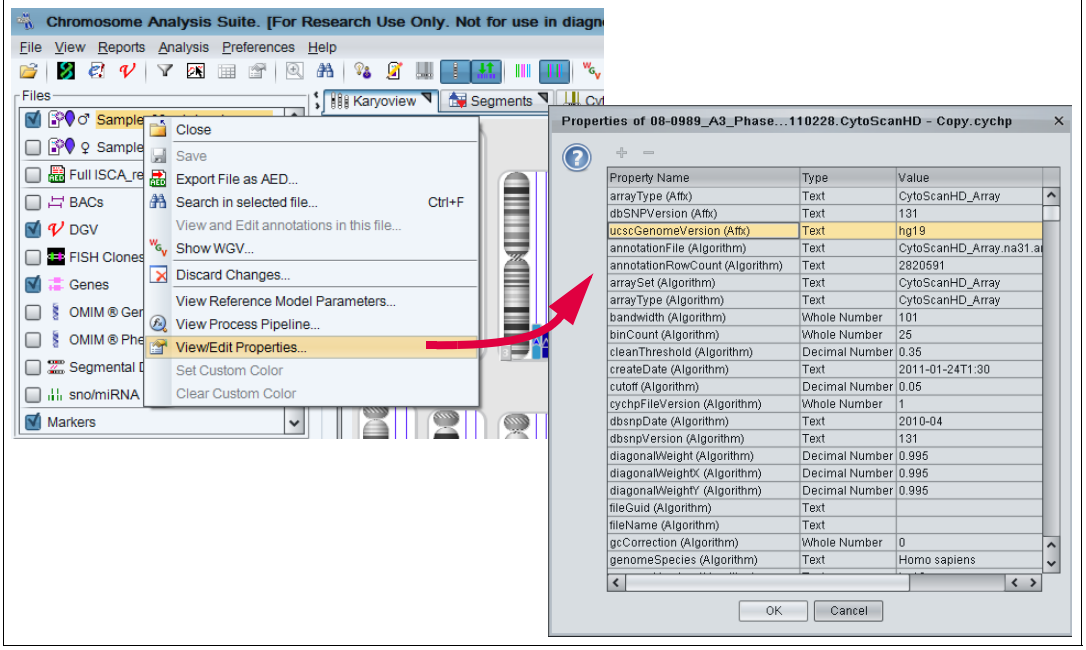
To change the order of the Sample lanes or reference annotations:

- In the Files list, click a file name and drag it to a new position.

To view data properties:

1. Right-click a file and select **View/Edit** Properties on the shortcut menu (Figure 7.32). The Properties dialog box appears.

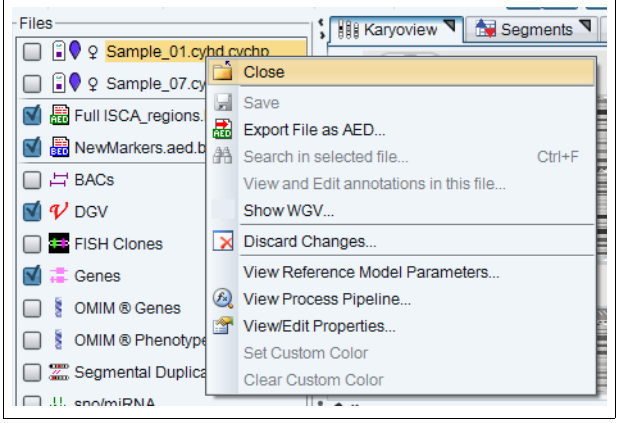
Figure 7.32 View data properties



To close a file:

1. Right-click on the file you want to close.
2. Select **Close** from the menu (Figure 7.33).

Figure 7.33 Closing a file



The file is removed from the Files list and the data is no longer displayed.

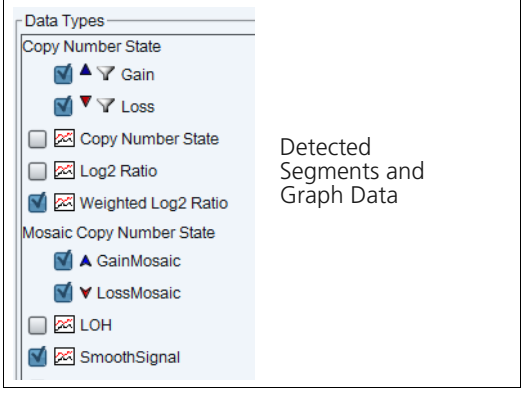
NOTE: If you change the smoothing or joining parameters, the new rules are applied to the original, raw segments of CxCHP files which have not had modifications, Calls or Interpretations made.

CxCHP files which have had segments modified or had Calls or Interpretations made will have had their smoothing and joining parameters fixed and will not adopt changes made to the array type's Segment Data smoothing and joining settings.

Selecting Data Types for Display

The Data Types list (Figure 7.34) shows the data types that can be displayed in ChAS.

Figure 7.34 Data Types list



It displays a list of data that can be displayed in the Karyoview, Selected Chromosome View, Detail View, and tables. The available data types may vary, depending upon the type of sample data available.

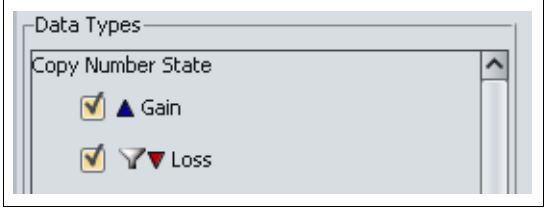
It allows you to select from Segments data and CN/LOH discrete graph data.

The Segments data is displayed in:

- Karyoview
- Selected Chromosome View
- Detail View

If segment parameter filters have been applied to a segment type, a funnel symbol appears next to the segment type name (Figure 7.35).

Figure 7.35 Funnel symbol for Loss segments



The Graph data is displayed only in the Detail View.

To select and deselect data types for display:

- Click in the checkbox next to the Data Type name.

To change the order of the data types:

- In the Data Types list, click a file name and drag it to a new position.

  The selections made here can be saved with a Named Setting (see *Named Settings* (on page 289)).

To turn the symbols used for segments on or off:

- From the View menu, select or deselect Segment Symbols.

Changing the Grouping of Samples and Data Types

A unique color is assigned to each sample and used for the lanes in the Karyoview, Selected Chromosome View, and Detail View.

Each segment type is assigned to its own lane and has its own symbol.

You can use the following options:

- *Lanes Grouped by Sample* (on page 119)
- *Tracks Grouped by Data Type* (on page 122)

This allows you to do different types of comparisons between samples and segment types.

To change the grouping:

- From the View menu, select Group by Sample or Group by Type; or

  In the tool bar, click the Group by Sample or Group by Type button.

Lanes Grouped by Sample

When the lanes are grouped by sample, the different segment types for each sample are kept together in the Karyoview and Selected Chromosome View (Figure 7.36) and in the Detail View (Figure 7.37).

Figure 7.36 Lanes grouped by sample in Karyoview

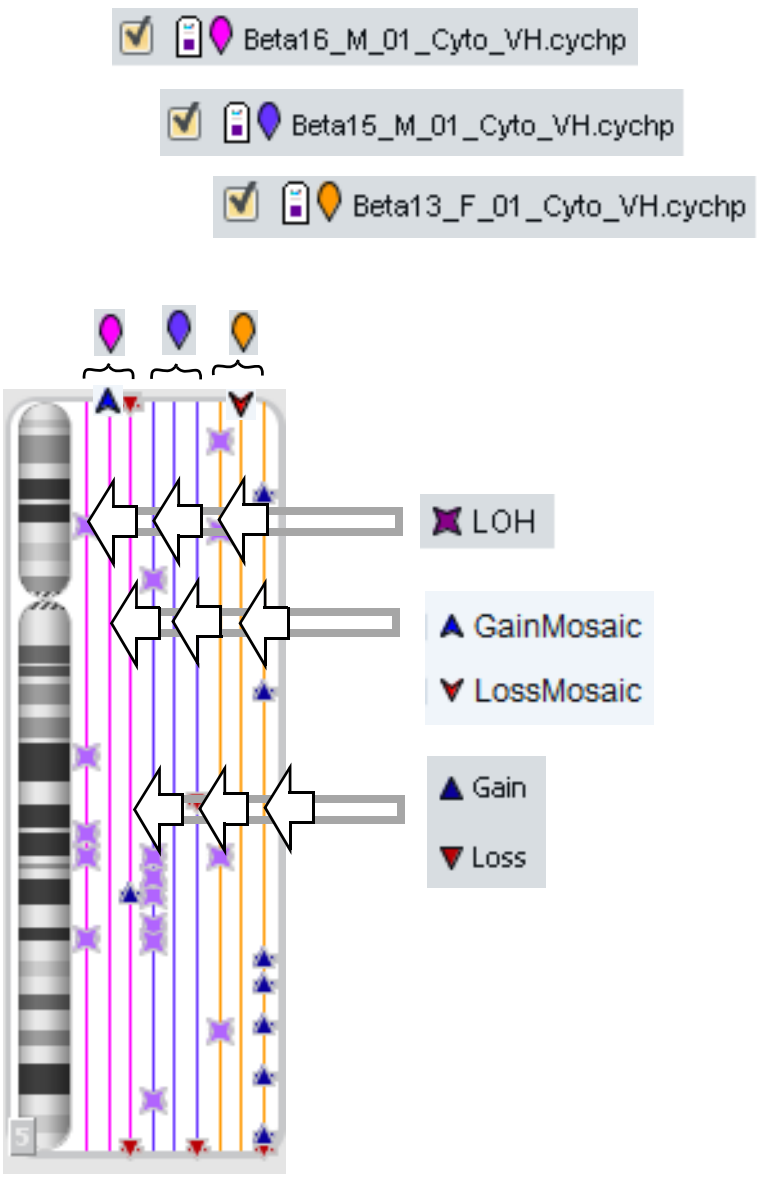


Figure 7.37 Lanes grouped by Sample in Detail view

Tracks Grouped by Data Type

When the lanes are grouped by data type, the lanes for different samples are kept together for each segment or graph type in the Karyoview and Selected Chromosome View (Figure 7.38) and in the Detail View (Figure 7.39).

Figure 7.38 Lanes grouped by data type in Karyoview

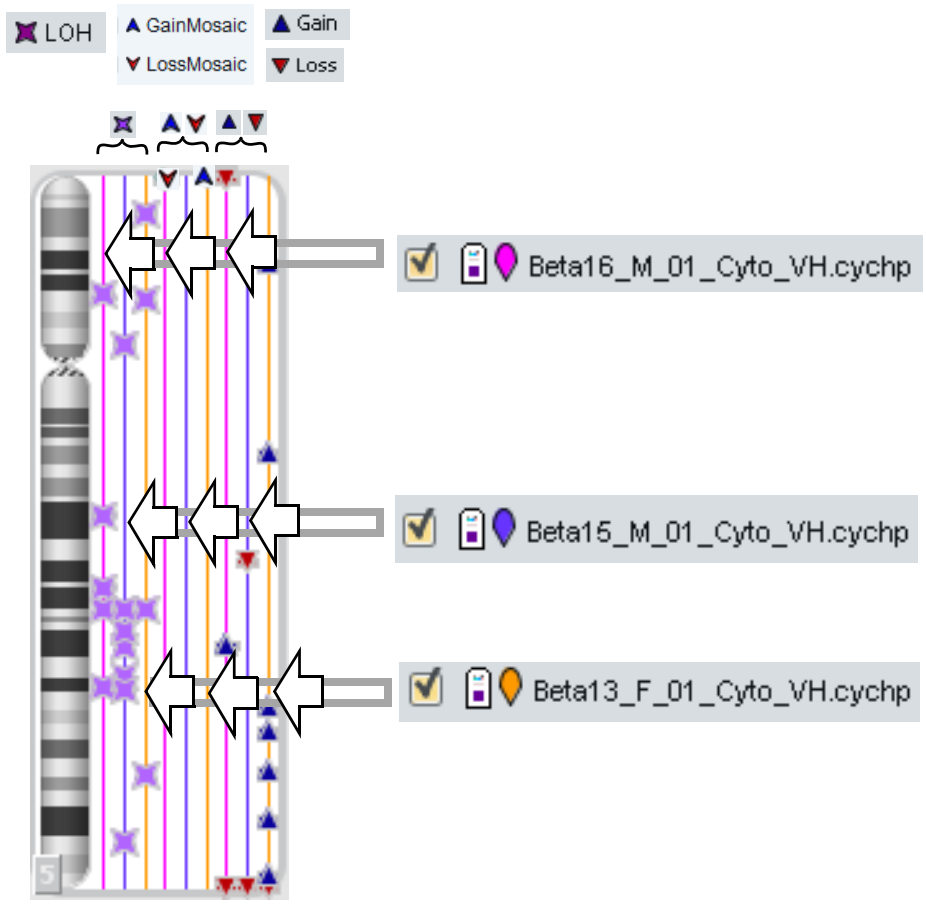
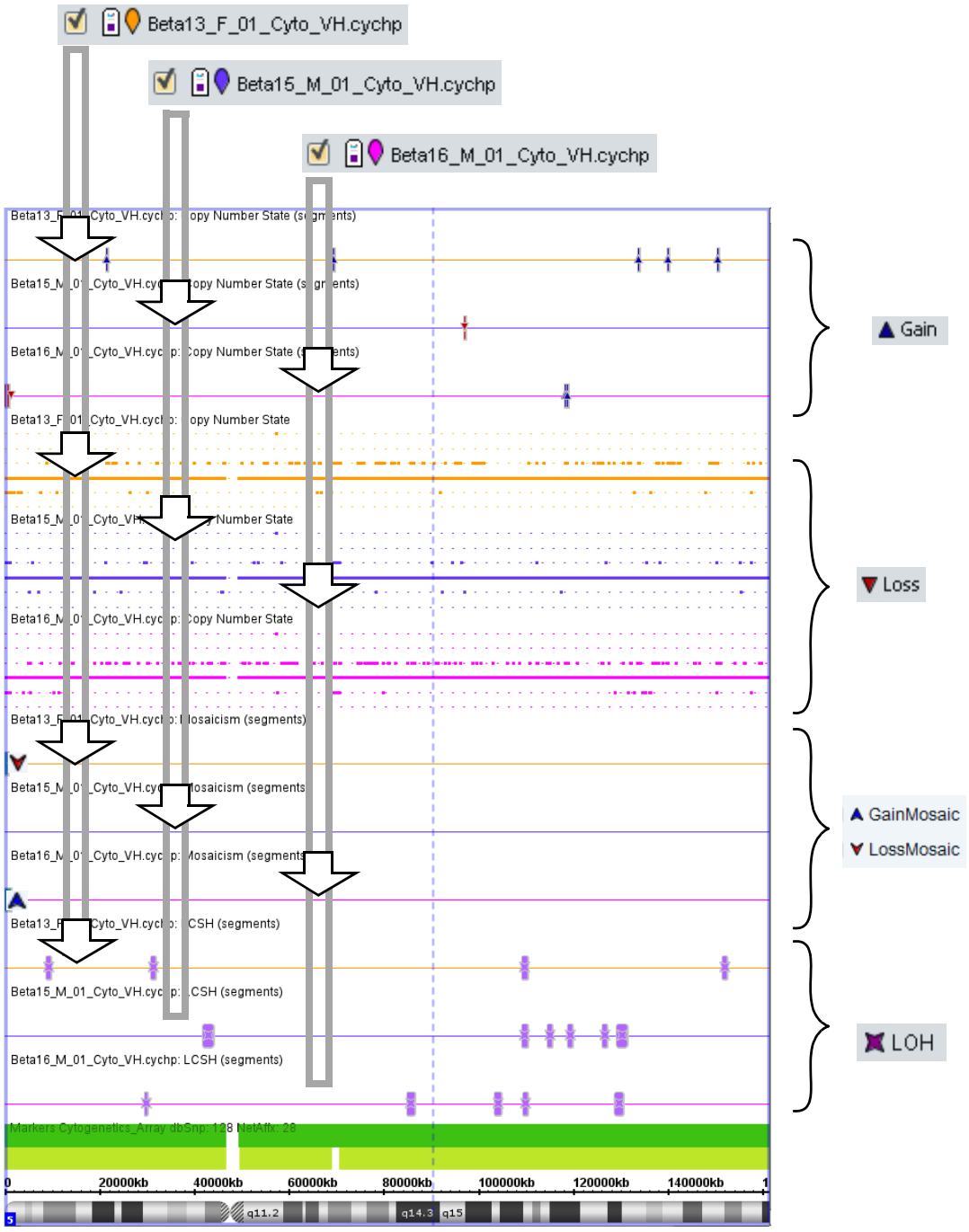


Figure 7.39 Lanes grouped by Data Type in Detail View



You can change the order of samples and Data types in the views by clicking and dragging in the Files and Data Types list.

Selecting Light or Dark Schemes for Display

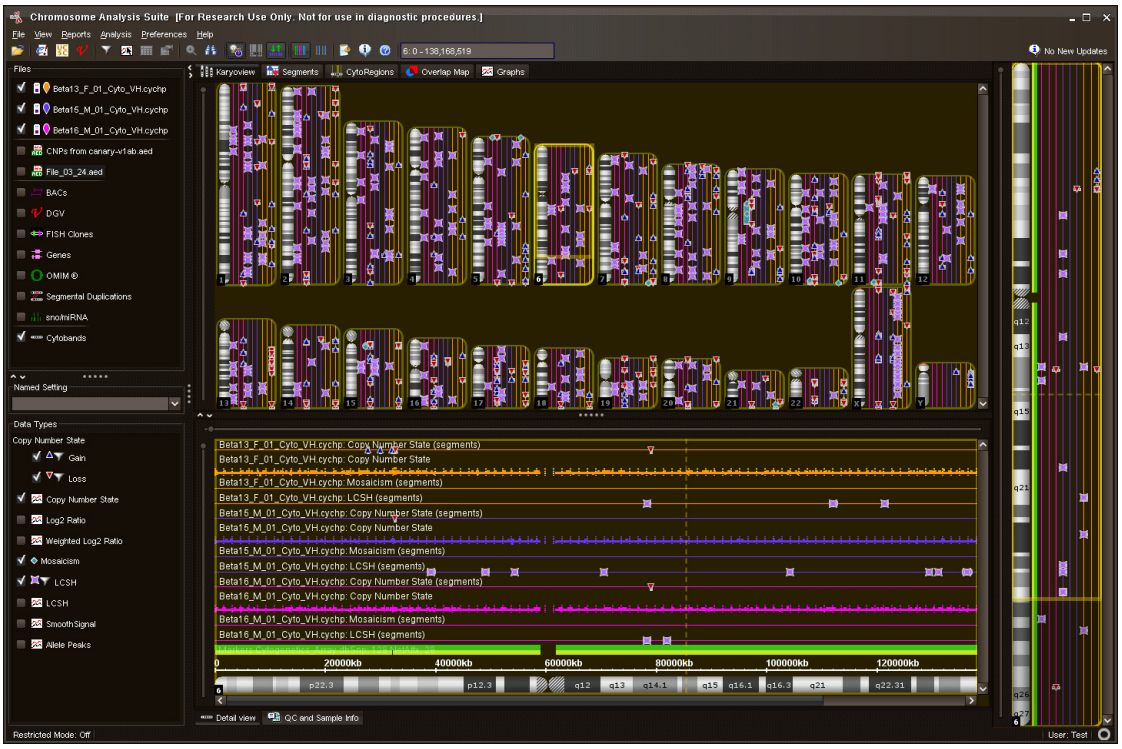
You can choose to display the graphics using a light background or a dark background (called Dark Scheme).

To select Dark Scheme for display:

- From the View menu, select **Use Dark Scheme**; or
- Click the **Dark Scheme** button in the Main Tool Bar.

  The Dark/Light scheme is used (Figure 7.38).

Figure 7.40 ChAS with Dark Scheme selected



Expanding and Contracting Annotations

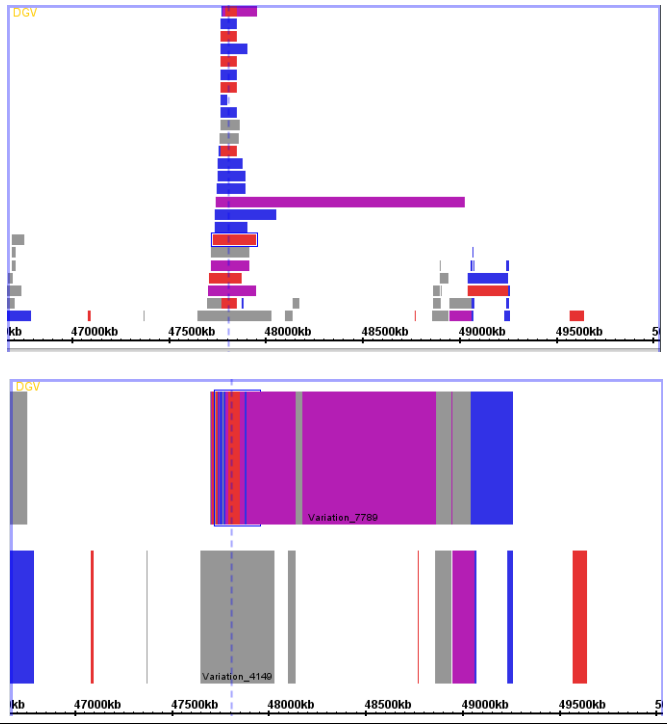
For tracks containing multiple rows of annotations, collapsing tracks consolidates all rows within a track into two rows. Any annotations after the first one will be placed in the second row (Figure 7.41).

When there are multiple annotations of one type at the same coordinate, the separate annotations will be shown on separate rows.

Collapsing tracks is useful if you don't need to see all the details. However, be aware that in collapsed tracks larger annotations may obscure smaller ones; annotations with introns may be obscured by annotations that don't show the intron.

NOTE: The maximum number of tracks that can be displayed for any reference annotation is 25.

Figure 7.41 DGV annotations, collapsed (top) and expanded (bottom)



To toggle between collapsed and expanded display of annotations:

- From the **View** menu, select **Expand/Collapse Annotations**

  or

- Click the **Expand/Collapse** button on the main tool bar.

Changing Graph Appearance

You can modify many properties of the graphs in the Detail View. ChAS provides options for:

- *Selecting Different Graph Styles* (on page 127)
- *Changing Graph Attributes* (on page 130)
- *Changing Scale* (on page 131)

Settings and adjustments that are specific for graphs can be made using the **Graph Settings** dialog box.

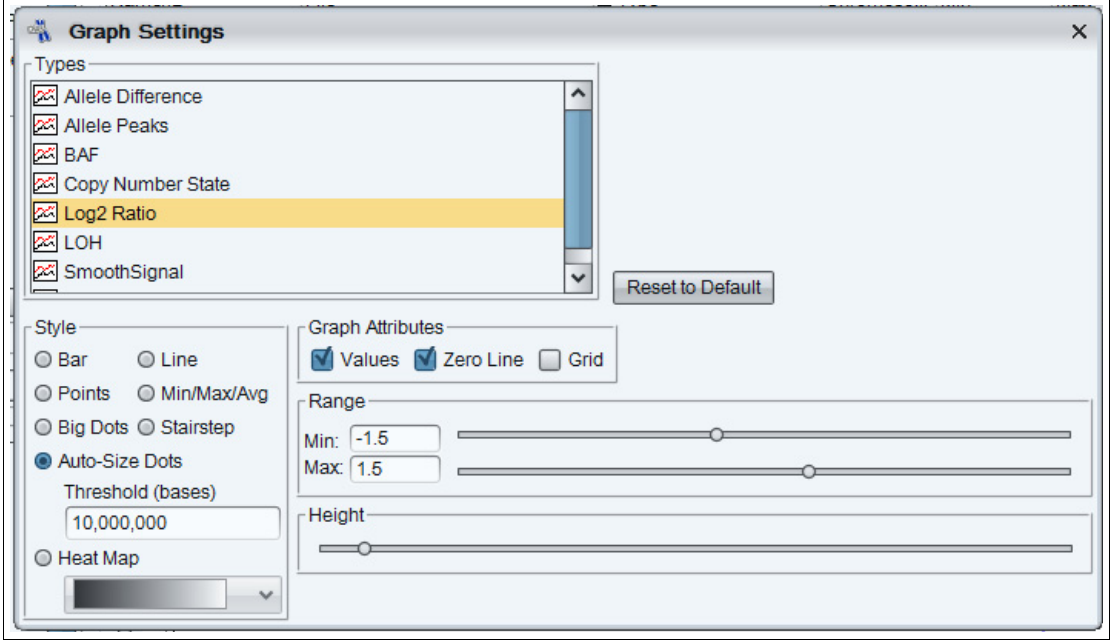
To open the Graph Settings dialog box:

- From the **View** Menu, select **Graph Settings**; or

  Click on the **Graph Settings** button in the Graphs Tab tool bar.

  The Graph Settings dialog box opens (Figure 7.42).

Figure 7.42 Graph Settings dialog box



The Types box displays the graph data types being displayed in the Detail View.

To change the settings for a graph type:

1. Click on a data type to change the settings for that type.
2. Make changes to the graph settings by typing in new values or by operating sliders in the Graph Adjuster panel. For details, see:
   - *Selecting Different Graph Styles* (on page 127)
   - *Changing Graph Attributes* (on page 130)
   - *Changing Scale* (on page 131)

   Any changes you make to the values in the Graph Settings dialog box will apply to all currently selected graph types.

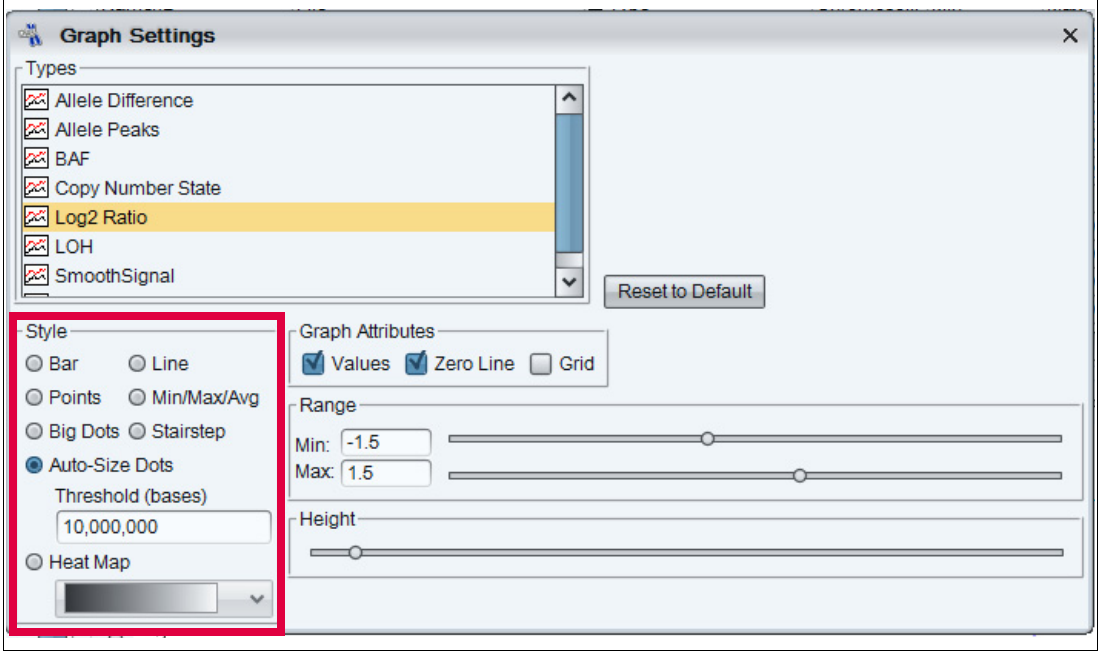
Selecting Different Graph Styles

Graphs can be shown in various representational styles. The type of graph that is most appropriate depends on the type of question being asked about the data. For example, when comparing trends and patterns, it is very useful to use the line graph display method. The user is encouraged to experiment with the different display types to find out which method works best for specific purposes and at specific zoom in magnifications.

To change the graph style:

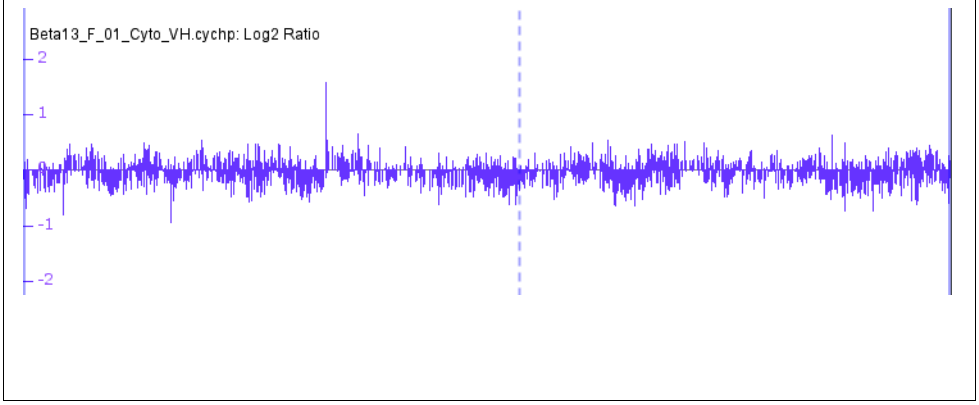
- In the **Style** section choose one of the options (Figure 7.43).

Figure 7.43 Graph Settings dialog box



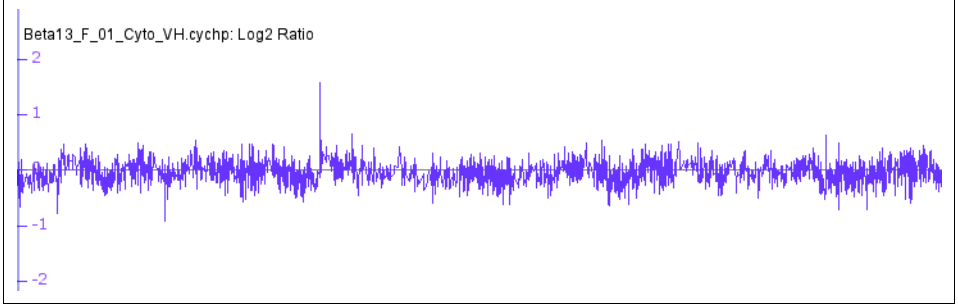
- **Bar** – Individual values are shown as vertical bars that are one base wide for position graphs (Figure 7.44).

Figure 7.44 Bar



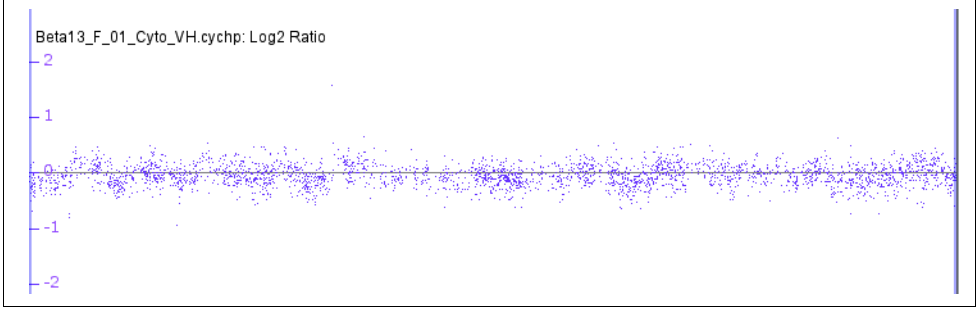
- **Line** – Subsequent values are linked with a line. Even if the input file was not sorted, the values will be connected in order along the genomic coordinate axis (Figure 7.45).

Figure 7.45 Line



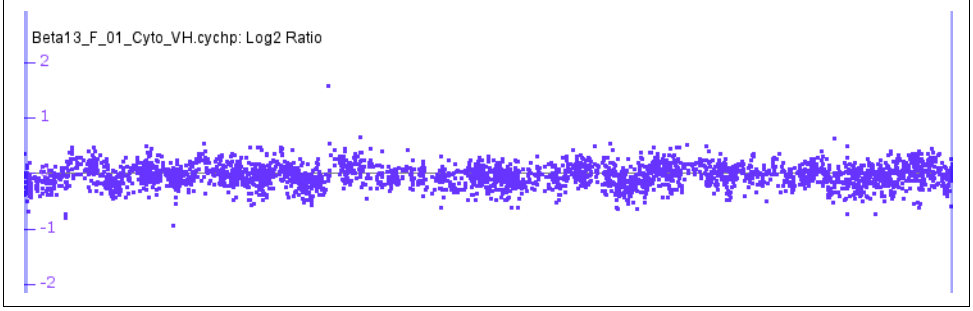
- **Points** – Shows a single dot for each data value (Figure 7.46).

Figure 7.46 Points



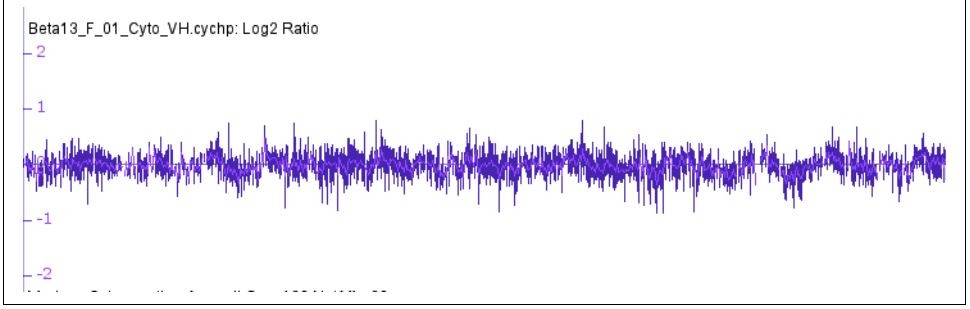
- **Big dots** – Shows a single big dot for each data value (Figure 7.47).

Figure 7.47 Big Dots



- **Min/Max/Avg** – This style is especially useful for showing very densely populated graphs with data points for large numbers of positions (Figure 7.48).

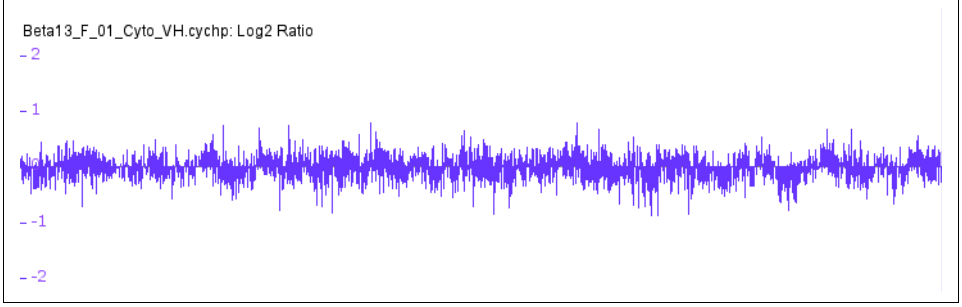
Figure 7.48 Min/Max Average



When Detail View is zoomed all the way in, the display is equivalent to the Line style. When zooming out, ChAS starts to summarize values. When the scale of the display reaches the point where individual x-values are associated with multiple score values, ChAS picks the maximum and minimum values and draws a vertical bar between them. In addition, ChAS draws lines through the average of all the data points represented at each x value.

- **Stairstep** – Similar to the bar graph style, except that bar widths along the horizontal axis are stair-stepped (Figure 7.49).

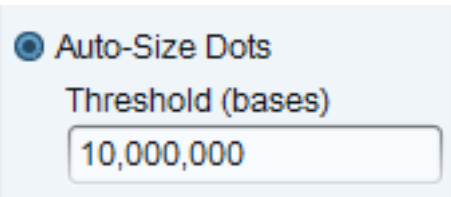
Figure 7.49 Stairstep



For example, if position 100 has a value of 50 and position 200 has a value of 75 and there are no values in between, then ChAS will draw a bar of height 50 that starts at position 100 and stops at position 200. Then, at position 200, ChAS will draw a new bar of height 75 that terminates at the next location with a value.

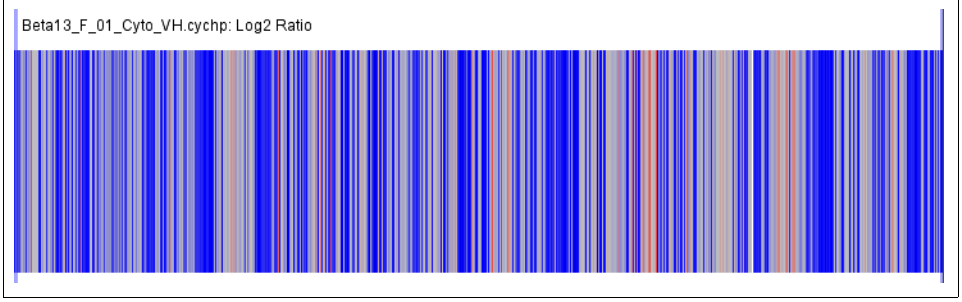
- **Auto-Size Dots** - Transition from Points to Big Dots when zooming in the Detail View. You can select the window size (in base pairs) in which the transition occurs. (Figure 7.50)

Figure 7.50 Auto-Size dots



- **Heat map** – Instead of showing relative intensity via the height of the line at each pixel or coordinate as in most other graph styles, a heat map shows expression levels via color or brightness of the line at each pixel or coordinate (Figure 7.51). This graph style is useful if you want areas of unusual values to jump out at you. If a graph does not render or is hard to see, adjust the visible bounds of the graph until features are readily visible. Several heat map color schemes are available to choose from.

Figure 7.51 Heat Map



There are now two Red/Gray/Blue heat maps.

One is designed to look good for copy number data scaled from 0 to 4 (or 1 to 3), with 2 = normal, and the other is designed to look good for copy number data scaled from 0 to 5 with 2 = normal.

The user must be careful to test that the heat map scaling is appropriate for the data.

Changing Graph Attributes

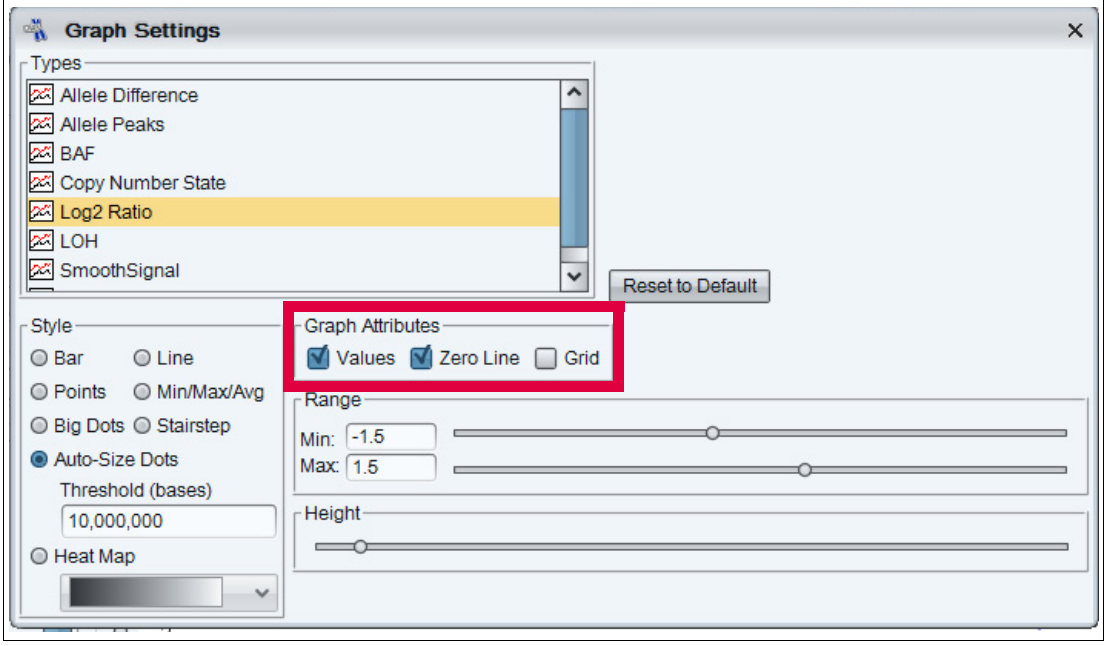
You can display graphs with:

- Value scale
- Zero Line
- Grid

To change the graph attributes:

- Select the attributes you want to display in the Detail View (Figure 7.52).

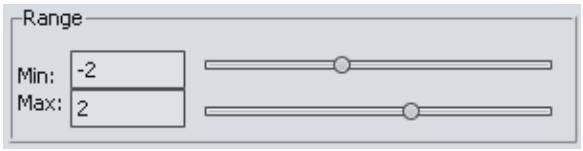
Figure 7.52 Graph Settings dialog box



Changing Scale

Changing the visible bounds involves changing the scale of the graph by setting the maximum and minimum values to be displayed.

To set these visible bounds, use the **Range** section of the **Graph Properties** dialog.



To set specific minimum and maximum values:

- Use the sliders, or type in values to the boxes.

  These values will be applied to each selected graph. You are free to set maximum and minimum values that cover a range smaller or larger than the actual range of your data.

NOTE: The algorithm detects CN state up to 4 for Genome-Wide Human SNP Array 6.0 CNCHP files and CytoScan® HD Array CYCHP files.

The TuScan algorithm detects CN State > 50 for OncoScan FFPE Assay OSCHP files. In the Graphs Settings, the Copy Number State range can be set to what is appropriate to the data.

To change the vertical height of a graph:

- Use the Height Slider to stretch all the graph type in the vertical direction.

The graph height slider is used to increase or decrease the size of a given graph type.

The size is specified in a relative manner. The final graph size will depend on the number of other graphs and annotations being displayed.

Learning More About Features

You can use the following tools to learn more about features in the different views:

- *Pop-ups*
- *Right-click Menu Options* (on page 133)
- *Selection Details Table* (on page 135)
- *Going to External Websites* (on page 140)

Pop-ups

You can mouse over a feature in any of the views to display a popup box with information on the feature. The information provided depends on the type of data that the mouse arrow is on.

Pop-ups are available for:

- Cytobands
- Detected segments
- Graph data
- Marker position indicators
- Reference annotations

NOTE: You should expand the reference annotations before selecting one to avoid selecting multiple annotations. See *Expanding and Contracting Annotations* (on page 124).

- Displayed Region files, including
  - Overlap Regions
  - CytoRegions

The information displayed differs depending upon the type of feature selected, as shown in the samples below.

For all segments, the segment start coordinates are always lower by one bp from the coordinate for the starting probe of the segment as reported in the graphs table while the end coordinate matches the coordinate for the ending probe as reported in the graphs table (see Appendix E, *Genomic Position Coordinates* on page 313).

You can learn more about the terms used in the pop ups in *Selection Details Table* (on page 135).

To turn pop-ups on or off:

- From the **View** menu, select **Mouse-over Pop-ups**.

The information (Figure 7.53, Figure 7.54, Figure 7.55) can include custom properties created by a user (see *Viewing and Editing Annotations* (on page 207) for more information).

Figure 7.53 Pop-up for CN State Graph

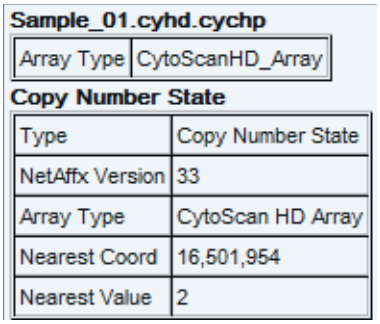


Figure 7.54 Pop-up for a Gain Segment

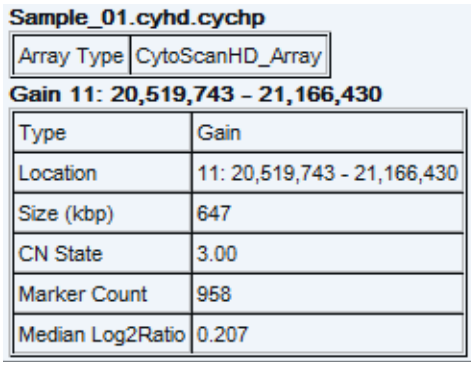
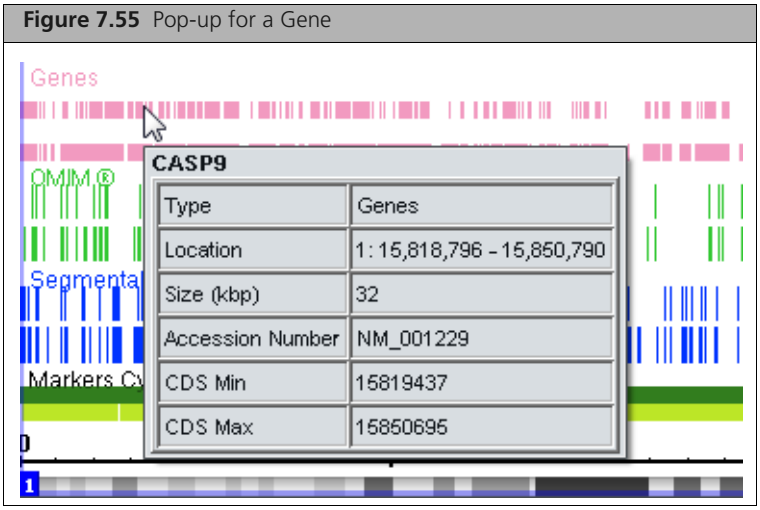


Figure 7.55 Pop-up for a Gene



Right-click Menu Options

You can right-click on any of the following types of features to open a menu with options for learning more about the feature:

- Detected segments (Figure 7.56)
- Reference annotations, including cytobands (Figure 7.57)

NOTE: You should expand the reference annotations before selecting one to avoid selecting multiple annotations. See *Expanding and Contracting Annotations* (on page 124).

- Displayed Region files, including (Figure 7.58):
  - Overlap Regions
  - CytoRegions

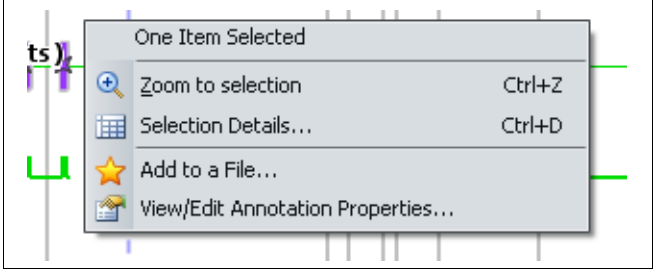
Not all options are available for the different feature types.

NOTE: In the Detail View you can select multiple items of different types. The available options will differ, depending upon the number and types you have selected.

Some menu options are common to the different types of features:

- Number of items selected (if more than one item is selected, the options available may differ, depending upon the type and number of items).
- Zoom to selection: See *Zoom to a Selected Item* (on page 111).
- Selection Details: Opens the Selection Details box with information about the feature. See *Selection Details Table* (on page 135).
- Add to a file: Add the selected segment, annotation, or region to a region file. See *Adding Regions to an Existing AED File* (on page 201).
- View/Edit Annotation Properties: displays the Annotations Properties dialog box for the selected feature. You may or may not be able to edit the properties. See *Viewing and Editing Annotations* (on page 207).

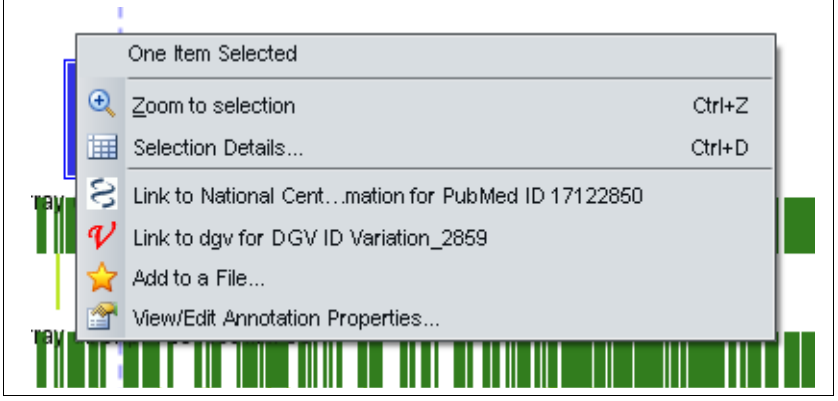
Figure 7.56 Segment right-click menu



Special options for annotations include:

- Link to remote web site for more information. (Figure 7.57)

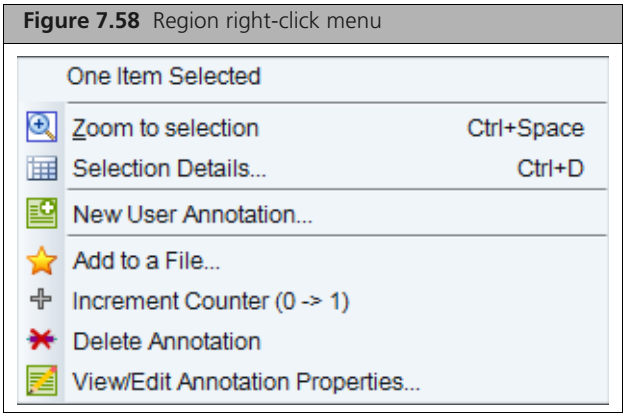
Figure 7.57 Annotation right-click menu



Custom functions for regions include (Figure 7.58):

- New User Annotation: Opens the New User Annotation dialog box, which allows you to add user annotation to a region. For more information, see *New User Annotations* (on page 215).
- Increment counter: Increments the counter in the annotation properties, allowing you to track the number of times a feature has been seen. For more information about the annotation properties, see *Viewing and Editing Annotations* (on page 207).

Figure 7.58 Region right-click menu



Selection Details Table

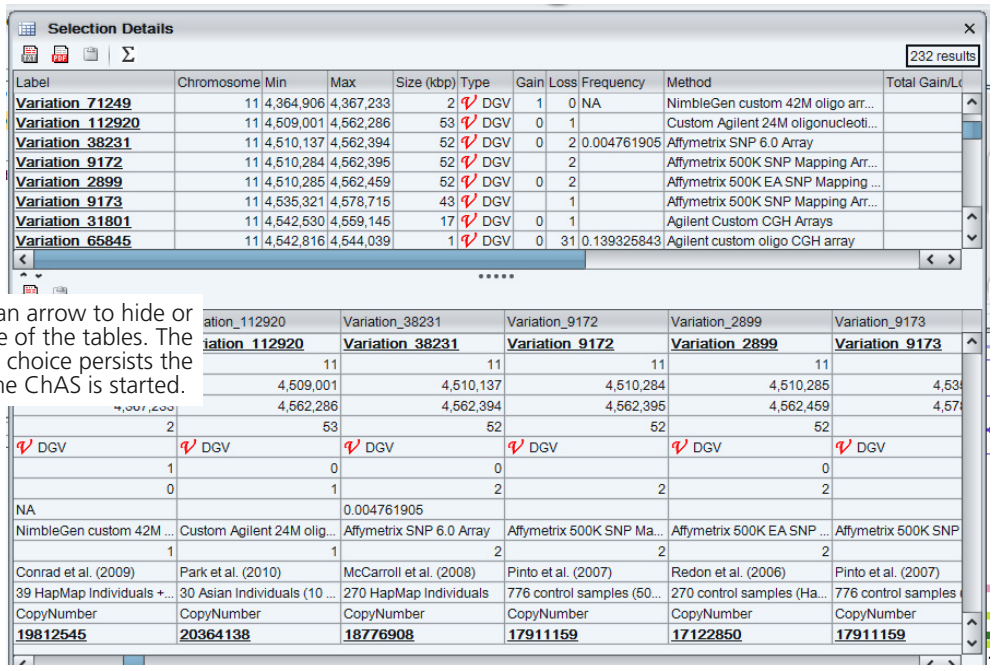
The Selection Details table (Figure 7.59) displays information available for items selected in the graphic display views (Karyoview, Chromosome View, and Detail View). It is accessed by right-clicking on an item in one of the views and selecting **Selection Details**.

The information is presented in two tables:

- Upper table shows one item per row – Provides summation feature and PDF or tab-separated text file export capabilities.
- Lower table shows one item per column – Provides ability to export the table to a tab-separated text file.

To reorder the columns in the upper table, drag a column header left or right. The corresponding row in the lower table is automatically moved to the new location in the table.

Figure 7.59 Example Selection Details table



Upper table:
One item per row

Lower table:
One item per column

Click an arrow to hide or show one of the tables. The display choice persists the next time ChAS is started.

The information displayed in the Selection Details table includes:

Common

Label — Identifier for the item.

Chromosome — Chromosome on which the item is located.

Min — Zero-based index position of the first base pair in the sequence.

Max — Zero-based index position of the last base pair in the sequence, plus one. Adding one ensures that the length of any (hypothetical) segment containing a single marker would be one, and ensures that the coordinates match the coordinate system used in BED files. For all segments, the segment start coordinates are always lower by one bp from the coordinate for the starting probe of the segment as reported in the graphs table while the end coordinate matches the coordinate for the ending probe as reported in the graphs table (see Appendix E, *Genomic Position Coordinates* on page 313).

Size (kbp) — Size of the item.

Type — Type of segment (Gain, Loss, GainMosaic, LossMosaic LOH) or annotation.

Segments

File — File the segment was detected in

CN State — Copy Number State (not displayed for LOH segment types).
The expected Copy Number State on the X chromosome in normal males is not constant over its entire length. This is due to the structure of the sex chromosomes. See *LOH Segments on the X and Y Chromosomes* (on page 47) for more information.

Type — Type of segment, for example, LOH. When sorting by this column, the segments of a particular sample are listed in the same order that they appear in the Data Types window pane.

Size (kbp)	Size of the segment.
Mean Marker Distance	Length of the segment in base pairs divided by the number of markers in the segment.
Interpretation	User-editable field for free-text interpretation on the segment
Call	User-editable field populated by a user-configurable drop list of Calls
Call & Interpretations By	The current computer Operting System login ID and ChAS user profile name at the time that the Call or Interpretation field was last edited.
Call & Interpretation Time	The time and date when the Call or Interpretation field was last edited.
Materially Modified Segment	Indication that segment was previously merged, deleted, or had its start or end boundary, type, or state altered by a ChAS user. (ChAS-based processes of Smoothing and Joining are not "Modifications", nor are making Calls or Interpretations, in this context).
Materially Modified By	The current computer Operating System login ID and ChAS user profile name at the time that the segment was last materially modified.
Materially Modified Time	The time and date when the Segment was last materially modified.
Max % Overlap	The highest percentage by which some item(s) in the Overlap Map overlaps the segment. Segments completely overlapped by an Overlap Map item are 100% overlapped. This number is used for Filtering Segments out by "Overlap".
Overlap Map Items (% of Segment overlapped)	Item(s) in the Overlap Map which overlap the segment, followed by the percentage by which the segment is overlapped by that Item.
CytoRegions	Names of the CytoRegions with which the segment shares coordinates.
Use in Report	Allows manual selection of Segments for export to a Segments Table PDF, rather than all segments in the table.
Marker Count	Number of markers in the segment.
Cytoband Start	Cytoband in which the segment begins.
Cytoband End	Cytoband in which the segment ends.
Genes	List of RefSeq genes from the Genes track that share coordinates with the segment. Identically named gene isoforms are NOT repeated.
Gene Count	A count of the gene names listed in the Genes column
DGV	List of DGV variations that share coordinates with the segment.
FISH Clones	List of FISH clones that share coordinates with the segment.
sno/miRNA	List of sno/miRNA features that share coordinates with the segment.
BACs	List of BACs that share coordinates with the segment.
Old OMIM®	The column formerly known as "OMIM" in the ChAS software when NetAffxGenomicAnnotation Browser files version NA32.1 and earlier are loaded. This column's information is now out of date and has been superceded by the newer OMIM Genes and OMIM Phenotype Loci columns present in NA32.3 and above browser annotation files.
OMIM® Genes	List of OMIM Genes that share coordinates with the segment.
OMIM® Gene Count	A count of the OMIM Gene names listed in the OMIM Genes column.
OMIM® Phenotype Loci	List of OMIM Phenotype Loci that share coordinates with the segment.
Segmental Duplications	List of Segmental Duplications that share coordinates with the segment.

Smoothed/Joined	Indication that segment was created by smoothing or joining two or more segments in the initial segment detection.
Segment Label	A label comprised of the segment's Type, State, and Filename.
Segment Name/ID	File-specific identifier assigned to the detected segment.
Start Marker	The array marker name which marks the beginning of the segment.
End Marker	The array marker name which marks the end of the segment.
Preceding Marker	The array marker just above the segment in the data track used as input for the segment. Note: This column is only applicable to CNState Gain and Loss segments.
Preceding Marker Location	The coordinate location of the array marker just above the segment in the data track used as input for the segment. Note: This column is only applicable to CNState Gain and Loss segments.
Following Marker	The array marker just below the segment in the data track used as input for the segment. Note: This column is only applicable to CNState Gain and Loss segments.
Following Marker Location	The coordinate location of the array marker just below the segment in the data track used as input for the segment. Note: This column is only applicable to CNState Gain and Loss segments.
Mean Log2 Ratio	The mean of all the Log2 Ratio values contained in the segment.
Mean Weighted Log2 Ratio	The mean of all the Weighted Log2 Ratio values contained in the segment.
Microarray Nomenclature	An ISCN-based description of the segment.
Sample UUID	Unique identifier for the CHP file.
Max % Coverage	The highest percentage by which a segment covers some item(s) in the Overlap Map.
Number of Overlap Map Items	Number of Overlap Map items which share genomic coordinates with the segment.
% of Overlaps Map Item covered by Segment	Overlap Map Item and the percentage by which it is covered by the segment.
Full Location	Chromosome Start and Stop in a user-friendly format for use in external databases.

Genes

Accession Number	Unique identifier assigned to the sequence in GenBank.
CDS Min	Minimum position of the coding sequence (BED-style coordinates).
CDS Max	Maximum position of the coding sequence (BED-style coordinates).
Strand	The sequence strand of the item.

FISH Clones

Labs	Lab where clone was produced.
accNames	Accession associated with the clone.
accCount	Number of accessions associated with the clone.
beCount	Number of BAC end sequences associated with this clone.
placeCount	Number of times FISH'd.
bandEnds	End FISH band.
stsCount	Number of STS markers associated with this clone.

stsNames	Names of STS markers.
placeType	How clone was placed on the sequence assembly.
bandStarts	Start FISH band.
beNames	Accession IDs of BAC ends

DGV

Database of Genomic Variants

Frequency	Usually defined by the lab that submitted the variant based on number of samples tested.
PubMed ID	Link to PubMed.
Gain	Number of gains in the sample population.
Loss	Number of losses in the sample population.
Total Gain/Loss	Number of total Gains and Losses observed in the sampled population.
Method	Method used to identify the variant.
Sample	Sample(s) used to identify the variant.
Reference	PubMed ID for paper.
VarType	Variation type.
Strand	The sequence strand of the item.

OMIM® Genes

OMIM Gene Title	The title of the gene associated with the OMIM entry.
OMNI Gene Symbol List	A list of genes associated with the OMIM entry.
OMNI Disorder	Disorder associated with the OMIM entry.
OMNI Phenotype Key	Indicates how this phenotype was placed on the map.
OMNI Gene Symbol	Symbol of the gene based on gene title.

OMIM® Phenotype LOCI

OMIM Phenotype ID	Unique identifier to an OMIM phenotype.
OMNI Phenotype Map Key	Indicates how this phenotype was placed on the map.
OMNI Phenotype Locus Description	Describes the phenotype or disorder associated at the OMIM Phenotype Loci.

Segmental Duplications

Score	Score based on the raw BLAST alignment score. The score for segmental duplications is set to zero in NetAffx annotation 31 and higher.
FracMatch	The fraction of matching bases.
FracMatchIndel	The fraction of matching bases with Indels.

NOTE: Affymetrix does not generate or verify the information for genes, FISH clones, Segmental Duplications, sno/miRNAs, DGV annotations, or OMIM data.

NOTE: Segmental Duplication and sno/miRNA annotations do not have any unique terms; but sno/miRNA annotations use the "type" field to indicate subtypes like "cdBOX" and "HAcaBOX".

Some information may not be displayed, depending upon the feature type. The information can include custom properties created by a user (see *Viewing and Editing Annotations* (on page 207)).

You can export data from the table using the standard table export tools (see *Exporting Table Data* (on page 274)).

You can perform multi-column sorts (see *Sorting by Columns* (on page 229)).

Going to External Websites

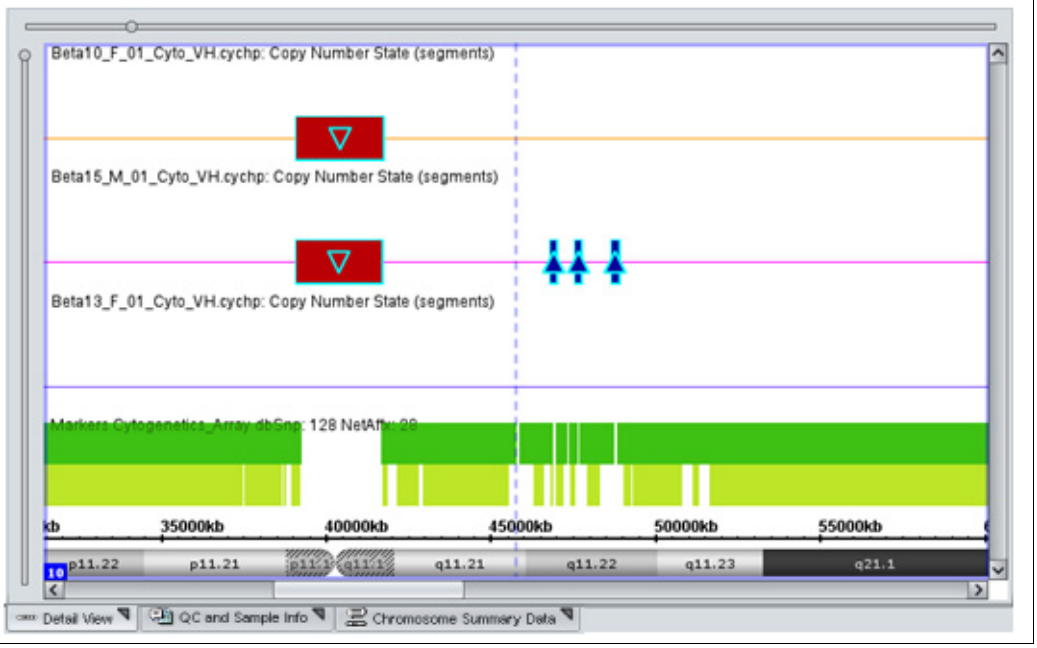
You can view select an area in the Detail View for viewing at one of the following public sites:

- UCSC
- Ensembl
- Toronto DGV

To view the selected area at a public site:

1. In the Detail View, zoom and scroll to the area of interest (Figure 7.60).

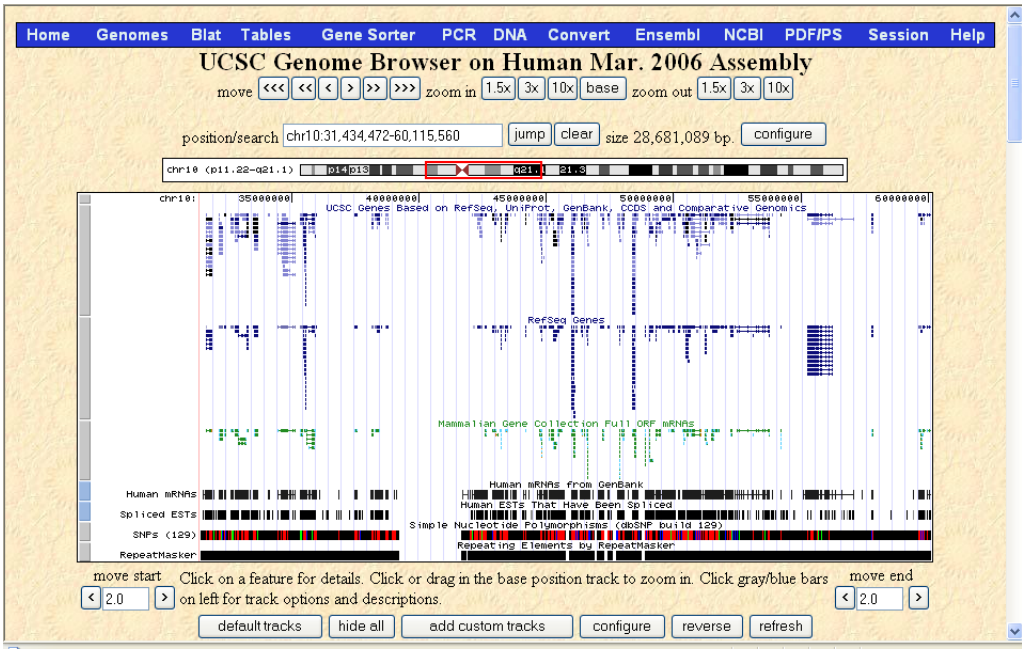
Figure 7.60 Selected area in Detail View



2. From the View menu, select **View Region at [site name]**. Alternatively, click a tool bar button:
   - UCSC
   - Ensembl
   - Toronto DGV

   A browser opens, displaying the selected area of the chromosome (Figure 7.61).

Figure 7.61 Selected area displayed in UCSC Genome Browser



Filtering Segments

ChAS enables you to filter the detected segments using different segment parameters, concealing segments that do not meet requirements for significance for:

- Marker Count
- Length

You can apply these filters to different segment types, using different parameters for each type. The filtering is done on the fly, with changes to the parameters reflected in the different views as they are made.

A segment must pass all filter requirements for the segment type to be displayed.

You can apply different filter values for areas inside CytoRegions and areas outside the CytoRegions (genomewide). *See Using Filters with CytoRegions (on page 185)*.

The Overlap Map filter is described in *Using the Overlap Map* (on page 188).

Filter settings are saved when a Named Setting is created and can be reapplied. *See User Profiles and Named Settings (on page 289)*.

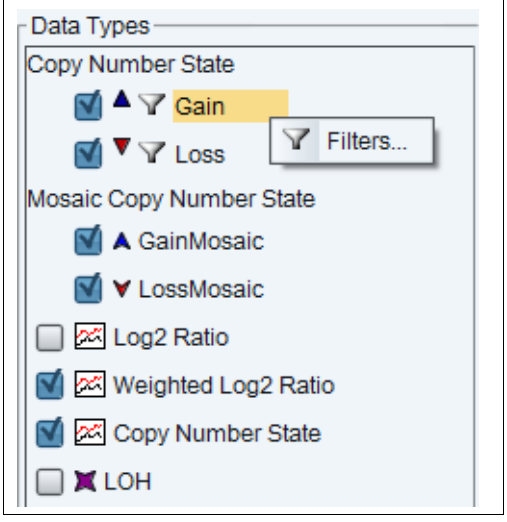
Applying Segment Parameter Filters

To open the Segment Filters dialog box, do one of the following:

- Select **View** → **Segment Filters** on the menu bar
- Click the **Segment Filters** tool bar button
- Right-click on a segment type in the Data Types list and select **Filters…** from the right-click menu (Figure 8.1).

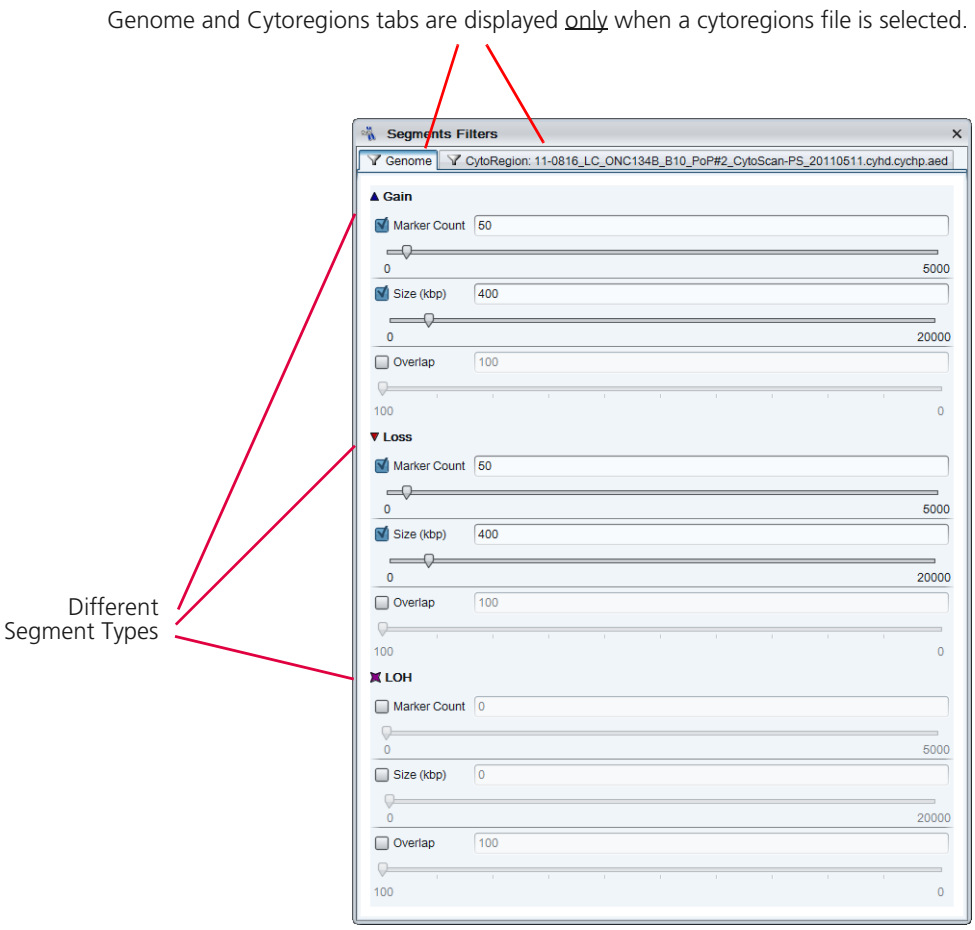
NOTE: If you use the right-click menu option, only the filter settings for the selected segment type are displayed.

Figure 8.1 Right-click menu



The Segments Filters dialog box opens (Figure 8.2 on page 143).

Figure 8.2 Segment Filters dialog box



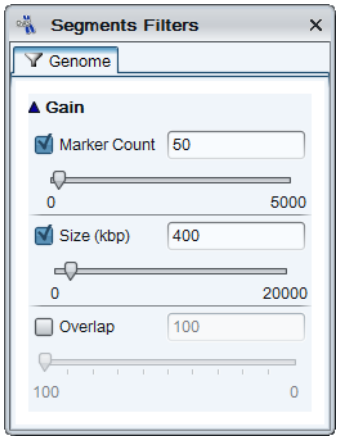
NOTE: The Overlap Map filtering parameter is set using the same dialog box. The Overlap Map function is described in *Using the Overlap Map* (on page 188).

You can filter using the following Segment Parameters:

Marker Count The number of markers the segment encompasses from start to finish. A segment must have at least as many markers as you specify to be displayed.
Each marker represents a probe which represents a sequence along the genome at a particular spot. Markers are probe sequences of DNA, each sized from 12-50 base pairs long, depending on the type of array data. The 12-50 bp sequence is unique to that one spot on the genome it represents.

Size Based on the start and end markers of a segment. Because each segment represents a single place in the genome, you can measure from start to end, in DNA base pairs, and by filtering, demand a segment be at least that long to be visualized.

Figure 8.3 Gain Segment filter settings



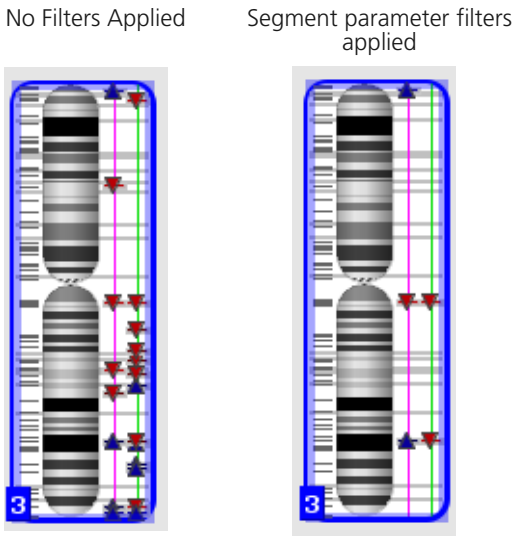
To use the Segment Parameter filters:

1. Open the Segment Filters dialog box (Figure 8.3).
2. Put a check mark next to the parameters that will be used as filters.
3. Set the value for the parameters with the slider or enter a value in the box.

   Moving sliders from left to right progressively removes more and more segments.

   The filtered results are seen instantly in all tables and graphs (Figure 8.4).

Figure 8.4 Filtering results



See *Using the Overlap Map* (on page 188) for information about using the Overlap setting.

See *Using Filters with CytoRegions* (on page 185) for information on using different filtering settings in CytoRegions.

Modification and Interpretation of Segment Data

Editing Segment Data

- ChAS 2.1 allows you to edit data segments; merge, delete, draw *de novo*, adjust boundaries, and change the Copy Number State of CHP data segments, or undo your changes.
- The instant a file is modified by editing of one or more segments, having a segment interpreted, or called, ChAS auto-generates a “sidecar” or CHP Change Archive (CHPCAR) file which saves these modifications. ChAS then uses this CHPCAR file for further user-defined edits and modifications, while your original (native) xxCHP file (CYCHP, CNCHP, or OSCHP) is safeguarded (remains un-touched).
- The smoothing and joining settings for each CnCHP file are locked at the time of CHPCAR file creation, while unmodified CHP files without CHPCAR files will still respond to array type-specific changes in smoothing and joining settings.
- In order for the changes stored in the CHPCAR file to show up in the CHP file data displayed in the ChAS Browser, the CHP file and CHPCAR file must be stored in the same directory/folder (the xxCHP file is still the only file that is required to be “loaded” into the browser, the CHPCAR file's edits, calls and interpretations will load when the CHP file is loaded).
- CHPCAR files are named using the entire CHP file's name, and contain the extension *.chpcar”.
- Changing the name of either file and not the other to match will disrupt the ability for the files to be recognized as associated with each other in the ChAS Browser.
- CHP files which have associated CHPCAR files detected (and in use), display as a special “CHP” icon in the **File** → **Open** dialog window, and in the **Files** *tree* of the Browser.
- Please move the CHP and CHPCAR files together when moving or archiving data.
- Two people using CHAS on different systems should NOT attempt to edit a CHP file at the same time.

Using Edit Mode

> **IMPORTANT:** Make sure the Edit Mode feature is turned on BEFORE you start editing segments.

For your convenience, Edit Mode is accessible in 3 places:

- The Browser's icon row (top).

- The Detail View's icon row (top right)

- Click **View** → **Edit Mode**

By default, Edit Mode is OFF. Click (located on the Browser's top icon row) to turn Edit Mode ON.

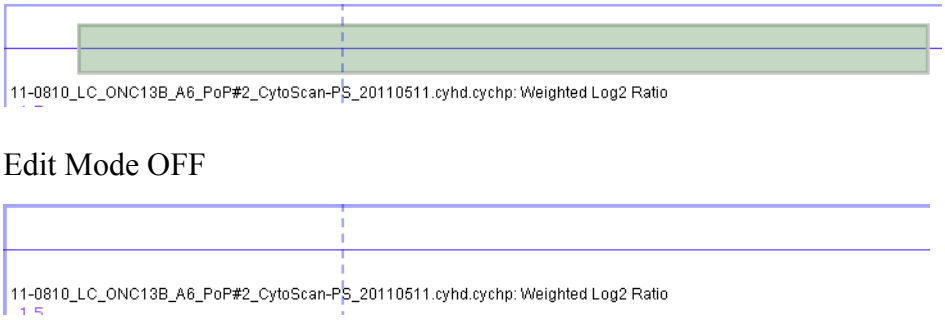
Click to turn Edit Mode OFF and remove all visual indications of your segment changes.

- When Edit Mode is **ON**, deleted segments are visible, and edited segments appear distinct from non-edited segments.
- When Edit Mode is **OFF**, deleted segments are invisible, and edited segments look identical to non-edited segments

.

> **IMPORTANT:** Turn Edit Mode OFF, before generating a PDF report of your data.

Figure 9.1 Edit Mode ON/OFF Deleted Segment example



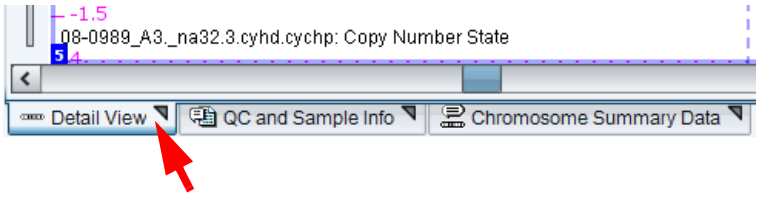
Types of Segment Editing

- *Merging Segments* (on page 147)
- *Deleting Segments* (on page 153)
- *Editing the Start/End Coordinates of a Segment* (on page 155)
- *De Novo Segment Drawing* (on page 157)

- *Changing the Copy Number* (on page 161)

NOTE: Before you start any segment editing, make sure the Detail View tab is selected. (Figure 9.2 on page 147)

Figure 9.2 Detail tab (bottom left)



Merging Segments

There are 2 ways to merge segments:

- *Merging Segment Groups*
- *Segment to Segment Merge* (on page 150)

Merging Segment Groups

NOTE: Merging segment groups together, cancels out any previously assigned Calls. However, un-doing the group of merged segments **(page 152)** reinstates their original Calls.

1. Click **File** → **Open**.

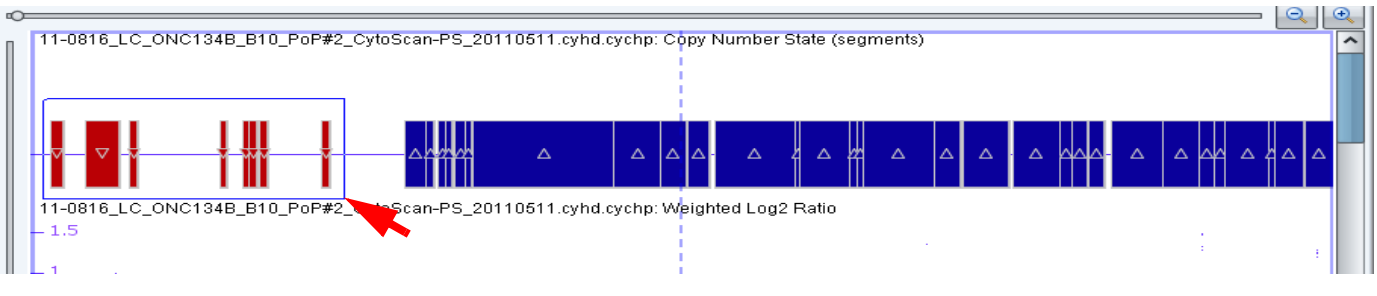
   Your Sample File data folder window appears.

2. Click to select the file you want to edit, then click **Open**.

   The file appears in the ChAS browser's Detail view.

3. Left-click, hold, then move the mouse to frame the segments you want to merge together. (Figure 9.3)

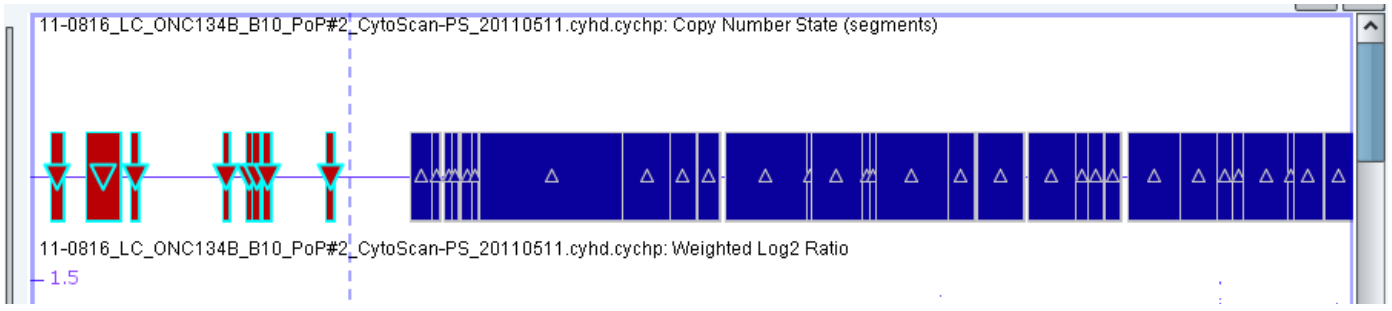
Figure 9.3 Framing your segments



4. Release the mouse button.

Your selected segments, including their *Loss* symbols are now highlighted in **blue**, as shown in the example below. (Figure 9.4)

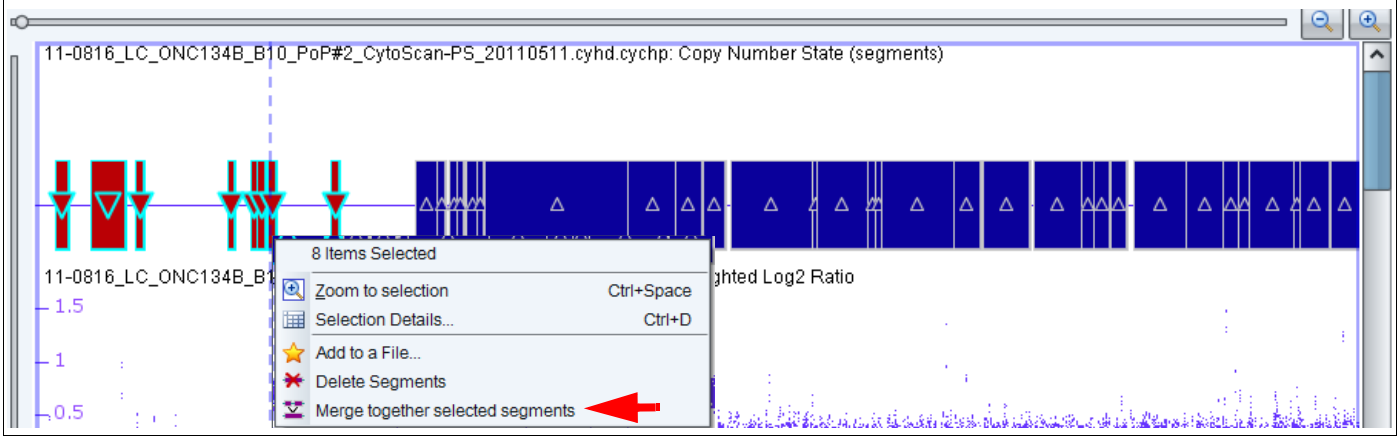
Figure 9.4 Framed segments now highlighted



5. Right-click on a highlighted segment.

   The following menu appears. (Figure 9.5)

Figure 9.5 Right-click menu



6. Click **Merge together selected segments**.

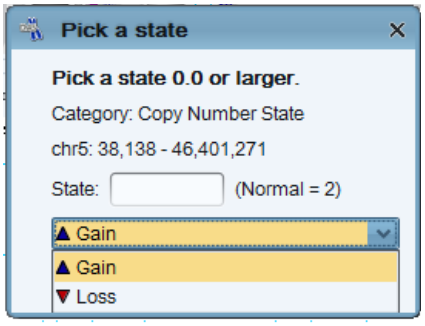
   A message regarding your planned merge may appear. (Figure 9.6)

7. Click **OK** to acknowledge it.

Figure 9.6 Merge All? message

The Pick a state window appears. (Figure 9.7)

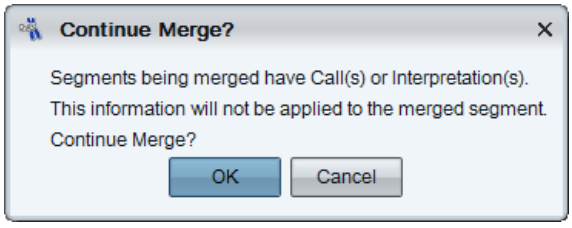
Figure 9.7 Pick a state window



8. In the State field, enter a Copy Number, then use the drop-down menu to select a Type (Gain/Loss) for the new segment. (Figure 9.7)
9. Click **OK**.

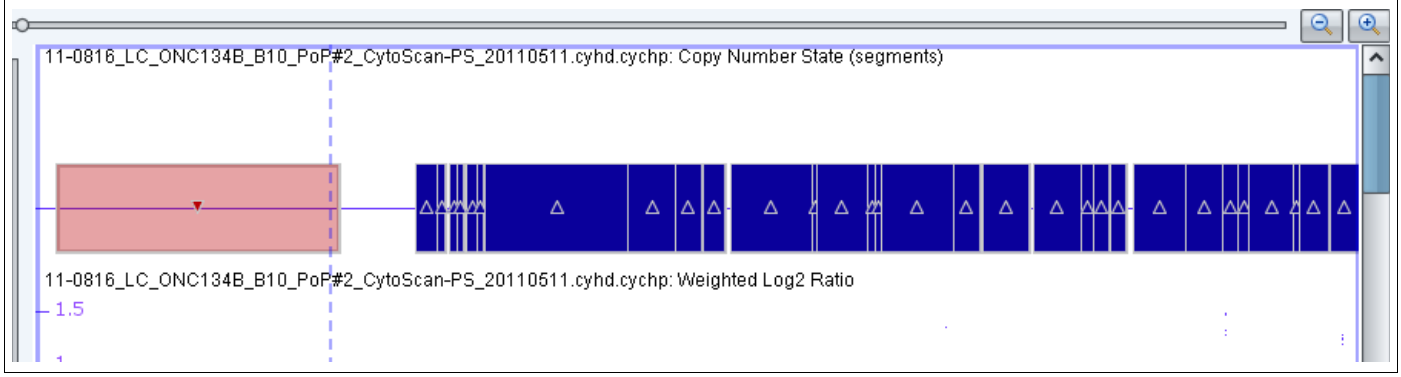
   In cases where the segment to be merged into a group contains a **Call or Interpretation**, the following message appears: (Figure 9.8)

Figure 9.8 Merge Warning message



10. Acknowledge the message, then click **OK**.message appears.
11. Your selected segments are now merged together and appear as shown. (Figure 9.16)

Figure 9.9 Merged segments



Segment to Segment Merge

IMPORTANT: When merging segments during the editing of the start or end of one particular segment, only the segment whose start or end you are editing has the option of having its Call or Interpretation saved (or not).

Segments which are being engulfed by the edit start/end procedure being performed will not have the option of having their Call or Interpretations placed on the resulting segment. However, un-doing the edit resulting in this type of merging (page 152) will reinstate previous Call and Interpretation information for all the original segments involved.

1. Click **File** → **Open**.

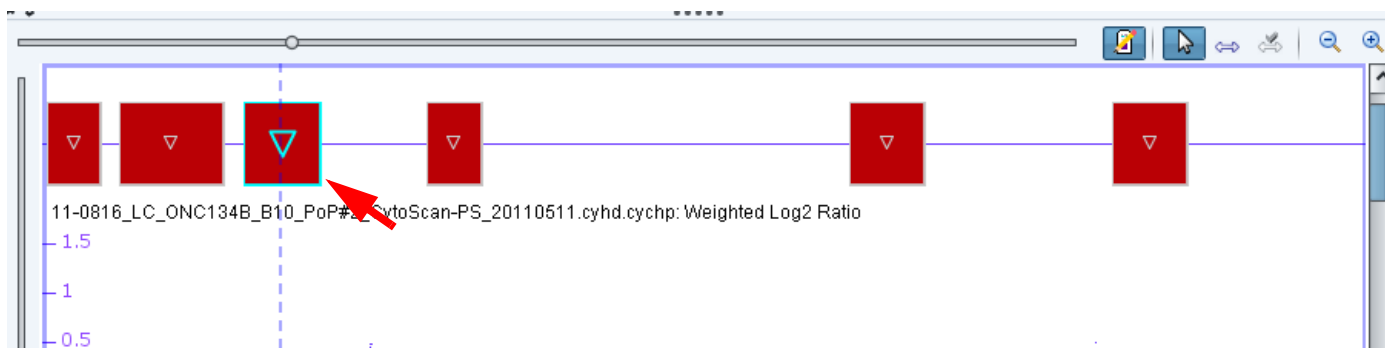
   Your Sample File data folder window appears.

2. Click to select the file you want to edit, then click **Open**.

   The file appears in the ChAS browser's Detail view.

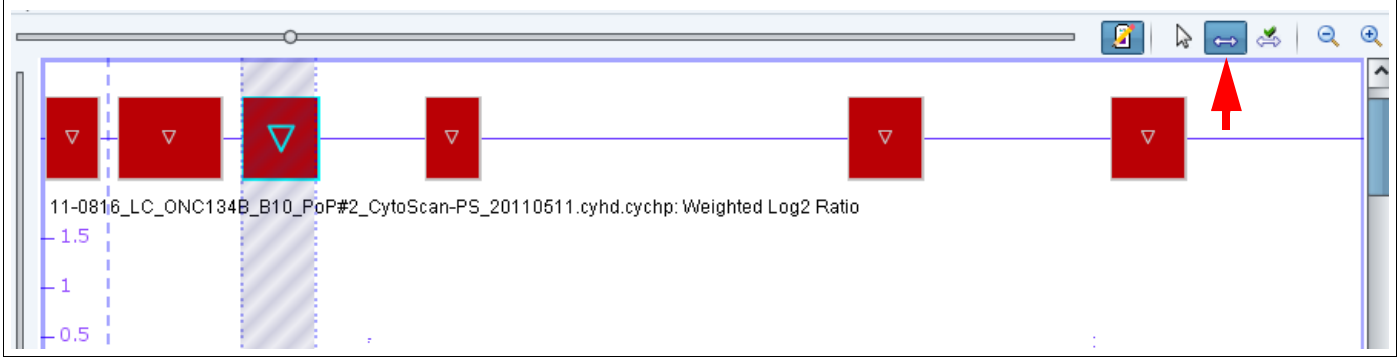
3. Click to select the segment you want to merge with another. (Figure 9.10)

Figure 9.10 Selected segment to merge with another



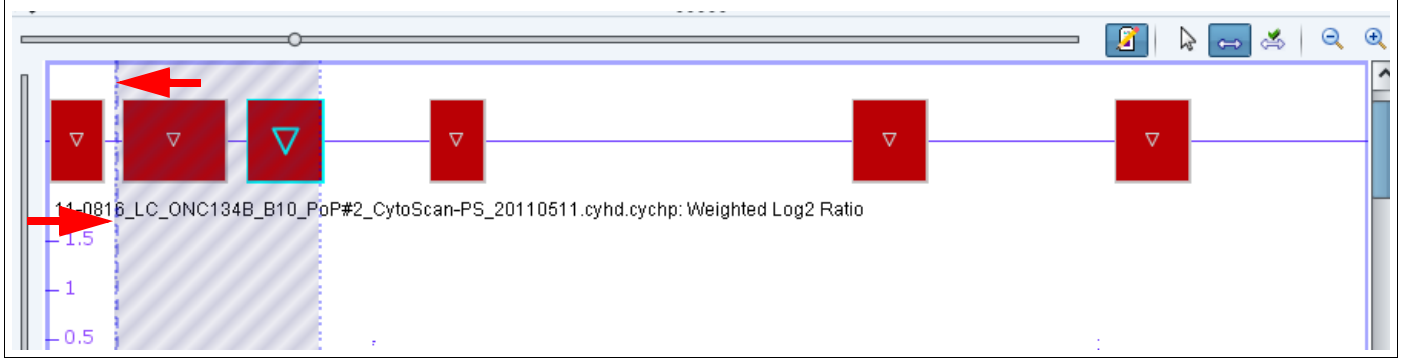
4. Click the Start Editing button.
5. A shaded area over your selected segment appears. (Figure 9.11)

Figure 9.11 Click Start Editing button



6. Click, then drag the shaded area over the segment you want to merge it with. Make sure you overlap your target segment just slightly, as shown. (Figure 9.12)

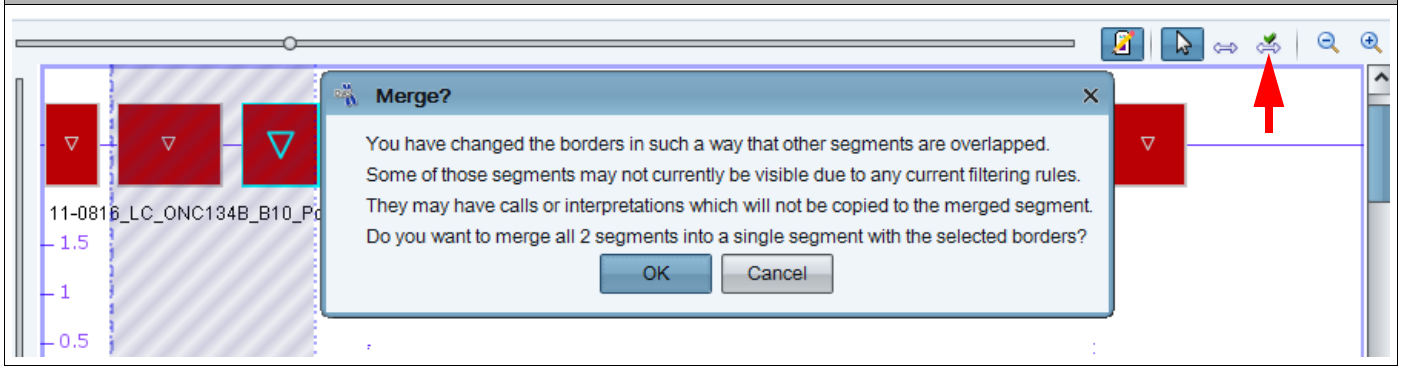
Figure 9.12 Click Start Editing button



7. Click the **Start/End** button.

   The following message appears: Figure 9.13

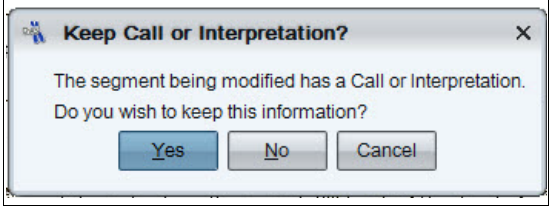
Figure 9.13 Click Start Editing button



8. Acknowledge the message, then click **OK.**

   If the segment whose Start or End you are editing contains **Call or Interpretation** information, the following message appears: Figure 9.14

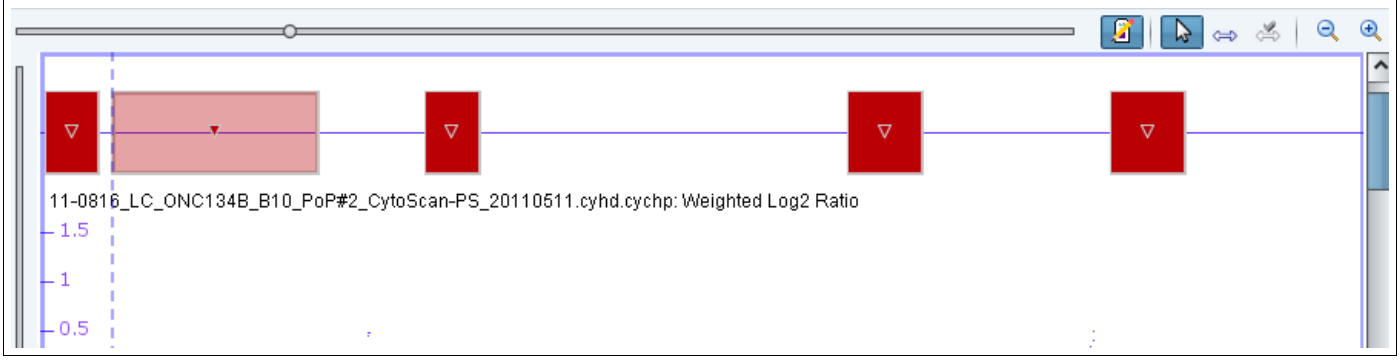
Figure 9.14 Keep Call or Interpretation message



9. Click the appropriate button.

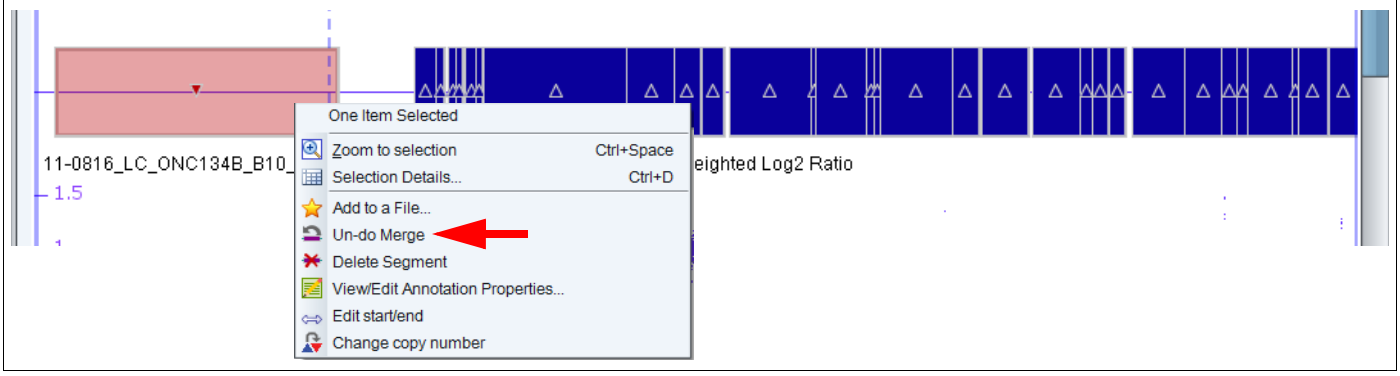
Your two segments are now merged together. While Edit Mode is **ON**, the segment appears in a distinct color. (Figure 9.15). (When Edit Mode is **OFF**, the edited segments look identical to non-edited segments.)

Figure 9.15 Click Start Editing button



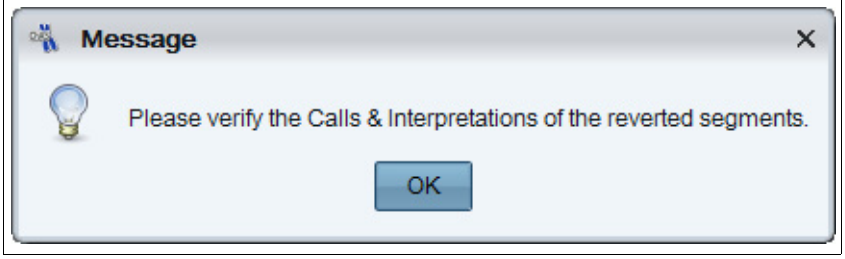
Un-Merging Merged Segments

Figure 9.16 Merged segments



1. Right-click on the newly merged segments (Figure 9.16), then click **Un-do Merge**
2. The following message appears. (Figure 9.17)

Figure 9.17 Confirm delation



3. Acknowledge the message, then click **OK**.

   Your previously merged segments return to their original states, including any Calls or Interpretations made on the original segments (prior to their merging).

Deleting Segments

1. Click **File** → **Open**.

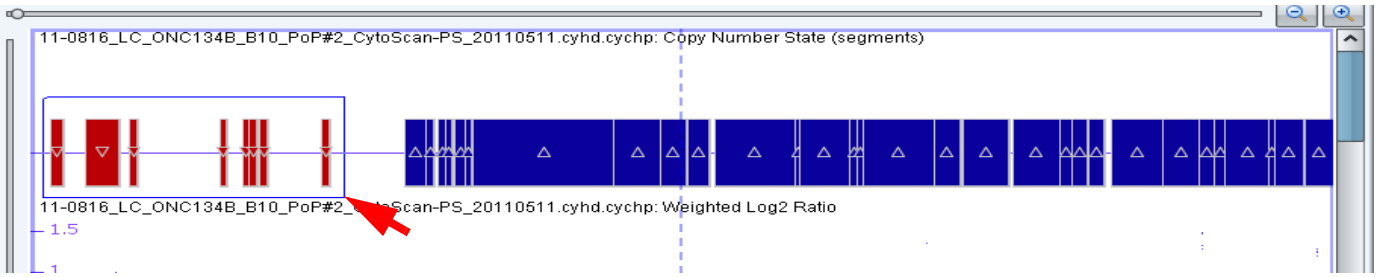
   Your Sample File data folder window appears.

2. Click to select the file you want to edit, then click **Open**.

   The file appears in the ChAS browser.

3. Left-click, hold, then move the mouse to frame the segments you want to delete. (Figure 9.18)

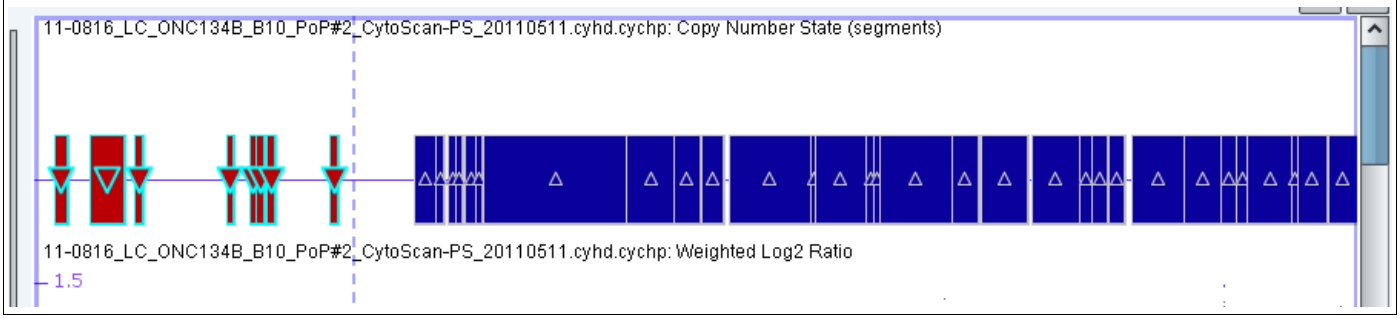
Figure 9.18 Framing your segments for delation



4. Release the mouse button.

   Your selected segments, including their *Loss* symbols are now highlighted, as shown in the example below. (Figure 9.19)

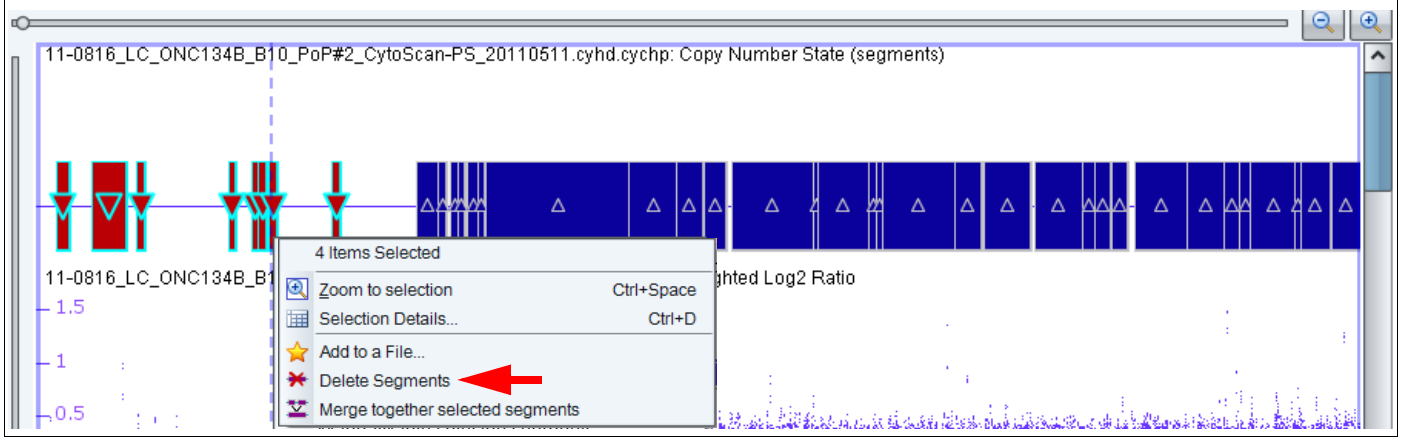
Figure 9.19 Framed segments now highlighted



5. Right-click on a highlighted segment.

The following menu appears. (Figure 9.20)

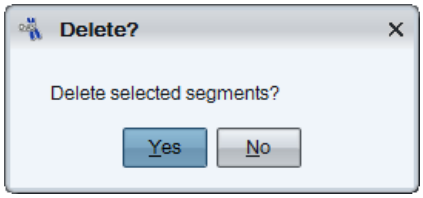
Figure 9.20 Right-click menu



6. Click **Delete Segments**.

   A message regarding your planned deletion appears. (Figure 9.21)

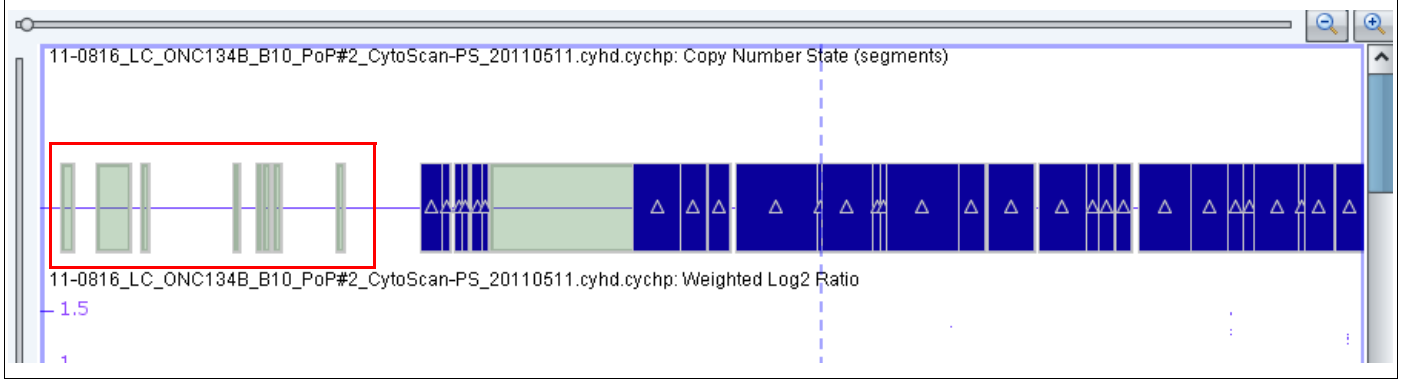
Figure 9.21 Confirm delation



7. Click **OK** to acknowledge it.

   Your selected segments are deleted and graphically represented as "ghosts" while Edit mode is **ON**. (Figure 9.22). When Edit Mode is OFF, deleted segments are invisible and NOT displayed in any view or table.

Figure 9.22 Deleted segments

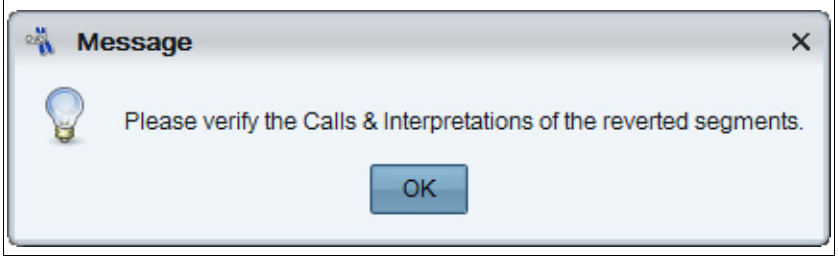


TIP: Even though your selected segments are deleted, ChAS preserves their data for future reference. See ***Right-click Menu Options* (on page 133)** for information on how to view and edit segment properties and view segment details.

Un-Deleting a Deleted Segment(s)

1. Right-click on the deleted segment(s), then click **Un-do Delete**.
2. The following message appears. (Figure 9.23)

Figure 9.23 Confirm delation



3. Acknowledge the message, then click **OK**.

   Your previously deleted segment(s) return to their original states, including any Calls or Interpretations made on the original segments (prior to their deletion).

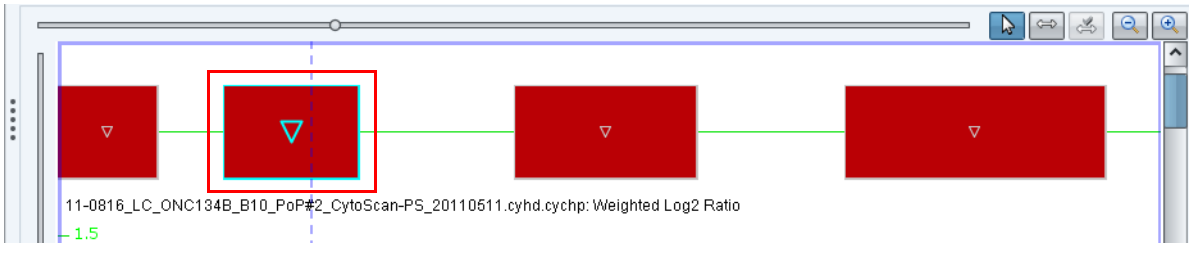
Editing the Start/End Coordinates of a Segment

IMPORTANT: Only the boundaries of 1 segment, can be adjusted/modified at one time.

NOTE: Segment start or end boundaries can be moved left and/or right. The Adjusting/ Modifying Segment Boundaries example that follows, shows how to move a segment end boundary farther to the right.

1. Using the zoom tools and scroll bars (if needed), identify the segment whose start or end (or both) you want to modify
2. Single-click on the segment to highlight it (Figure 9.24).

Figure 9.24 Selected segment

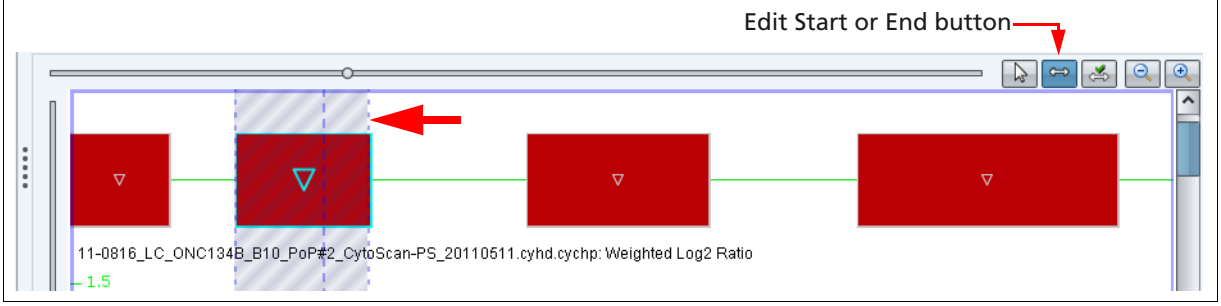


3. Click the **Start editing...** button.

   The following appears: (Figure 9.25 on page 156)

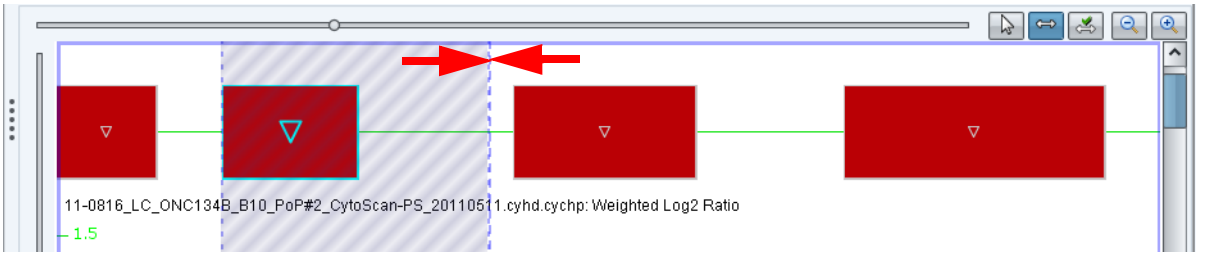
4. Place the mouse cursor on the very-right edge of the current segment boundary. (Figure 9.25)

Figure 9.25 Selected segment - Start editing



5. Click, then drag the segment's right-edge boundary to the right, then stop at an appropriate point. (Figure 9.26) You will ONLY be allowed to set the Start or End position to match the position of a marker probeset.

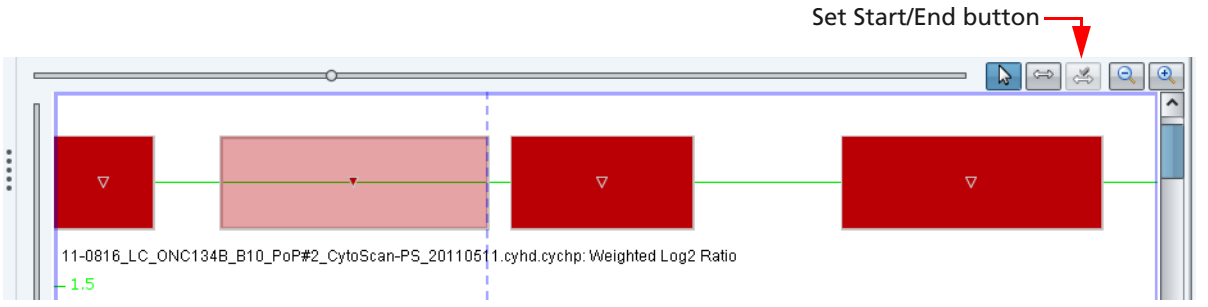
Figure 9.26 Selected segment - Start editing



6. Click the **Set start/end**... button. (Figure 9.27)

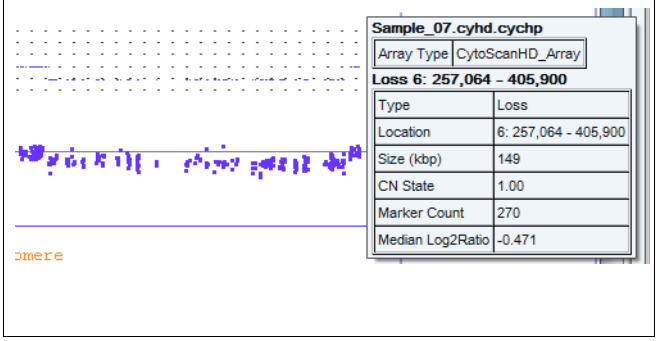
   The newly adjusted/modified segment boundary appears as shown. (Figure 9.27)

Figure 9.27 Selected segment - Start editing



7. ChAS auto-calculates various properties of the newly edited segment. Mouse over the segment to view its new details. (Figure 9.28)

Figure 9.28 Selected segment - Start editing



Un-Doing the Edited Start/End Coordinates of a Segment

1. Right-click on the previously adjusted/modified segment, then click **Un-do Edit**.

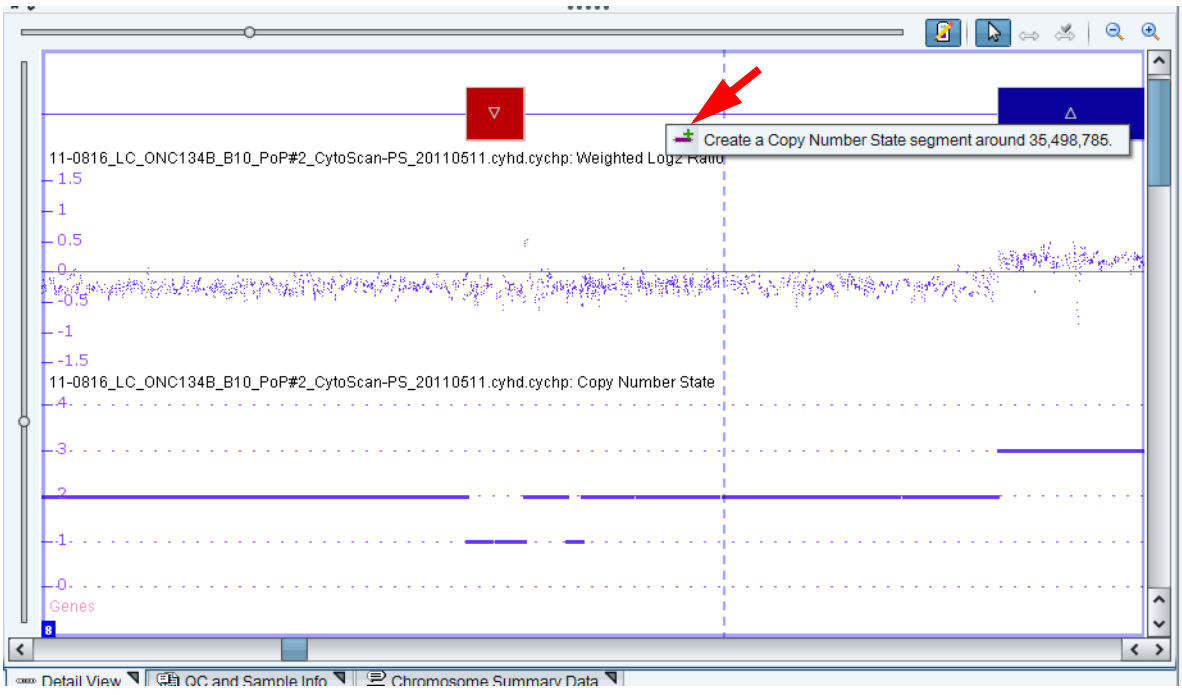
   The segment returns to its previous state, including, previous boundaries, calls, and interpretations.

De Novo Segment Drawing

1. Within an existing segment track, locate an empty space where you want to draw in a brand new segment.
2. Right-click inside this empty space.

   A **Create a Copy Number State segment around (*genomic coordinate of nearest probe*)** ribbon appears. (Figure 9.29)

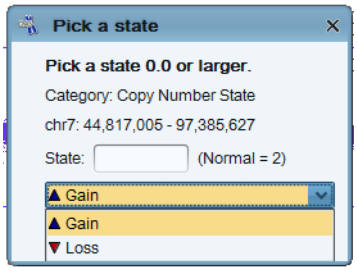
Figure 9.29 Selected segment - Change Copy Number - Pick a state window



3. Click on this ribbon. (Figure 9.29)

   The Pick a state window appears (Figure 9.30)

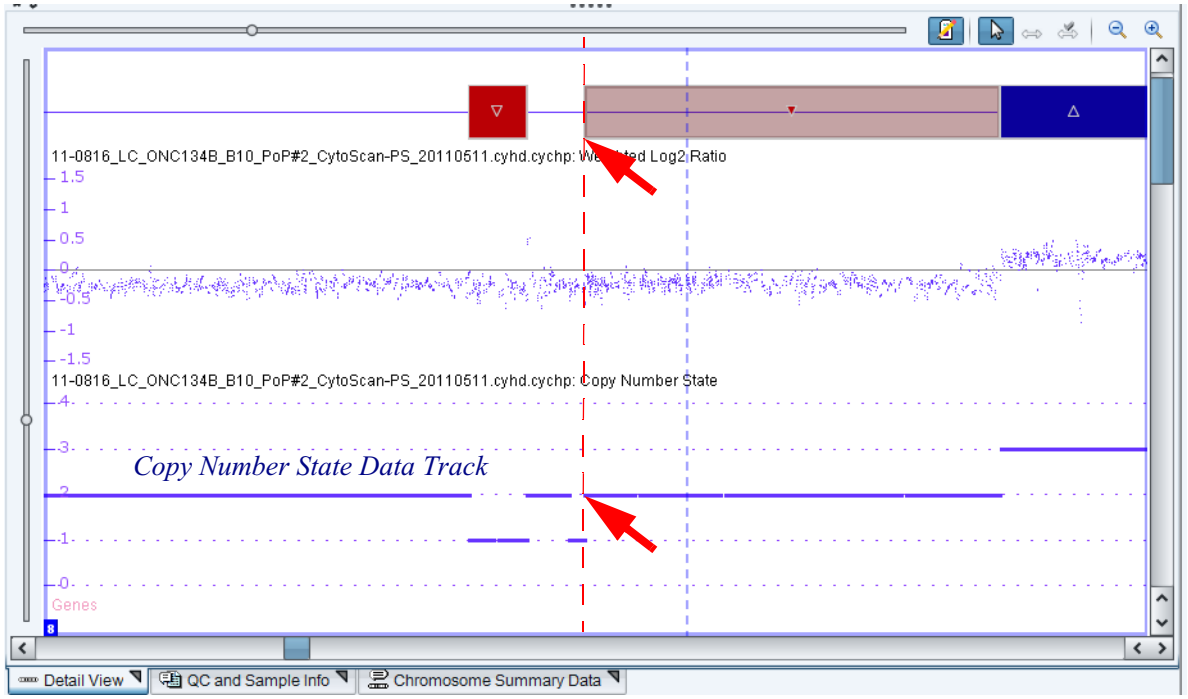
Figure 9.30 Drawing a new segment -- Pick a state window



4. Enter a **Copy Number** and **Type** for the new segment.
5. Click **OK**

   Your newly drawn segment appears. (Figure 9.31).

Figure 9.31 Newly drawn segment limited in size by the Copy Number State track

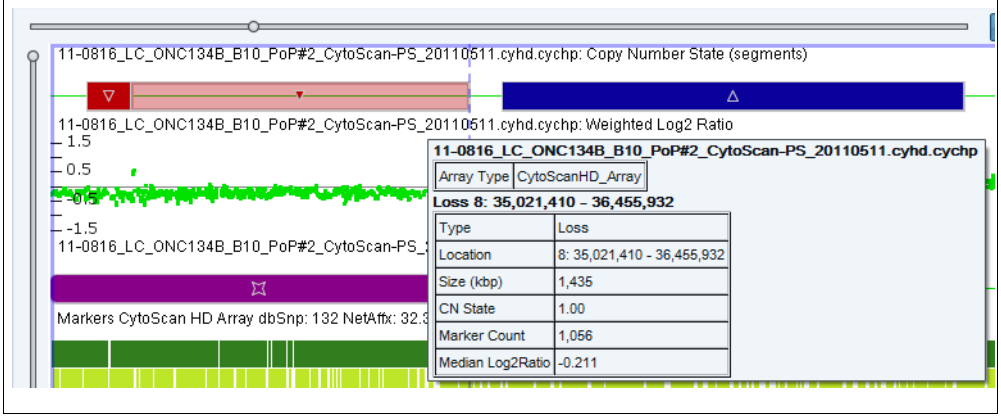


IMPORTANT: Newly drawn segments will stop when they encounter another segment, whether or not that segment is currently drawn or whether it is currently filtered out; see Figure 9.31 for an example in the CN State data track which is used to draw Copy Number Segments of Gain and Loss.

New segments will also stop when they encounter the last appropriate marker on the chromosome. Because of this, new segments drawn will vary in their initial size.

6. Move the mouse cursor over the newly drawn segment to reveal its properties. (Figure 9.32)

Figure 9.32 Newly drawn segment - Mouse over to see its properties



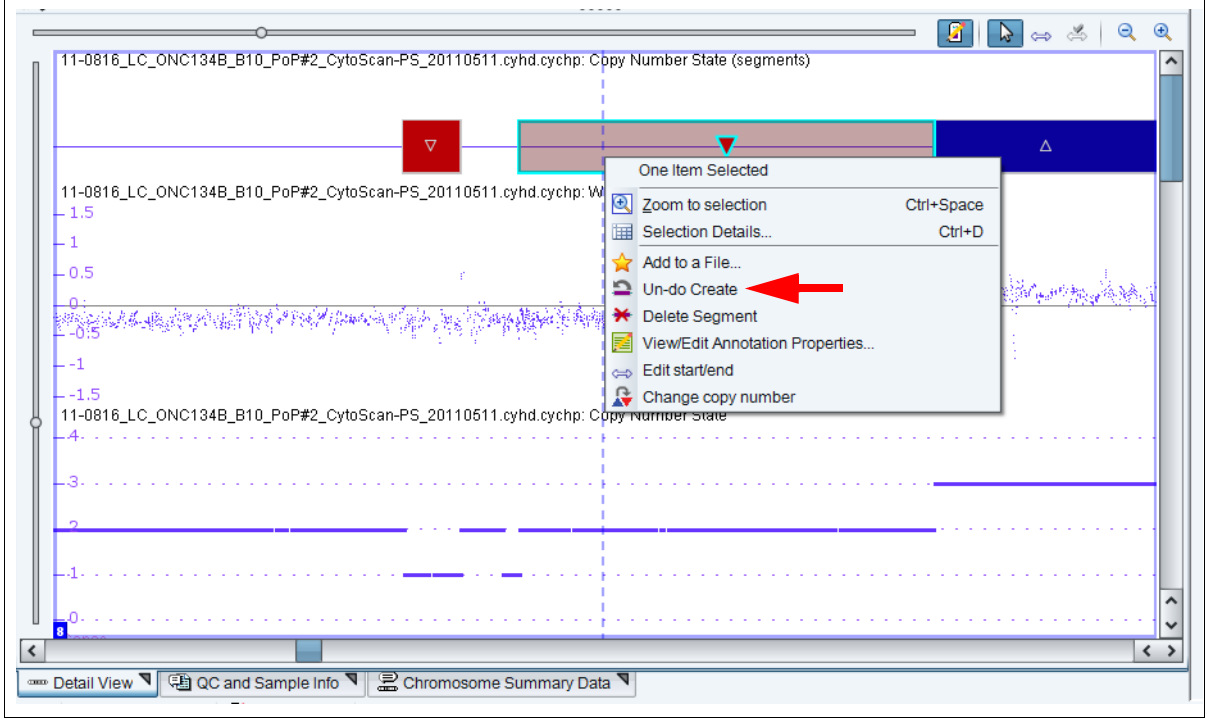
Un-Doing a Segment *De Novo* Drawing

To undo a drawn segment:

1. Right-click on the newly drawn segment.

   The following menu appears: (Figure 9.33).

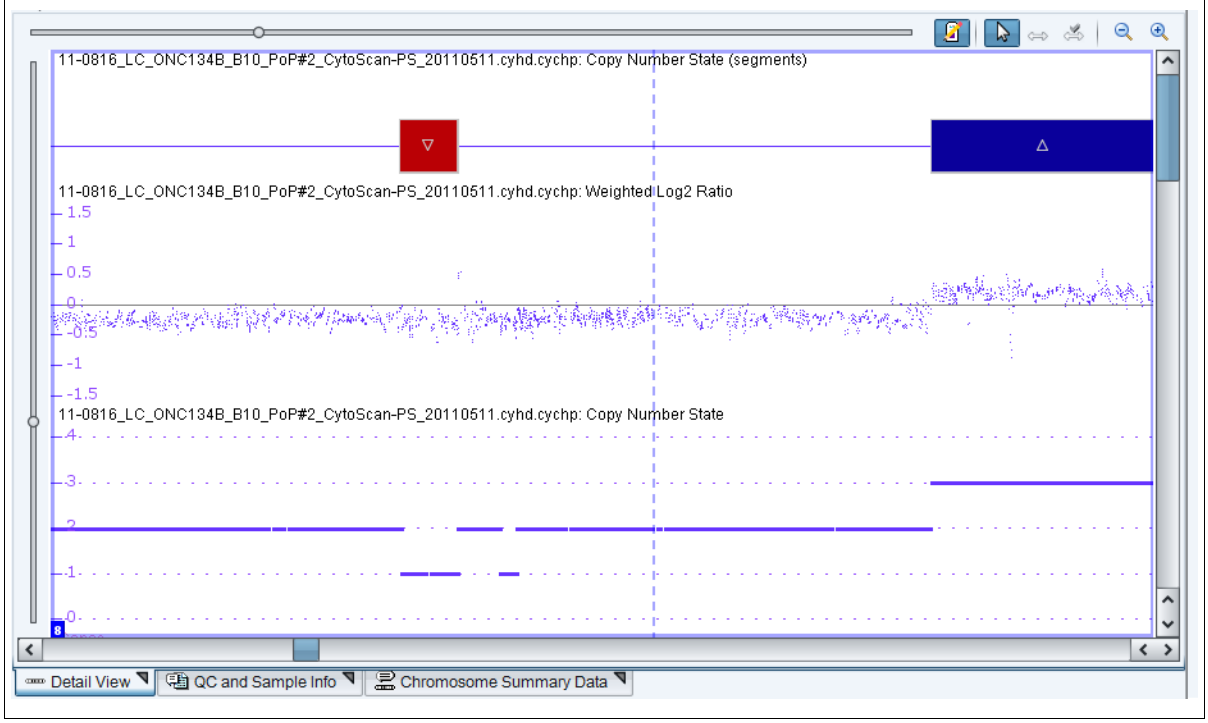
Figure 9.33 Newly drawn segment - Un-do drawn segment



2. Click **Un-do Create**.

The previously drawn segment is removed, along with any Calls or Interpretations it had been given. (Figure 9.34).

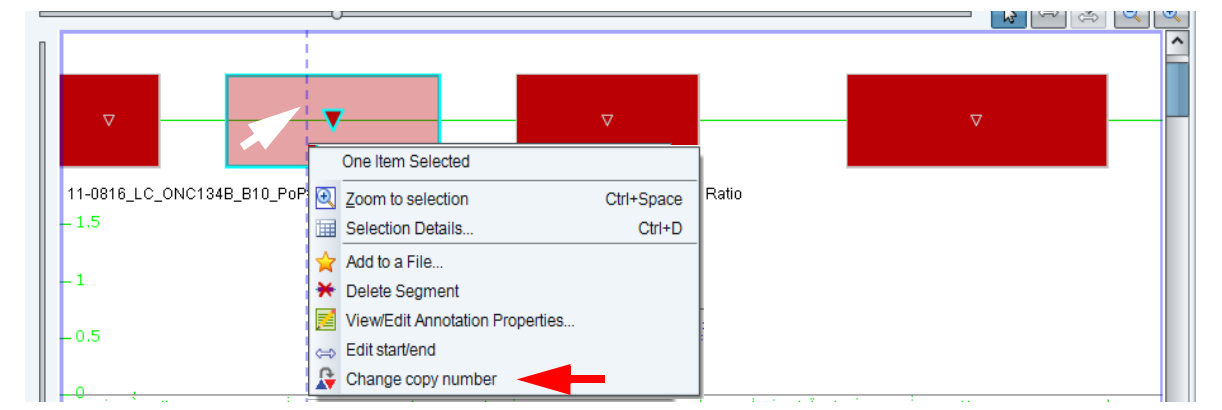
Figure 9.34 Newly drawn segment - Drawn segment removed



Changing the Copy Number

1. Right-click on the segment that contains the copy number you want to change
2. Click Change copy number. (Figure 9.35).

Figure 9.35 Selected segment - Change copy number



The **Pick a state** window appears: (Figure 9.36)

Figure 9.36 Selected segment - Change Copy Number - Pick a state window

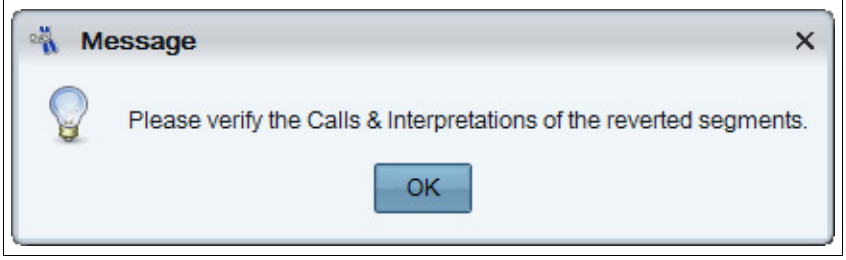
3. Enter a new Copy Number State value and Type (Gain/Loss).
4. Click **OK**.

   The segment now reflects your revised Copy Number State and Type.

Un-Doing a Copy Number Change

1. Right-click on the segment you performed a Copy Number State change on, then click **Un-do Edit**.
2. The following message appears. (Figure 9.37)

Figure 9.37 Confirm delation



3. Acknowledge the message, then click **OK**.

   The segment's original Copy State Number is reinstated.

Modified Segments in the Segments Table

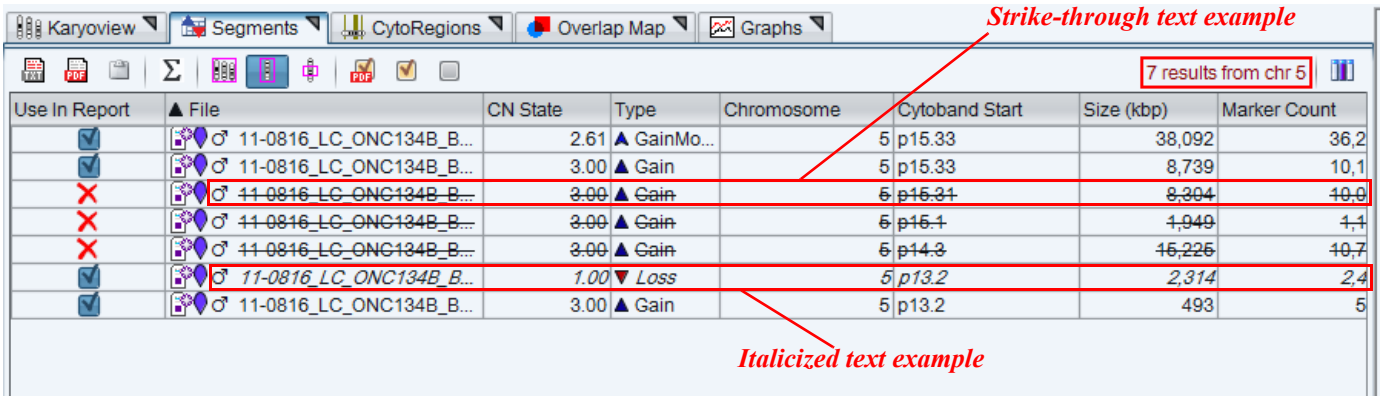
IMPORTANT: When exporting a Segments table to text (.txt), please note that deleted segments will be part of the export when Edit Mode is On, and will NOT be part of the export (when Edit Mode is Off). For more information, see *Export Tables as TXT File* (on page 280).

In PDF reports, deleted segments are never shown in graphical views, nor are they listed in the Segments Table, as Edit Mode is required to be OFF to generate a PDF report. For more information, see *Exporting Selected Segments to a PDF File* (on page 243).

When the Edit Mode is ON, the Modified segments appear differently within the Segments table. (Figure 9.38)

- Deleted Segments are represented with a **red X** and a **strike-through** line. Deleted segments do NOT show up in PDF reports because PDF reports cannot be created while Edit Mode is **ON**.
- Materially Modified Segments (including segments that have been merged, boundaries edited, or had their Copy Number States changed) are represented with ***italicized*** text.

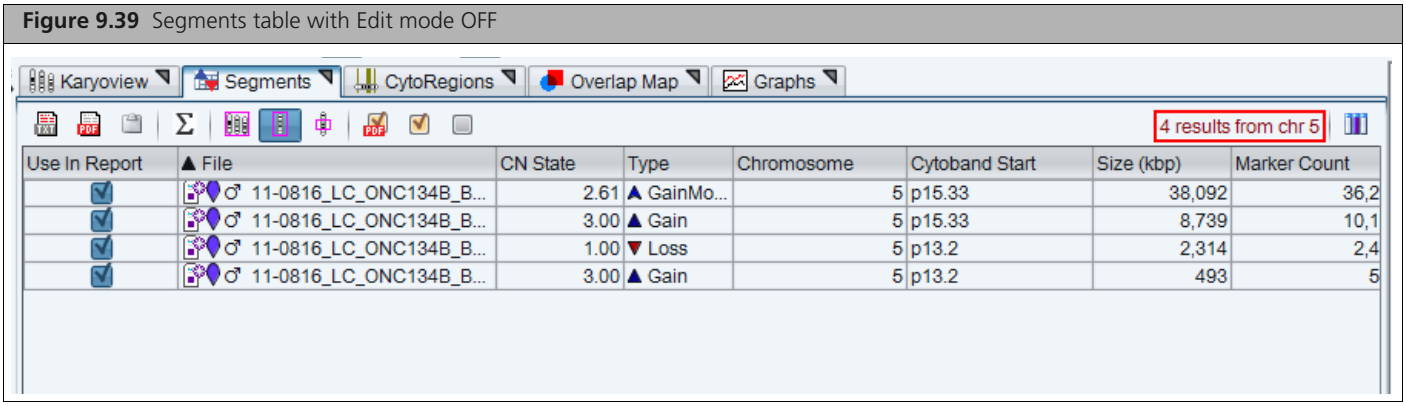
Figure 9.38 Segments table with Edit mode ON



Turning the Edit Mode OFF, displays the Segments table as follows: (Figure 9.39)

Note that the 3 deleted segment rows which were in strike-through text are removed from the table when Edit Mode is OFF. Also note, that the rows which indicated Modified segments are no longer italicized when Edit Mode is OFF.

Figure 9.39 Segments table with Edit mode OFF



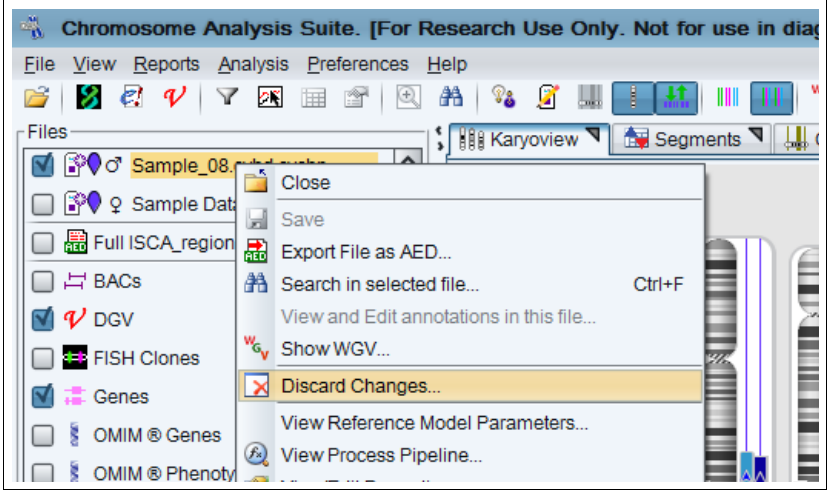
Use In Report	File	CN State	Type	Chromosome	Cytoband Start	Size (kbp)	Marker Count
☑	♂ 11-0816_LC_ONC134B_B...	2.61	▲ GainMo...	5	p15.33	38,092	36,2
☑	♂ 11-0816_LC_ONC134B_B...	3.00	▲ Gain	5	p15.33	8,739	10,1
☑	♂ 11-0816_LC_ONC134B_B...	1.00	▼ Loss	5	p13.2	2,314	2,4
☑	♂ 11-0816_LC_ONC134B_B...	3.00	▲ Gain	5	p13.2	493	5

Revert All Edits for a Sample

All edits done to a sample can be removed in a single step. To do this:

1. Right-click on the File name in the Files list, then select **Discard Changes** (Figure 9.40)

Figure 9.40 Discard Changes

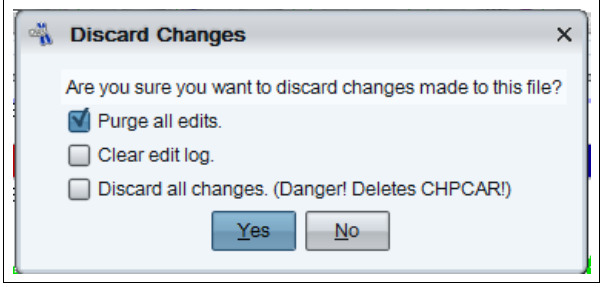


The Discard Changes window appears and displays the following options:

- **Purge All Edits**: Reverts all edits to segments and removes all calls and interpretations. Note: The log of the edits that have been performed on the sample can still be viewed.
- **Clear Edit Log**: Purges all edits and clear the log of any edits performed on the sample.
- **Discard all changes**: Purges all edits, clears the edit log, and deletes the CHPCAR file.

2. Check the appropriate box, then click **Yes**. (Figure 9.41)

Figure 9.41 Discard Changes Confirmation window



For more information on viewing Edits and Edit logs see *Tracking (Reviewing the Log File)* (on page 173).

Setting Up and Using the Calls Feature

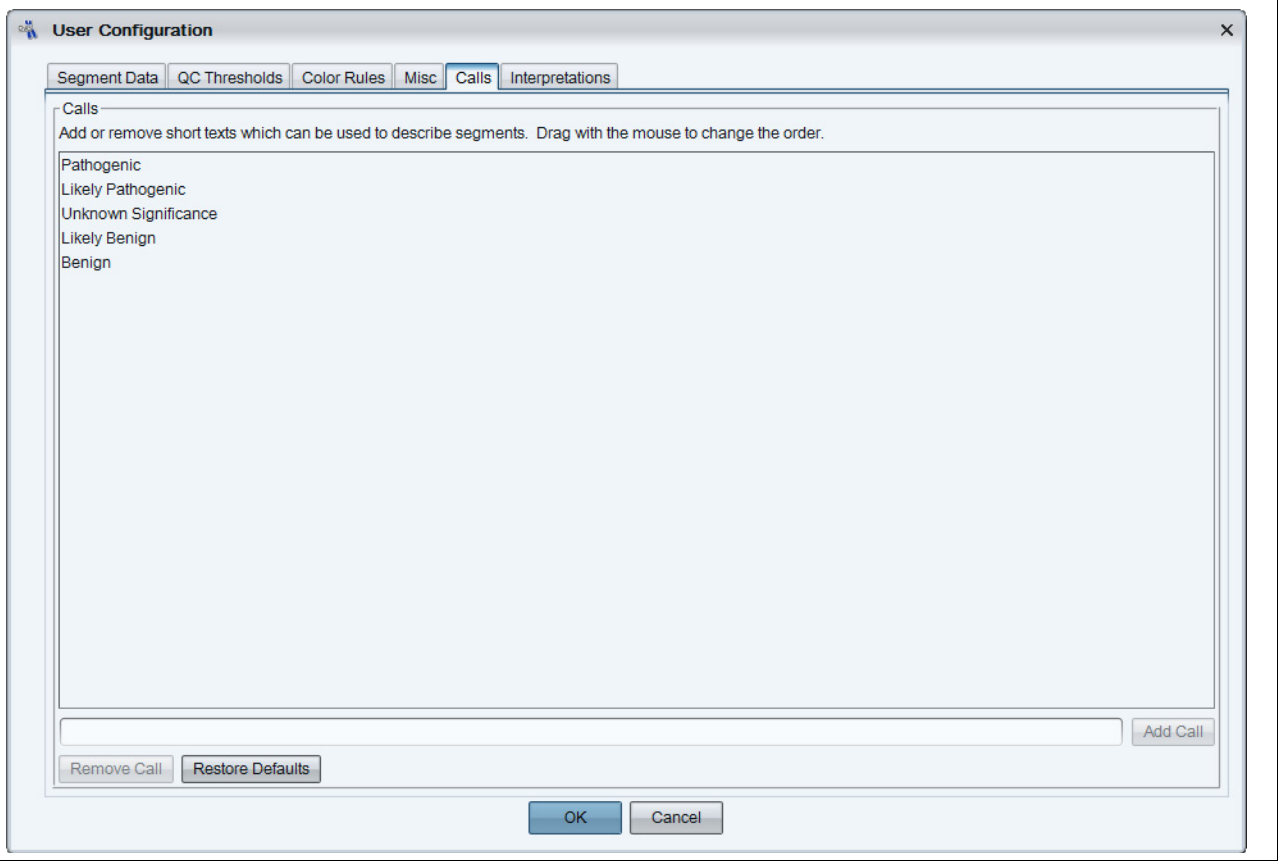
NOTE: If you are using a user profile from a previous version of ChAS, your default set of Calls will NOT appear in the Calls drop list until "Restore Defaults" button (Figure 9.42 on page 165) is clicked on in the *Edit User Configurations* → *Calls* window tab.

Setting Up Calls

1. Click **Preferences** → **Edit User Configurations** or click on the upper tool bar.
   The User Configuration window appears.
2. Click the **Calls** tab.

The Calls window appears. (Figure 9.42)

Figure 9.42 User Configuration window - Calls tab window



Adding, Deleting, and Re-arranging Calls

The Calls window allows you to Add, Delete and Re-arrange current Calls.

Do the following to add calls to the Call drop-down list:

1. Click inside the Add Call field, then enter your new Call.
2. Click **Add Call**.

   Your newly added Call now appears in the Call drop-down-list.

Do the following to delete calls from the Call drop-down list:

1. Click to highlight the Call you want to delete.
2. Click **Remove Call**.

   Your newly deleted Call is removed from the Call drop-down-list.

Do the following to re-arrange the order of Calls displayed in the Call drop-down list:

1. Click to highlight the Call you want to move.
2. Drag and drop it to its new location (order).

   Your Call is now in its revised position/order within the Call drop-down-list.
3. If needed, repeat steps 1-2 to re-arrange additional Calls.

Restoring Default Calls

If using a user profile started in a previous version of ChAS, you can populate the default Calls list to make them available in the Call drop-down list.

Do the following to restore default Calls:

1. Click **Restore Defaults**.

   Your previously default Call(s) are now populated in the Call drop-down-list.

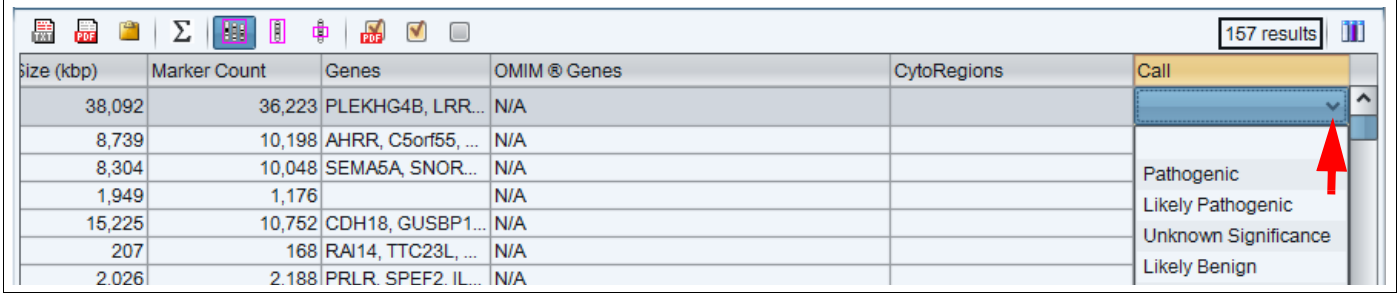
Using the Calls Feature

Method 1: At the Segments Table

1. Click the Browser's **Segments** tab.
2. Scroll the Segment table to the right until you see the **Call** column.
3. Locate the appropriate row.
4. Single-click inside the field.

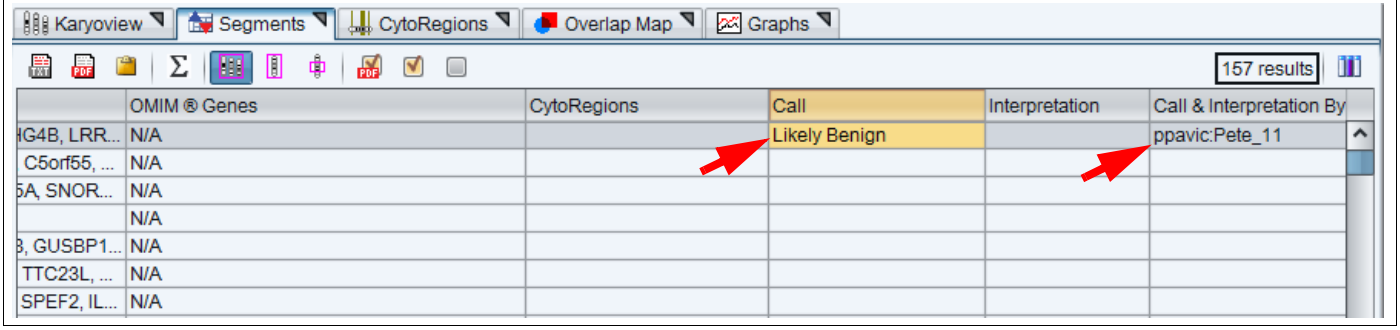
   A blue drop-down bar appears.
5. Click on the drop-down to reveal the list of available Calls. (Figure 9.43)

Figure 9.43 Calls drop-down menu barr



6. Click to select the appropriate Call.
7. Click outside the field or press **Enter** on the keyboard.
8. Your Call is entered. Note that the **Calls & Interpretation By** column is populated with the user's Windows login ID (left) and ChAS User Profile ID (right). (Figure 9.44)

Figure 9.44 Call is entered and Call &Interpretation By field is populated



Method 2: At the View/Edit Annotation Properties Window

1. Click **File** → **Open**.

   Your Sample File data folder window appears.

2. Click to select the file you want to edit, then click **Open**.

   The file appears in the ChAS browser.

3. At the Visualization window (lower pane) click the **Details View** tab.

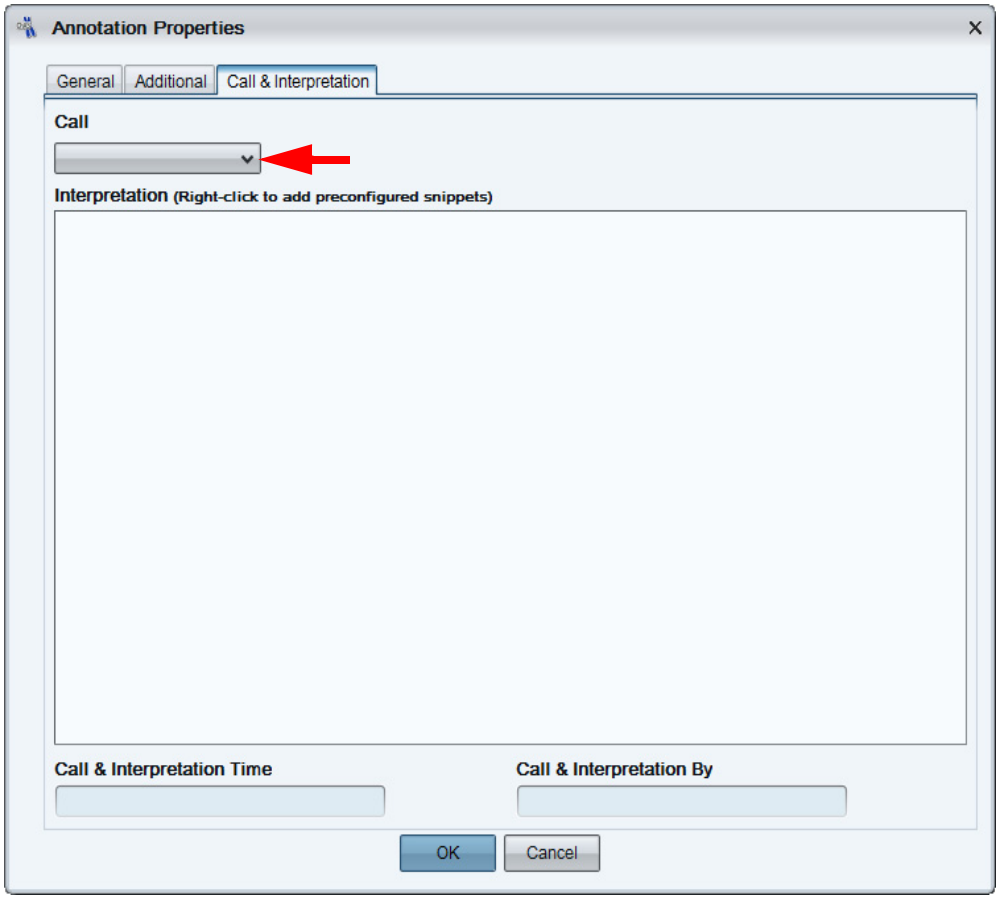
4. Right-click on the segment you want to add an interpretation to, then click the menu selection **View/Edit Annotation Properties**.

   The Annotation Properties window appears.

5. Click the **Calls & Interpretation** tab.

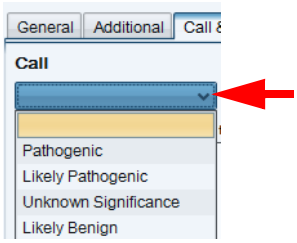
   The Interpretation window tab appears: (Figure 9.45)

Figure 9.45 Annotation Properties window - Interpretation window



6. Click the **Call** drop-down menu, then select your appropriate interpretation call. (Figure 9.46)

Figure 9.46 Call drop-down menu



Setting Up and Using the Interpretations and Snippets Feature

Setting Up the Interpretation Snippets Feature

The Snippets feature can be used in conjunction with the free-typing Interpretation field (below the Calls drop-down). It allows for a convenient "shortcut" when common words or phrases are used often in an interpretation.

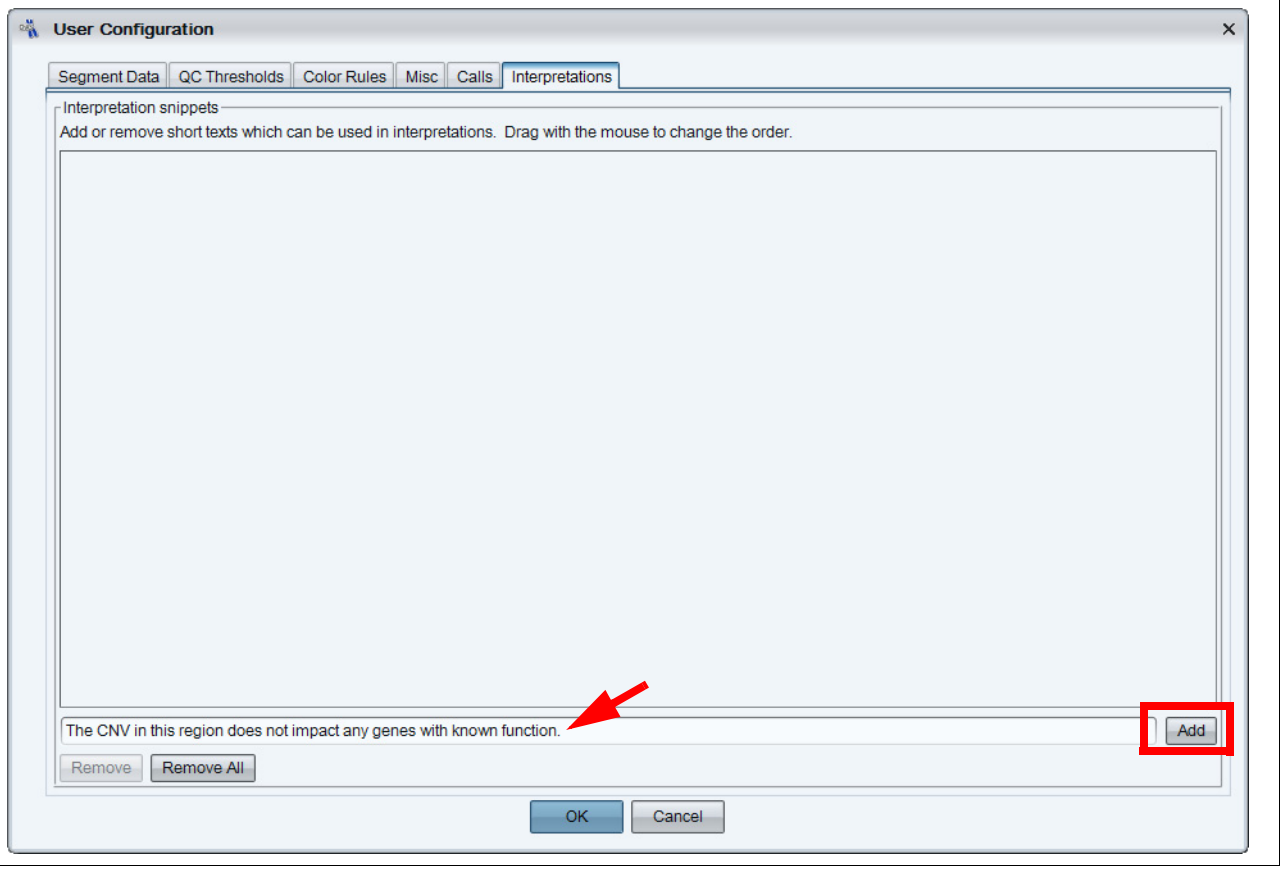
1. Click **Preferences** → **Edit User Configurations** or click on the upper tool bar.

   The User Configuration window appears.

2. Click the **Interpretations** tab.

   The Interpretations window appears. (Figure 9.47 on page 169)

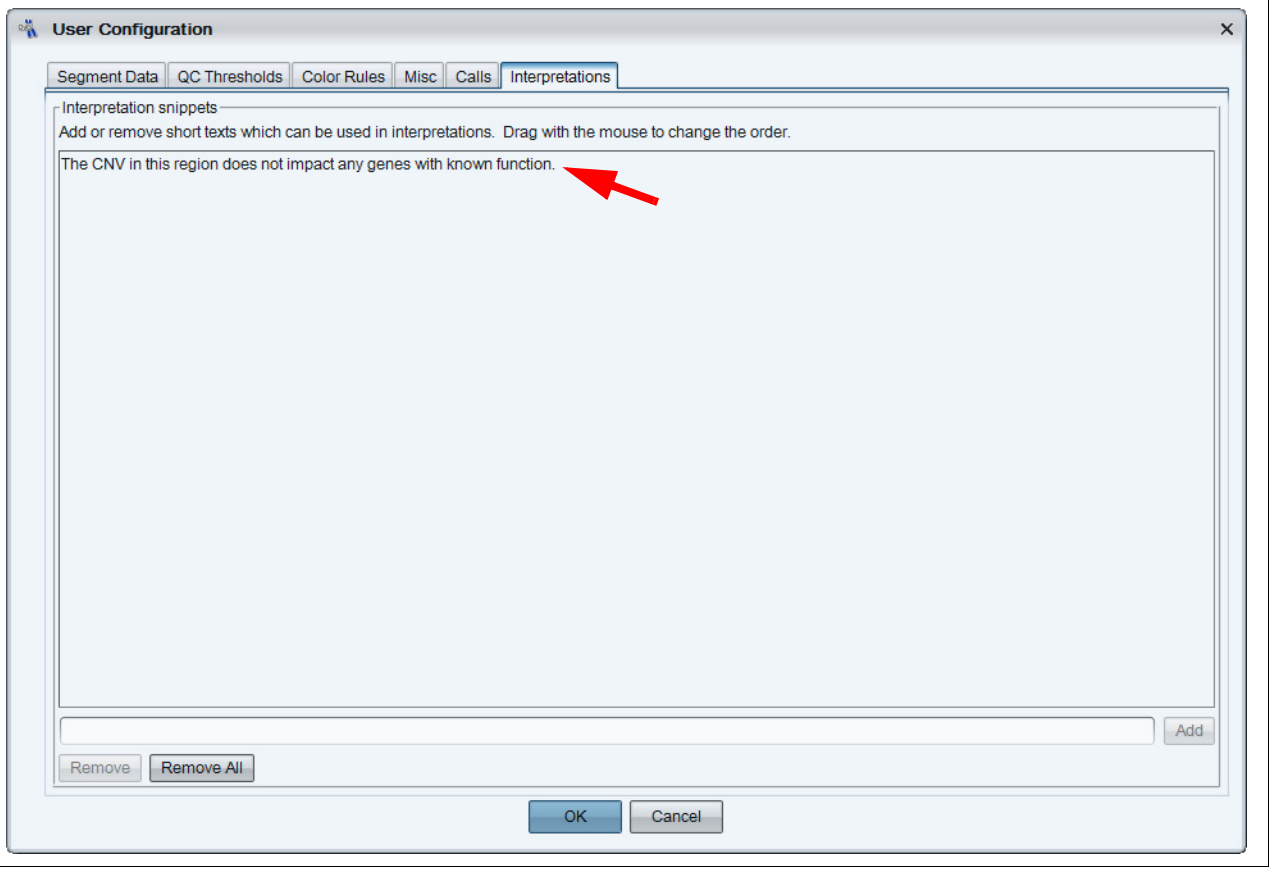
Figure 9.47 User Configuration window - Interpretation tab window



3. At the Interpretation window, click inside the field shown (Figure 9.47), then enter the snippet you want to use with your interpretation(s).
4. Click **Add**.

   The snippet now appears and is saved in the Interpretations List pane. (Figure 9.48 on page 170)
5. Repeat steps 3 and 4 to add additional snippets.
6. Click **OK**.

Figure 9.48 Interpretations window - Snippet added to Interpretations list



Removing Interpretation Snippets

To remove a saved snippet:

- Click to highlight the snippet you want to remove.
- Click **Remove**.

To remove multiple saved snippets:

- Shift click or Ctrl click to highlight each snippet you want to remove.
- Click **Remove All**.

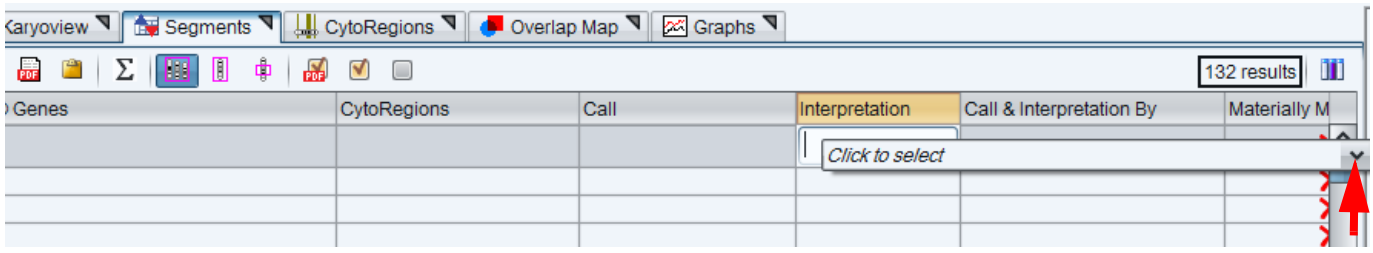
Using the Interpretation Snippets Feature

Method 1: At the Segments Table

1. Click the Browser's **Segments** tab.
2. Scroll the Segment table to the right until you see the **Interpretation** column.
3. Locate the appropriate row.
4. Double-click inside the field.
5. Right-click on the flashing cursor. If the field has existing text, place the flashing cursor to the point where you want to add a snippet, then right-click.

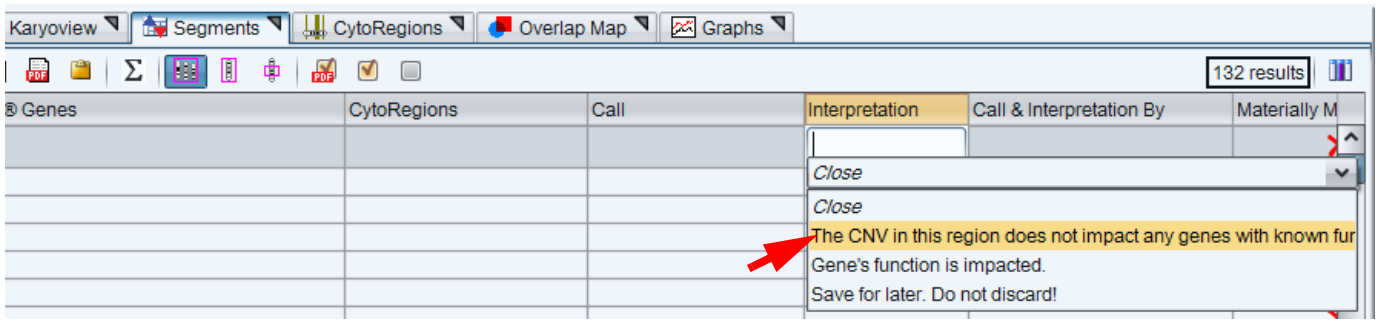
A floating drop-down menu bar appears. (Figure 9.49)

Figure 9.49 Floating Snippets drop-down menu barr



6. Click the snippet bar's drop-down to display the available snippets. (Figure 9.49)

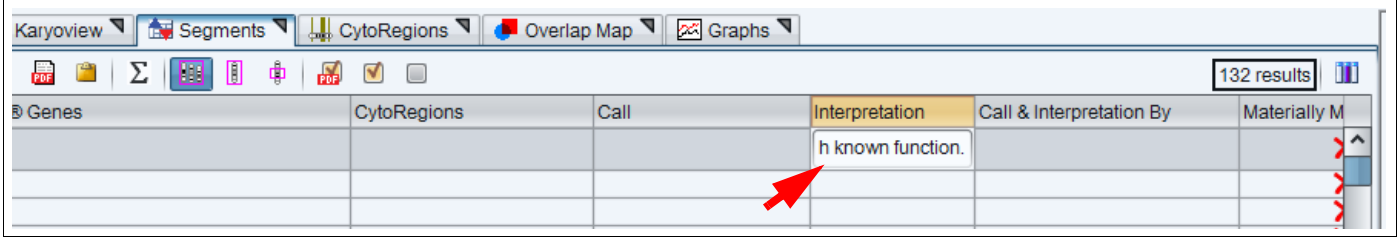
Figure 9.50 Snippets drop-down menu



7. Click on the appropriate snippet. (Figure 9.50)

   The snippet appears in your Interpretation row. (Figure 9.51)

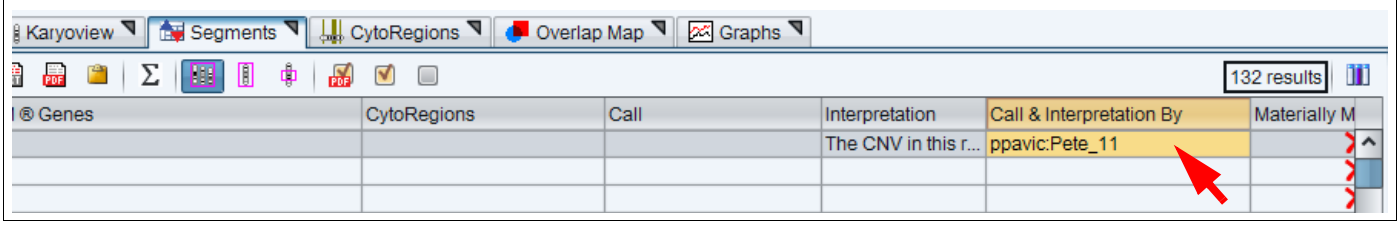
Figure 9.51 Snippet appears in field



8. Click outside the field or press **Enter** on the keyboard.

   Your snippet is entered. Note that the **Calls & Interpretation By** column is populated with the user's Windows login ID (left) and ChAS User Profile ID (right). (Figure 9.52)

Figure 9.52 Call & Interpretation By field now populated



Method 2: At the View/Edit Annotation Properties Window

1. From any graphical view (Karyoview, Selected Chromosome View, or Details View), right-click on the segment you want to add an interpretation to, then click the menu selection **View/Edit Annotation Properties**.

   The Annotation Properties window appears.

2. Click the **Call & Interpretation** tab.

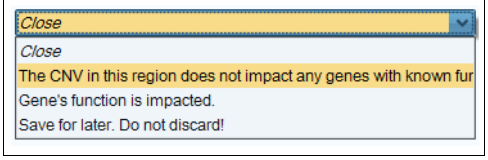
3. Type in your interpretation. If at any point you want to insert your preset snippet(s), right-click inside the interpretation window.

   A drop-down bar graphic appears. (Figure 9.53).

Figure 9.53 Drop-down menu bar

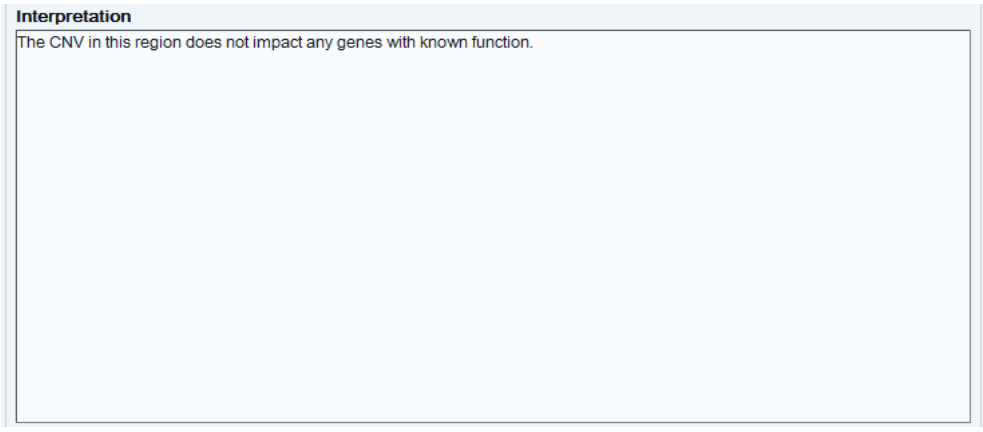
4. Click to select the appropriate snippet or click **Close** to exit. (Figure 9.54)

Figure 9.54 Snippet drop-down



Your snippet (preset word, sentence, or phrase) now appears “in line” with your typed text. (Figure 9.55)

Figure 9.55 Snippet in the Interpretation field pane



Tracking (Reviewing the Log File)

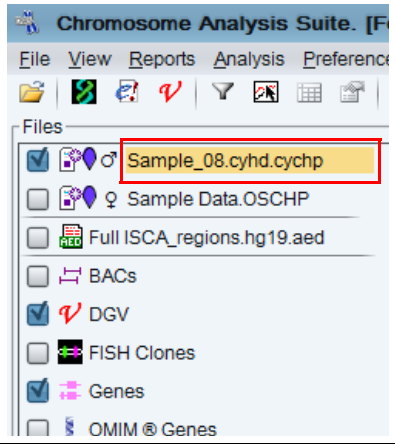
1. Click **File** → **Open**.

   Your Sample File data folder window appears.

2. Click to select the file you want to view the history of, then click **Open**.

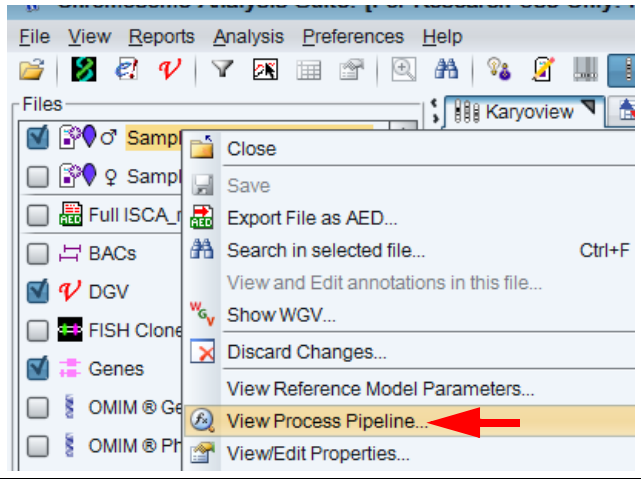
   The file appears in the ChAS browser and Files window pane (upper left). (Figure 9.56)

Figure 9.56 Files window



3. Right-click on the file, then click **View Process Pipeline**. (Figure 9.57)

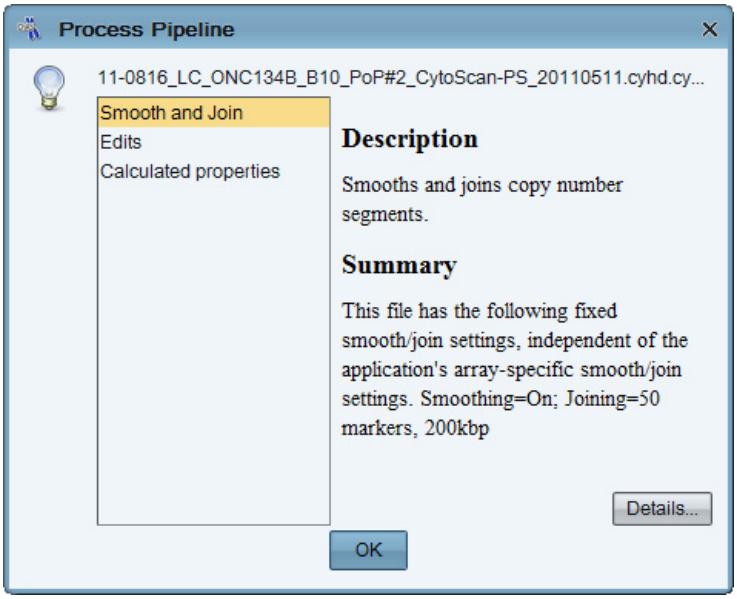
Figure 9.57 Right-click menu



The Process Pipeline window appears.(Figure 9.58 on page 174)

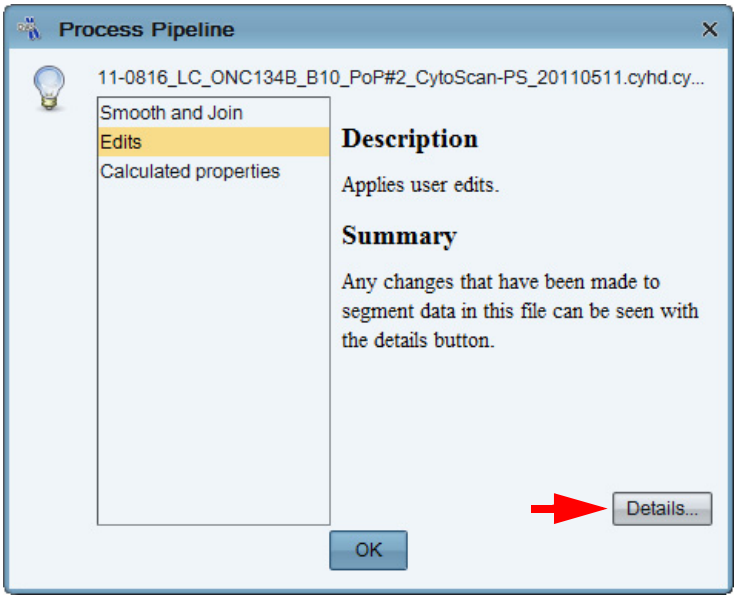
4. Click **Smooth and Join** for its description and summary.

Figure 9.58 Process Pipeline - Smooth and Join



5. Click **Edits** for its description and summary of the file's past user edits. (Figure 9.59)

Figure 9.59 Process Pipeline -Edits

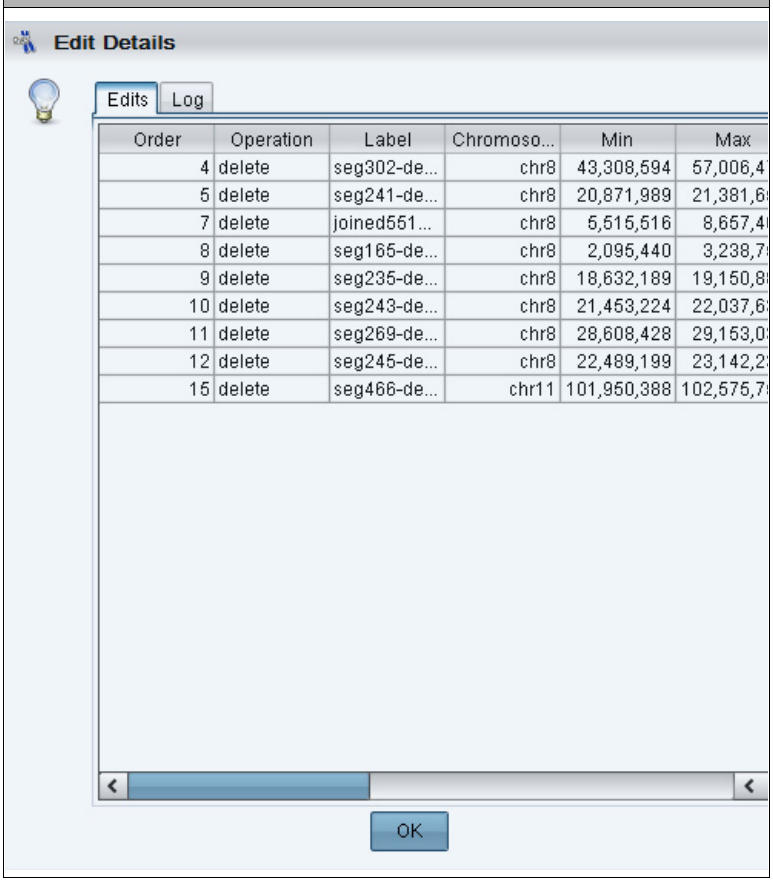


6. Click **Details.**

   A window featuring 2 tabs appears. (Figure 9.60 on page 175)

7. The **Edit** tab lists only recent edits that were NOT reverted. Use the horizontal scroll bar to reveal additional edit information. For a complete historic account of all edits made (including reversions of some edits), click on the **Log** tab.

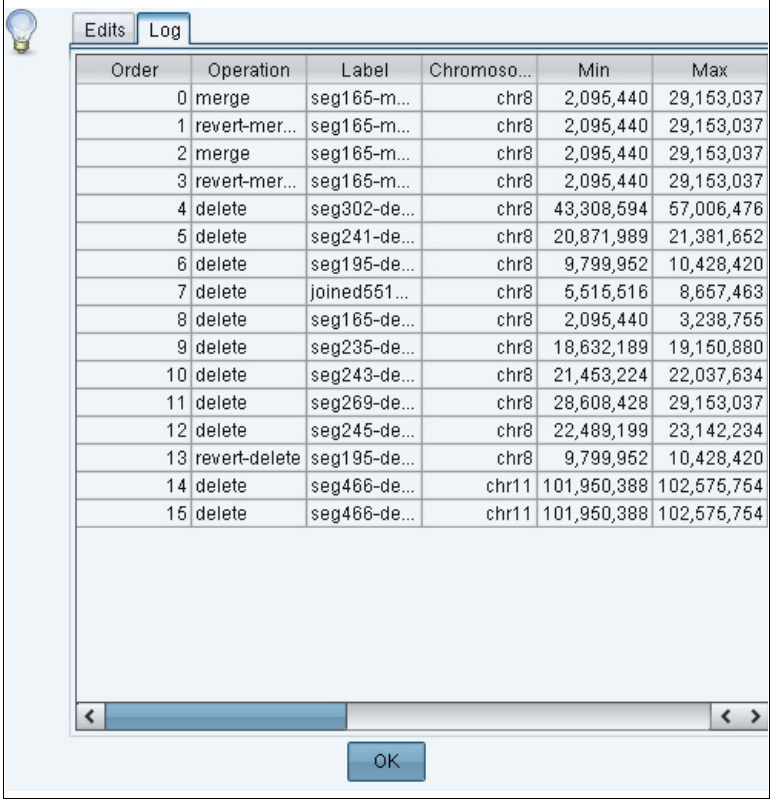
Figure 9.60 Process Pipeline - Edit tab (Figure 9.60)



8. Click **OK** to return to the Process Pipeline window or click the **Log** tab.

9. The **Log** tab displays an extensive summary of the file's edited history to date. Use the horizontal scroll bar to reveal additional log information. (Figure 9.61)

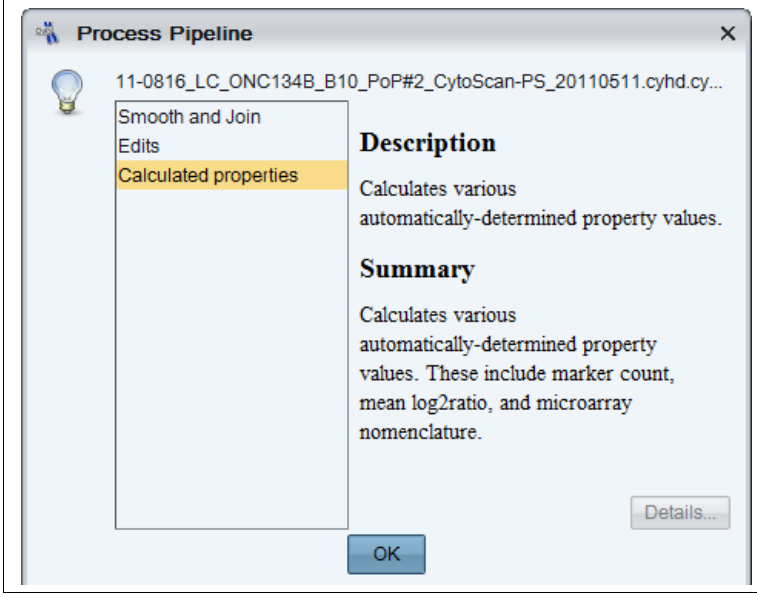
Figure 9.61 Process Pipeline - Log tab



10. Click **OK** to return to the Process Pipeline window.

11. Click **Calculated Properties** for its description and summary. (Figure 9.62)

Figure 9.62 Process Pipeline - Calculated properties



12. Click **OK** to return to the ChAS browser.

Using CytoRegions

The CytoRegions feature enables you to define parts of the genome that are of special interest to you. To use CytoRegions, you need to select a file with position information for regions of the genome as the CytoRegions file.

You can select a Region information file in AED or BED format. You can use a pre-existing Region file, or create a new one in AED format in ChAS and add regions to it by selecting segments, annotations, or regions in other loaded files.

You can add annotations to regions to help you track the information. See *Creating and Editing AED Files* (on page 196).

NOTE: The CytoRegions feature is intended for use with up to a few thousand regions. Larger numbers of regions can be used, but will impact performance. A reference annotation file, such as BAC or Genes, is not recommended for use as a CytoRegions file due to the large number of reference annotations.

After selecting a CytoRegions file, you can:

- Use the Restricted Mode to display only Segments and graph data that appear in those regions. Annotations are not hidden by CytoRegions or by the application of Restricted Mode.
- Use differential filtering options for these regions and for the rest of the genome.

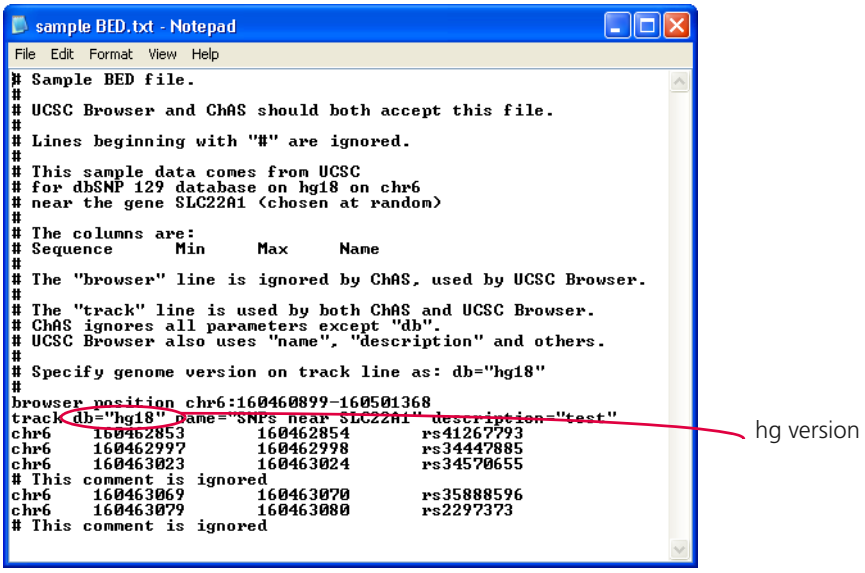
This chapter describes:

- *Selecting a CytoRegions Information File*
- *Viewing CytoRegions* (on page 180)
- *Searching CytoRegions* (on page 185)
- *Using Filters with CytoRegions* (on page 185)
- *Using Restricted Mode* (on page 186)

Selecting a CytoRegions Information File

Select the CytoRegions file from the available region information files. See*Loading Files* (on page 68). The software automatically checks the hg version of an AED or BED file before loading (see Figure 10.1 for an example BED file). The file will not be loaded if the hg version does not match what is loaded in the ChAS Browser. If an hg version is not found for the AED or BED file, a warning message is displayed.

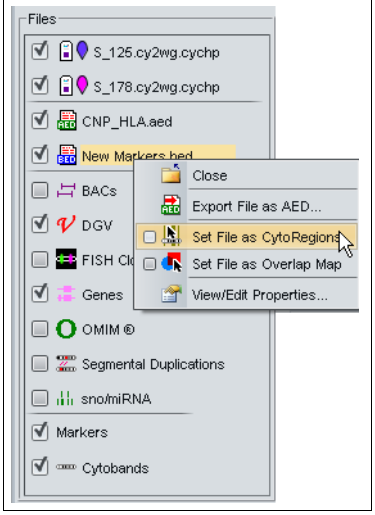
Figure 10.1 Example BED file



Do either of the following to select a CytoRegions file:

- In the files list, right-click a file and select **Set File as CytoRegions** on the shortcut menu (Figure 10.2).

Figure 10.2 Select a CytoRegions file from the file list



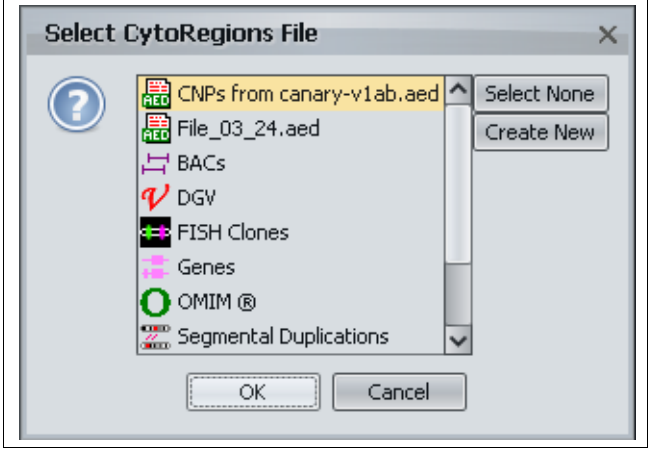
OR

1. Select **View** → **CytoRegions** on the menu bar.

   Alternatively, in the CytoRegions Tab of the Upper display, click the **Set CytoRegions** button.

   The Select CytoRegions dialog box opens (Figure 10.3).

Figure 10.3 Select CytoRegions File dialog box



2. Select a regions information file to use for cytoregions and click **OK**.

   The **Create New** feature is described in*Creating an AED File of Annotations* (on page 196).

NOTE: If a region file was selected for the CytoRegion file before shutting off the software, that file is automatically loaded as the CytoRegion file when the software is opened the next time with the same user profile.

NOTE: To clear a CytoRegions file map from ChAS, choose "Select None" in the dialog box (Figure 8.2). Alternatively, right-click the file in the Files list and select "Set File as CytoRegions" on the shortcut menu to toggle the check mark off.

After selecting a CytoRegions file, you can:

- *CytoRegions in the Graphic Views*
- *CytoRegions Table* (on page 182)
- *Using Filters with CytoRegions* (on page 185)
- *Using Restricted Mode* (on page 186)

Viewing CytoRegions

CytoRegions can be displayed in:

- *CytoRegions in the Graphic Views*
- *CytoRegions Table* (on page 182)

NOTE: CytoRegions that share genomic coordinates with a detected segment are listed in the "CytoRegions" column of the Segments table. See ***Segments Table* (on page 237)**.

CytoRegions in the Graphic Views

Regions specified in the CytoRegions file are displayed as dark gray stripes in the Karyoview and Chromosome Display (Figure 10.4) and Detail View (Figure 10.5).

Figure 10.4 Karyoview and Selected Chromosome View with CytoRegions Displayed

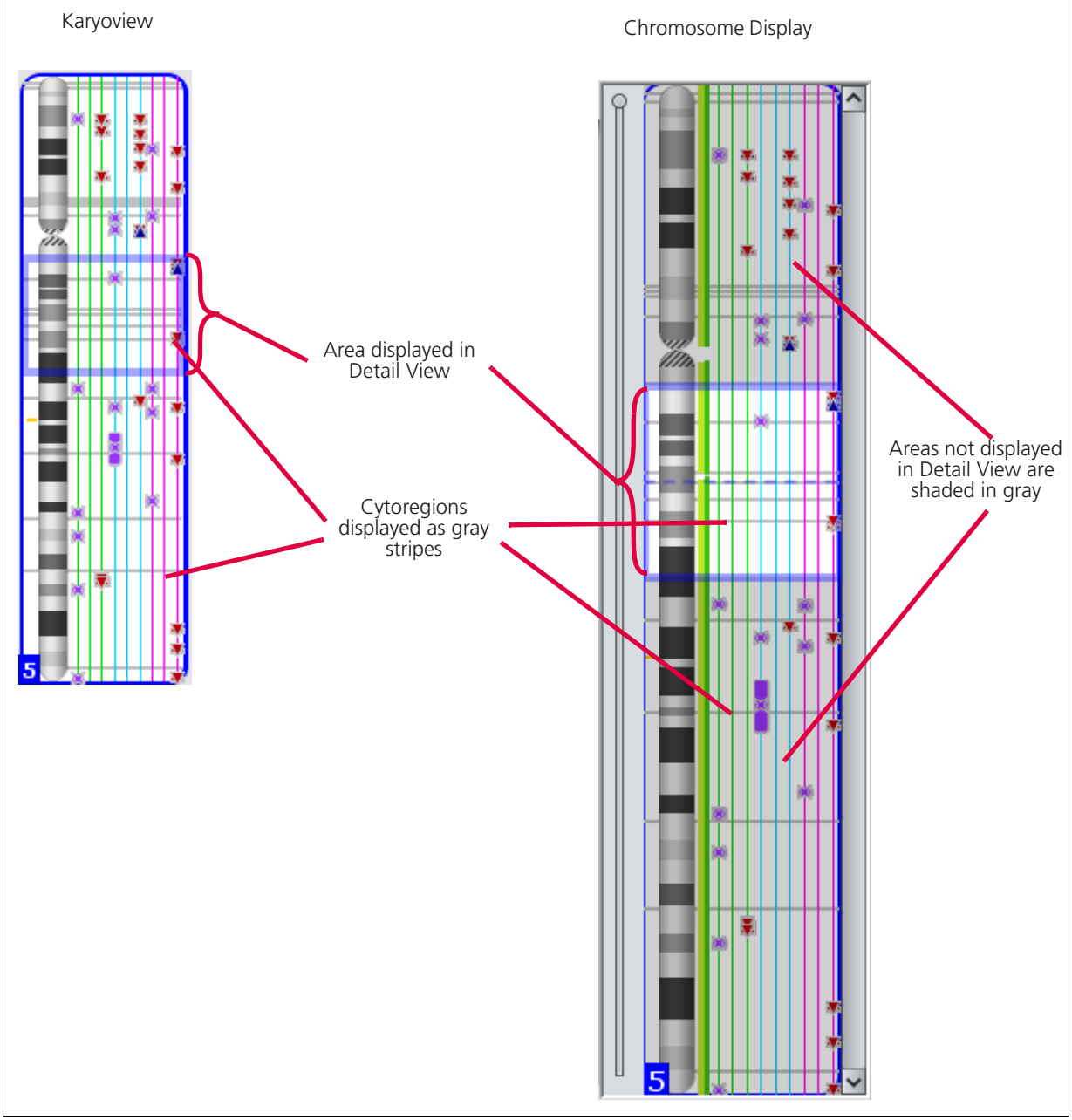
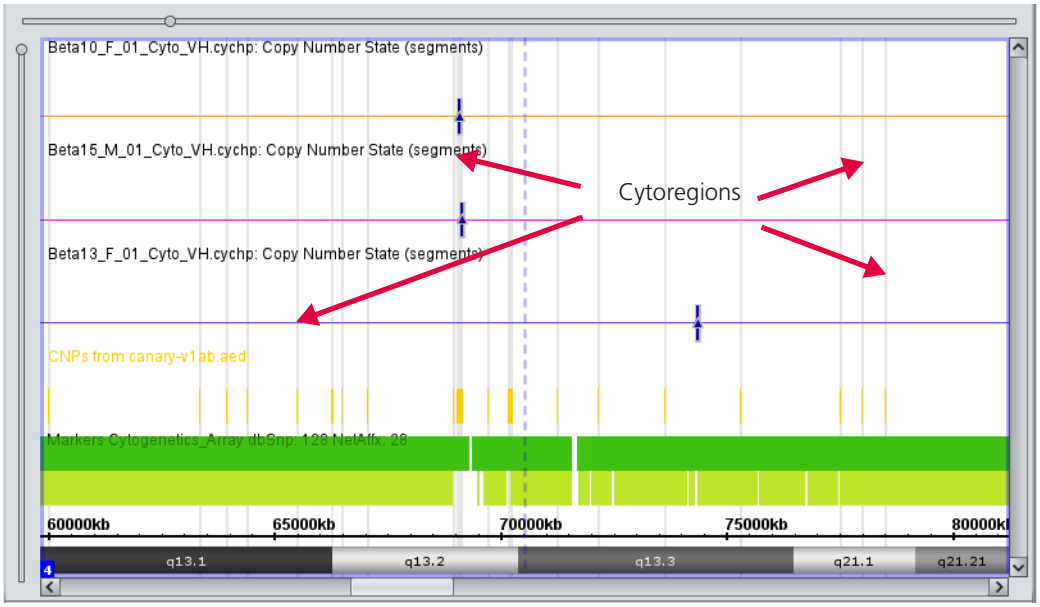


Figure 10.5 Detail View with CytoRegions displayed

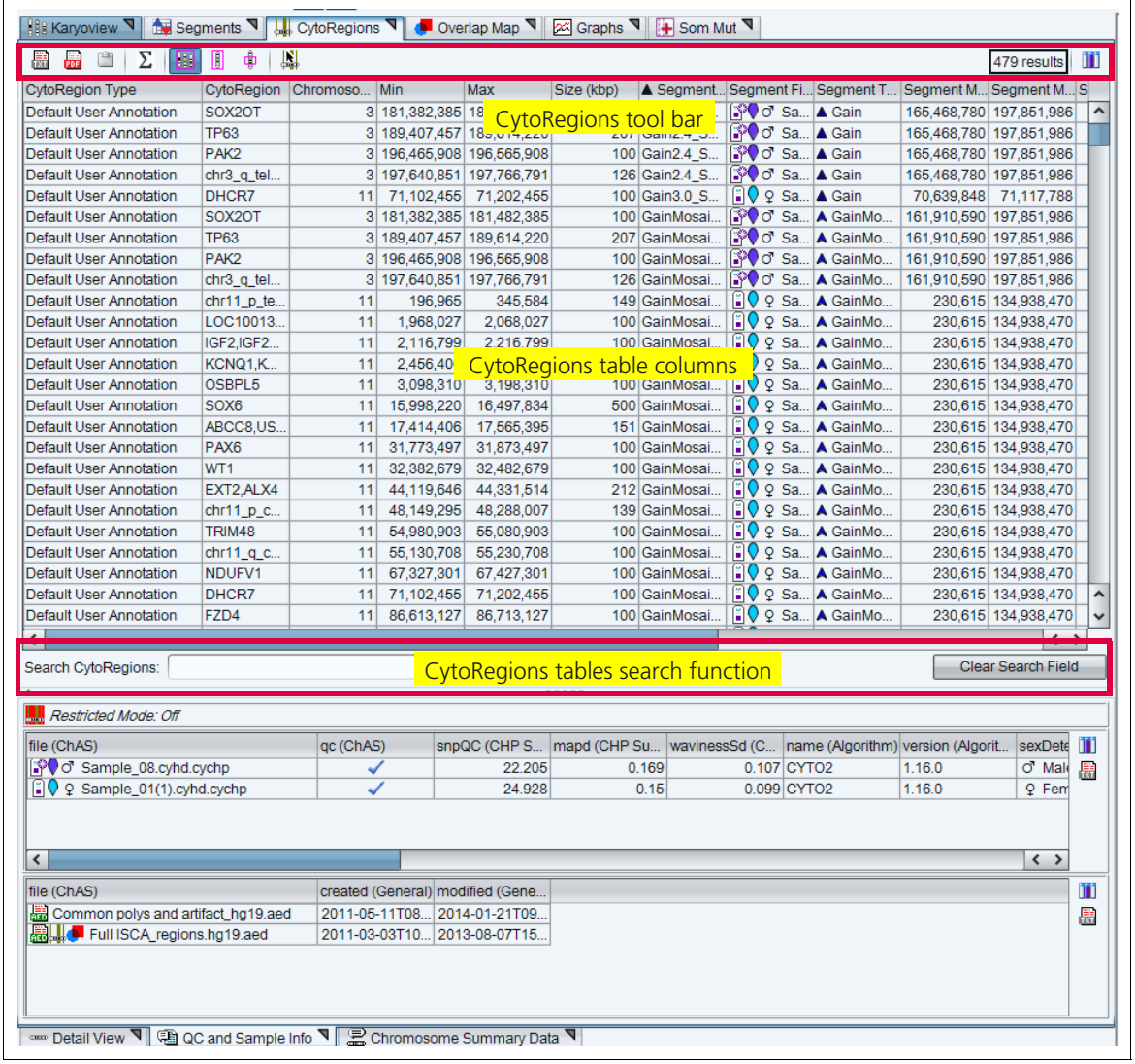


CytoRegions Table

The CytoRegions table (Figure 10.6) shows the intersections between regions in the designated CytoRegions file and the segments in the Segments table. Every region in the CytoRegions file will be listed on at least one line of the CytoRegions table, even if it does not intersect any segments. For those regions which intersect one or more segments, there will be one table row for each intersection. Depending on the columns which have been used to sort the CytoRegions table, these rows may or may not be near each other. A segment that intersects more than one region in the CytoRegions file appears multiple times in the CytoRegions table, one row for each intersection.

In the CytoRegions table, the “CN State” value corresponds to the state of copy-number and mosaicism segments that intersect the cytoregion. There is no CN State value for other types of segments that do not correspond to a copy-number call.

Figure 10.6 CytoRegions Table



Highlighting Regions in the CytoRegions Table and Details View

If the Details View displays the CytoRegions file (the CytoRegions file is check marked in the Files list), you can conveniently find and view items.

- Click a row in the CytoRegions table to select the corresponding annotation from the CytoRegions file. All of the lines for that region are highlighted in the table. The Details View zooms to the currently selected region.

> **NOTE:** Note: If the Details View does not automatically zoom to the selected region, confirm that the Auto-zoom option is selected (select "**View** → Auto-zoom to table selection" on the menu bar.)

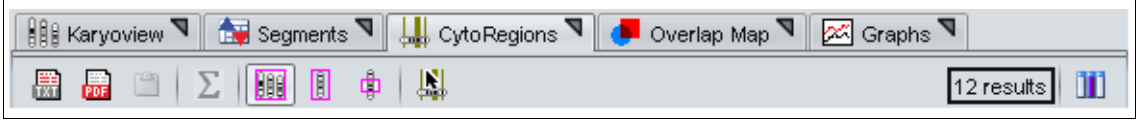
- In the Detail View, click a region or select multiple regions of the CytoRegion file to highlight all of the corresponding rows in the CytoRegions table. The CytoRegions table automatically scrolls to show at least one of the highlighted rows.

- To quickly find a particular Segment in the CytoRegions table, first double-click that segment in any of the views or in the Segments table (the current region will be set to that segment), then press the tool bar button in the CytoRegions table to show only the cytoregions in the current region.

CytoRegions tool bar

The tool bar (Figure 10.7) provides quick access to table functions. The standard functions are described in *The Standard Tool Bar Controls* (on page 228).

Figure 10.7 CytoRegions table tool bar



The tool bar has one specialized button.

Select CytoRegions file (see *Selecting a CytoRegions Information File* (on page 178)).

CytoRegions Table Columns

The table can display the following columns:

CytoRegion Type	Type of file or element from which the CytoRegion is derived. Default User Annotations are annotations derived from AED or BED file.
CytoRegion	Identifier for region.
Chromosome	Chromosome in which the region is located.
Min	Starting position of the region.
Max	Ending position of the region.
Size (kbp)	Size of the region.
Segment Label	Label assigned to the detected segment by ChAS.
Segment Name/ID	Name/ID of the Segment.
Segment File	Sample File that the segment was detected in.
Segment Type	Type of segment: CN loss or gain Mosaicism LOH
Segment Min	Starting position of segment.
Segment Max	Ending position of segment. For all segments, the segment start coordinates are always lower by one bp from the coordinate for the starting probe of the segment as reported in the graphs table while the end coordinate matches the coordinate for the ending probe as reported in the graphs table (see Appendix E, *Genomic Position Coordinates* on page 313).

Segment Size (kbp)	Size of the segment.
CN State	Copy Number State (Displayed for Gain, Loss, and Mosaicism segment types).
Shared Size	Size of the contact between segment and cytoregion.
% of CytoRegion touching Segment	How much of the CytoRegion is contacted by the Segment.
% of Segment touching CytoRegion	How much of the detected segment is contacted by the CytoRegion.

You can:

- Use the common table functions to control the display of data in the table. *See Common Table Operations (on page 228)*
- Export data in various formats. *See Exporting Table Data (on page 274)*.
- Search CytoRegions to display only regions of interest (see below).

Searching CytoRegions

The Search CytoRegions feature allows you to search the "CytoRegion" column for text strings that match a search string. You can use * as a wild card.

To search cytoregions:

1. Enter the search string in the Search CytoRegions box at the bottom of the CytoRegions table.
2. Press **Enter**.

   The table displays only annotations that match the search string.
3. To restore the table, click **Clear Search Field**.

Using Filters with CytoRegions

If you have CytoRegions information loaded into ChAS, you have the option to apply different segment filtering parameters to the parts of the genome that are defined as CytoRegions and the parts of the genome that are not within these regions (Genome). For segment filtering to function as described below, the CytoRegion segment filters should always be set identical to or less stringent than the counterpart Genome segment filter settings.

The filter settings for CytoRegions and for the rest of the genome interact in different ways, depending upon whether:

- A cytoregion information file is selected
- Restricted Mode is selected

If a cytoregion information file is selected, but Restricted Mode is not selected:

- Segments wholly within a cytoregion are filtered using the CytoRegions filter settings.
- Segments that don't touch a cytoregion are filtered using the Whole Genome settings.
- A segment that touches both Genome and CytoRegions must pass both the CytoRegions filter thresholds and Whole Genome filter thresholds, otherwise it will not be shown.

If Restricted Mode is selected:

- Segments in or straddling cytoregion boundaries are filtered using the CytoRegions filter settings.

These rules also apply when a region information file has been selected as an Overlap Map and the Overlap filters are used.

Using Restricted Mode

The Restricted mode allows you to view detected segments and graph data only in regions you have defined in advanced in the CytoRegions file.

NOTE: Restricted mode is not available unless a region information file or one of the Reference Annotations is selected for the CytoRegions function.

To select/deselect Restricted Mode:

- From the **View** menu, select **Restrict to CytoRegion**; or

  Click the Restricted Mode button in the main tool bar.

  Segments and graph data that are not in the defined cytoregions are concealed in the display.

Segments in or straddling cytoregions boundaries are filtered using the CytoRegions filter settings (Figure 10.8, Figure 10.9).

Figure 10.8 Karyoview in unrestricted and restricted mode

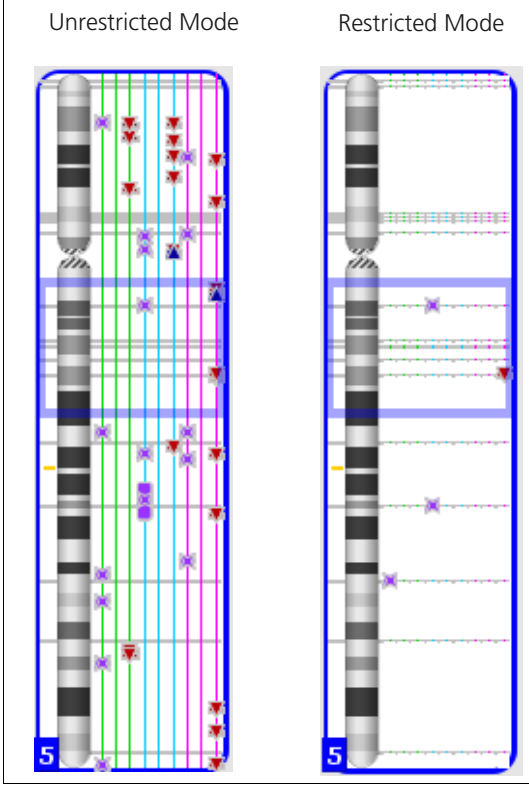
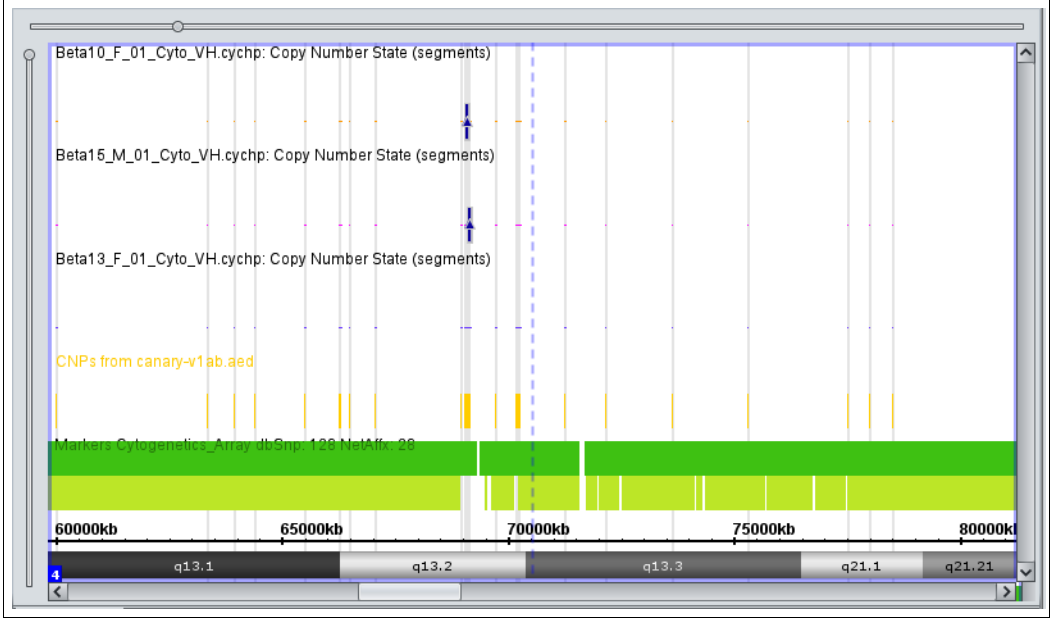
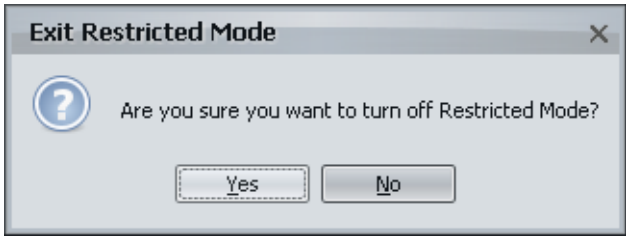


Figure 10.9 Detail View in Restricted Mode



When you deselect Restricted Mode, the following dialog box appears (Figure 10.10):

Figure 10.10 Exit Restricted Mode notice



Click **Yes** to exit Restricted Mode

Using the Overlap Map

The Overlap Map allows you to show or hide segments in areas of the genome that you aren't interested in, for example, in regions of Benign Copy Number Polymorphism, by:

- Specifying regions of the genome in an Overlap Map File.
- Optionally filtering out detected segments that are overlapped by these regions.

  You can specify the percentage of the segment (between 1 and 100%) that must be overlapped by the Overlap Map items before being filtered from display.

The Overlap Map filter operates separately from the CytoRegions features, but you can apply different overlap map filtering parameters to CytoRegions and to areas outside of CytoRegions.

To use the overlap filter:

1. Select a region information file for the Overlap Map (see *Selecting the Overlap Map File*, below).
2. Set Percent Overlap value for each segment type you want to filter in the Segment Filters dialog box See *Setting the Overlap Filter*(on page 194).
3. View the Overlap Map regions in:

- *Viewing Overlap Map Regions in the Graphic Displays*(on page 190)
- *Viewing the Overlap Map Table*(on page 192)

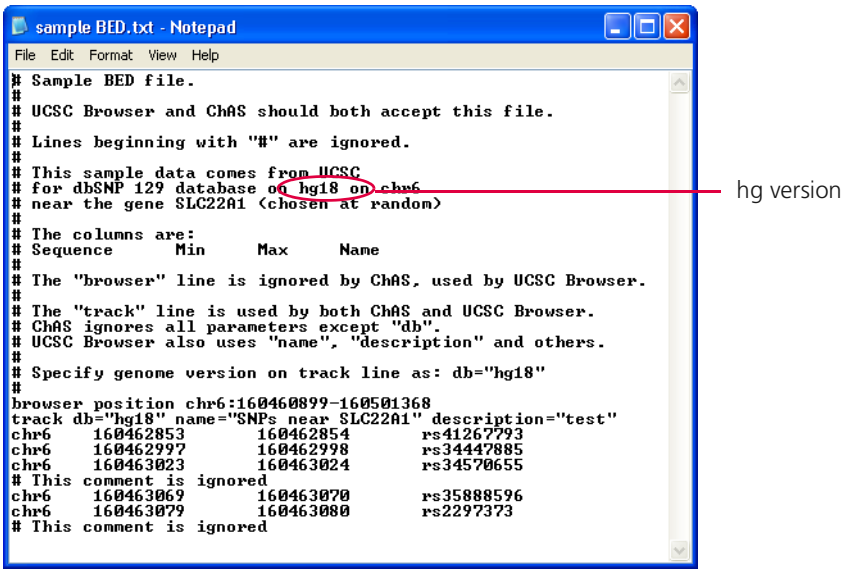
Selecting the Overlap Map File

You can use the following file types for the Overlap Map:

- Region information files in AED and BED format.
- Position information from Reference annotation files.

The software automatically checks the hg version of an AED or BED file before loading (see Figure 11.1 for an example BED file). The file will not be loaded if the hg version does not match what is loaded in the ChAS Browser. If an hg version is not found for the AED or BED file, a warning message is displayed.

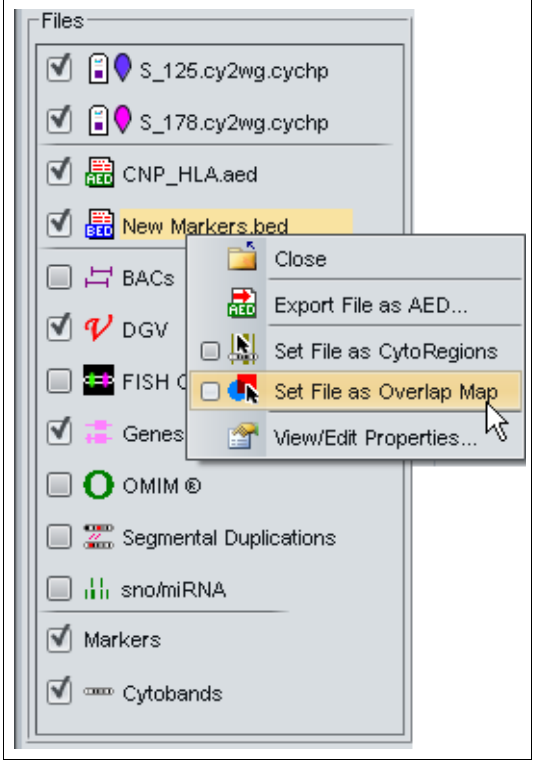
Figure 11.1 Example BED file



Do either of the following to select an overlap map:

- In the files list, right-click the file and select **Set File as Overlap Map** on the shortcut menu (Figure 11.2).

Figure 11.2 Select an Overlap Map from the files list



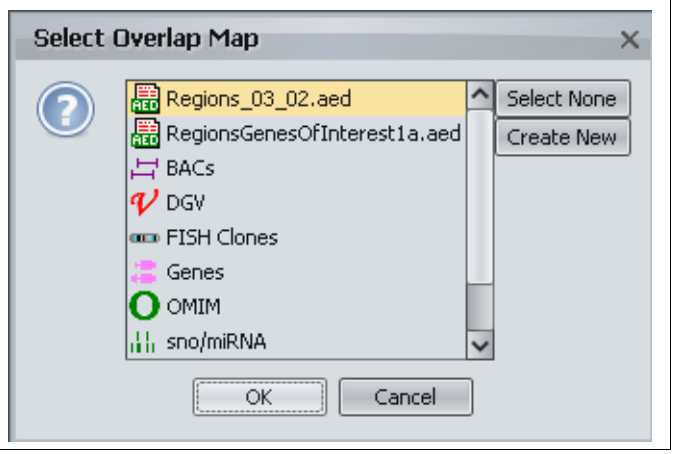
OR

1. **Select View → Overlap Map** on the menu bar.

Alternatively, in the Overlap Map tab in the upper display, click the Select Overlap Map tool bar button.

The Select Overlap Map dialog box appears (Figure 11.3).

Figure 11.3 Select Overlap Map dialog box

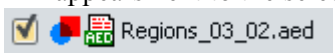


The dialog box displays a list of the region and annotation files you can select for an overlap map.

2. Select the file you want to use and click **OK**.

   The region file is loaded and the regions are displayed with overlap information in the Overlap Map tab.

   The Overlap Map icon appears next to the selected file in the File List.



NOTE: To clear an Overlap map, click "Select None" in the dialog box (Figure 10.2). Alternatively, right-click the file in the files list, and select "Set File as Overlap Map" on the shortcut menu.

Viewing Overlap Regions

You can view the Overlap Map regions:

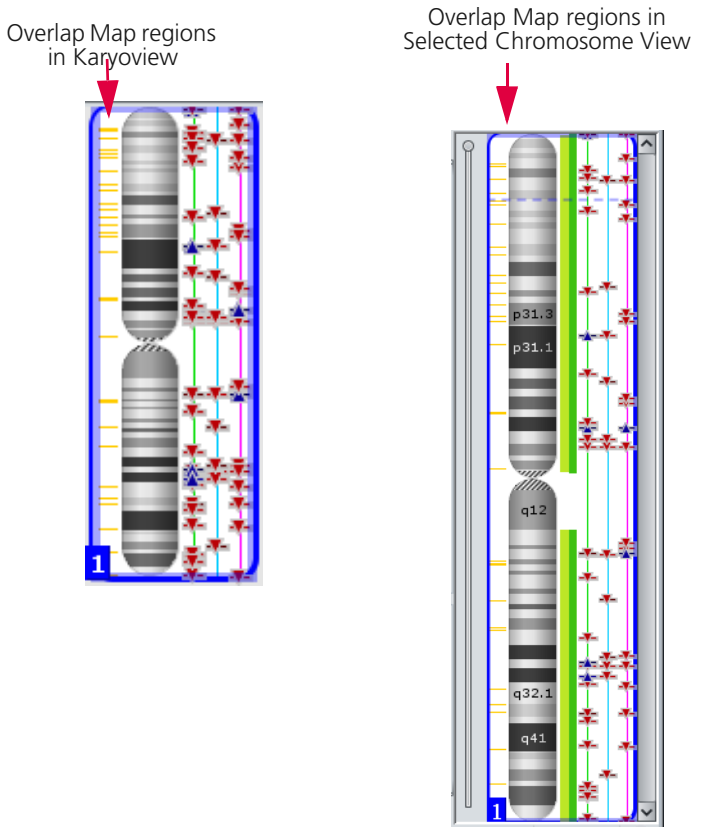
- In the Graphic Display views (desribed below)
- In the Overlap Map Tab (on page 192)

NOTE: Overlap Map items that are covered by a detected segment are listed in a column of the Segments table. See *Segments Table*(on page 237).

Viewing Overlap Map Regions in the Graphic Displays

In the Karyoview and Selected Chromosome View, regions specified in the Overlap Map file are displayed as short rectangles to the immediate left of the cytobands (Figure 11.4 on page 191).

Figure 11.4 Karyoview and Selected Chromosome View with Overlap Map Regions Displayed

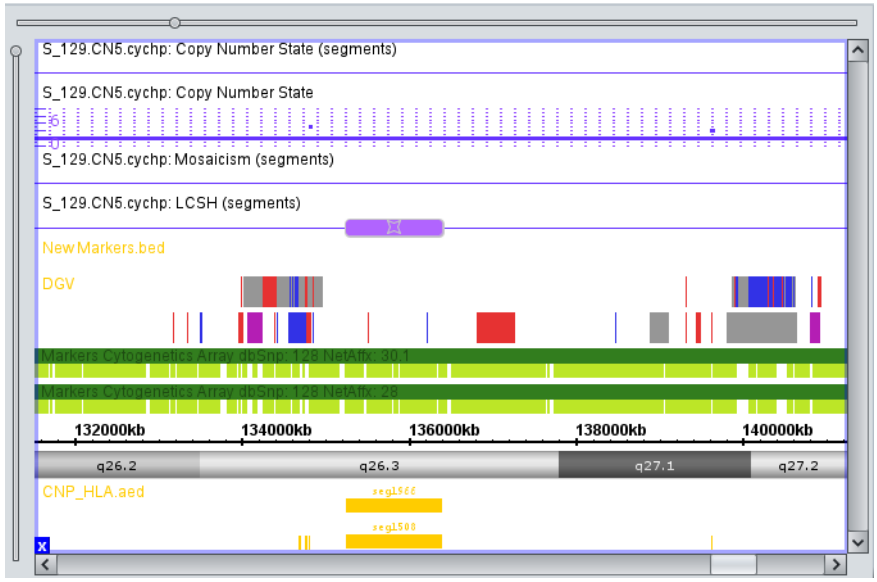


In the Detail View, the overlap map regions are displayed as rectangles below the cytobands (Figure 11.5). The default color is yellow, but you may select different colors for displaying regions in a region information file.

You can change the colors used to display items by using:

- Specifying color when adding the item to an AED file. *See Entering General Information (on page 208)*.
- Using the Color Rules. *See AED/BED Color Rules (on page 219)*.

Figure 11.5 Detail View with Overlap Regions displayed



Overlap Map regions

Viewing the Overlap Map Table

The Overlap Map table (Figure 11.6 on page 193) displays a list of overlapping items from the overlap map file and the Segments table. Each region in the Overlap Map file will be listed on at least one row of the table, even if it does not overlap any segments. For those regions which overlap one or more segments, there will be one for each overlap. Depending on the columns which have been used to sort the Overlap Map table, these rows may or may not be near each other. A segment that overlaps more than one region in the Overlap Map file will appear multiple times in the Overlap Map table, one row for each overlap.

Highlighting Overlap Regions in the Overlap Map Table and Details View

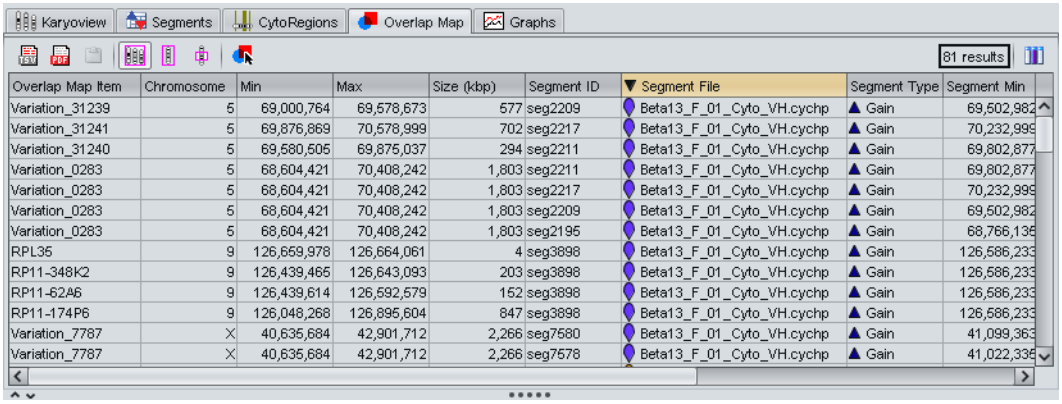
If the Details View displays the Overlap Map file (the Overlap Map file is check marked in the Files list), you can conveniently find and view items.

- Click a row in the Overlap Map table to select the corresponding annotation from the Overlap Map file. All of the rows for that region are highlighted in the table. The Details View zooms to the currently selected region.

NOTE: If the Details View does not automatically zoom to the selected region, confirm that the Auto-zoom option is selected (select "View → Auto-zoom to table selection" on the menu bar.)

- In the Detail View, click a region or select multiple regions from the Overlap Map file to highlight all of the corresponding rows in the Overlap Map table. The Overlap Map table automatically scrolls to show at least one of the highlighted rows.
- To quickly find a particular Segment in the Overlap Map table, first double-click that segment in any of the views or in the Segments table (the current region will be set to that segment), then press the tool bar button in the Overlap Map table to show only the overlaps in the current region.

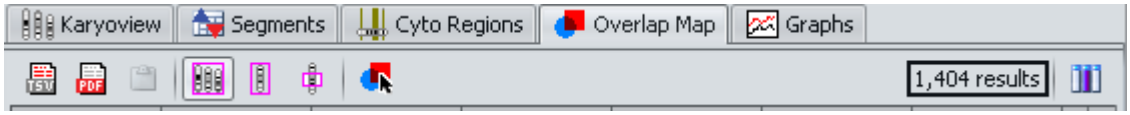
Figure 11.6 Overlap Map Table



Overlap Map Item	Chromosome	Min	Max	Size (kbp)	Segment ID	Segment File	Segment Type	Segment Min
Variation_31239	5	69,000,764	69,578,673	577	seg2209	Beta13_F_01_Cyto_VH.cychp	Gain	69,502,982
Variation_31241	5	69,876,869	70,578,999	702	seg2217	Beta13_F_01_Cyto_VH.cychp	Gain	70,232,999
Variation_31240	5	69,580,505	69,875,037	294	seg2211	Beta13_F_01_Cyto_VH.cychp	Gain	69,802,877
Variation_0283	5	68,604,421	70,408,242	1,803	seg2211	Beta13_F_01_Cyto_VH.cychp	Gain	69,802,877
Variation_0283	5	68,604,421	70,408,242	1,803	seg2217	Beta13_F_01_Cyto_VH.cychp	Gain	70,232,999
Variation_0283	5	68,604,421	70,408,242	1,803	seg2209	Beta13_F_01_Cyto_VH.cychp	Gain	69,502,982
Variation_0283	5	68,604,421	70,408,242	1,803	seg2195	Beta13_F_01_Cyto_VH.cychp	Gain	68,766,136
RPL35	9	126,659,978	126,664,061	4	seg3898	Beta13_F_01_Cyto_VH.cychp	Gain	126,586,233
RP11-348K2	9	126,439,465	126,643,093	203	seg3898	Beta13_F_01_Cyto_VH.cychp	Gain	126,586,233
RP11-62A6	9	126,439,614	126,592,579	152	seg3898	Beta13_F_01_Cyto_VH.cychp	Gain	126,586,233
RP11-174P6	9	126,048,268	126,895,604	847	seg3898	Beta13_F_01_Cyto_VH.cychp	Gain	126,586,233
Variation_7787	X	40,635,684	42,901,712	2,266	seg7580	Beta13_F_01_Cyto_VH.cychp	Gain	41,099,363
Variation_7787	X	40,635,684	42,901,712	2,266	seg7578	Beta13_F_01_Cyto_VH.cychp	Gain	41,022,335

Overlap Map tool bar

The tool bar provides quick access to table functions. The standard functions are described in *The Standard Tool Bar Controls*(on page 228).



The tool bar has one specialized button.

Select Overlap Map file (see *Selecting the Overlap Map File*(on page 188)).

Overlap Map Table Columns

The table can display the following columns:

Map Item Type Source of the position information (CN Gain or Loss segment, reference annotation, etc.)

Overlap Map Item Identifier for the item.

Chromosome Chromosome in which the item is located.

Min Starting position of the item.

Max Ending position of the item.

Size (kbp) Size of the item.

Segment Label Label assigned to the detected segment by ChAS.

Segment Name/ID Name/ID of the Segment.

Segment File Sample File that the segment was detected in.

Segment Type Type of segment:
- CN loss or gain
- Mosaicism
- LOH

Segment Min Starting position of segment.

Segment Max	Ending position of segment. For all segments, the segment start coordinates are always lower by one bp from the coordinate for the starting probe of the segment as reported in the graphs table while the end coordinate matches the coordinate for the ending probe as reported in the graphs table (see Appendix E, *Genomic Position Coordinates* on page 313).
Segment Size	Size of the segment.
CN State	Copy Number State (Not displayed for other segment types).
% Overlap	How much of the detected segment is covered by the Overlap Map item. A Segment which has 20% of its length somehow encompassed within the boundaries of an Overlap Map item has an Overlap value of 20%. This percentage value is used to filter segments out of the displays and tables when filtering segments by "Overlap" in the filter slider dialogs.
% Coverage	How much of the Overlap Map item is covered by the Segment.
Shared Size	Size of the overlap between segment and Overlap Map item. Coverage values are not presently used in filtration of Segments from the displays or tables.

Setting the Overlap Filter

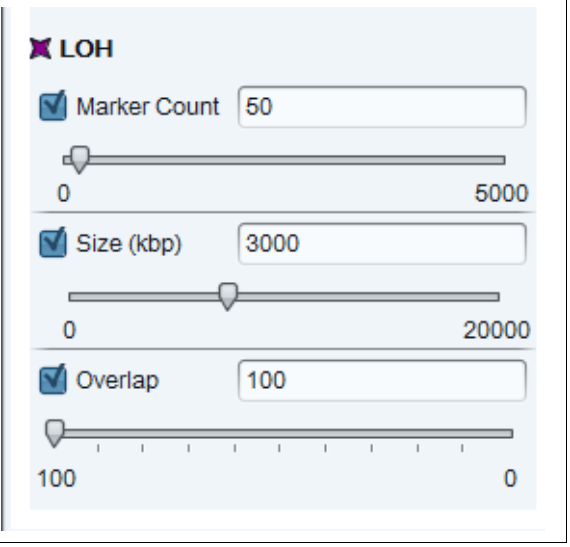
After selecting an Overlap Map Region file, you have the option to set Overlap filters for the different segment types.

To use the Overlap filters:

1. Select a region information or Reference Annotation file for the Overlap Map. *See Selecting the Overlap Map File (on page 188)*.
2. Open the Segment Filters dialog box (click the button) (Figure 11.7).

   See Applying Segment Parameter Filters (on page 142).

Figure 11.7 LOH filter settings, Overlap selected



3. Select the Overlap check box(es) for the segment types you want to use to filter.
4. Set the value for the parameters with the slider or enter a value in the box.

- Moving the slider farther to the right (or entering smaller values in the box) will cause more and more of the Overlapped segments to be removed from the display.

- The detected segments must share at least the specified percentage of their length with the Overlap Map region to be filtered out and hidden from display. A Segment which has 20% of its length somehow encompassed within the boundaries of an Overlap Map item has an Overlap value of 20%.
- The minimum value of the setting is 1%.
- The results of filtering are seen instantly in all tables and graphs.
- *See Using Filters with CytoRegions (on page 185)* for information about the interactions of the Overlap Map filter with the CytoRegions.

Creating and Editing AED Files

You can create AED files that contain:

- User-selected annotations. The annotations in an AED annotation file can be edited in ChAS.
- Position information on regions of the chromosome, as well as additional annotation information on the regions. AED region information files can be used for CytoRegions or Overlap Map functions.
- Third party reference annotations converted to AED file format

User-generated AED region information files can be used for CytoRegions or Overlap Map functions. You can add a region to a new or previously created AED file by selecting the following feature types:

- Detected Segments
- Reference Annotations
- Regions in previously loaded files

This chapter discusses:

- *Creating an AED File of Annotations* (on page 196)
- *Creating an AED File of Regions* (on page 198)
- *Viewing or Editing AED File Properties* (on page 203)
- *Viewing and Editing Annotations* (on page 207)
- *Viewing and Batch Editing AED File Contents* (on page 216)
- *AED/BED Color Rules* (on page 219)
- *AED/BED Color Rules* (on page 219)
- *Exporting Information in AED or BED Format* (on page 224)

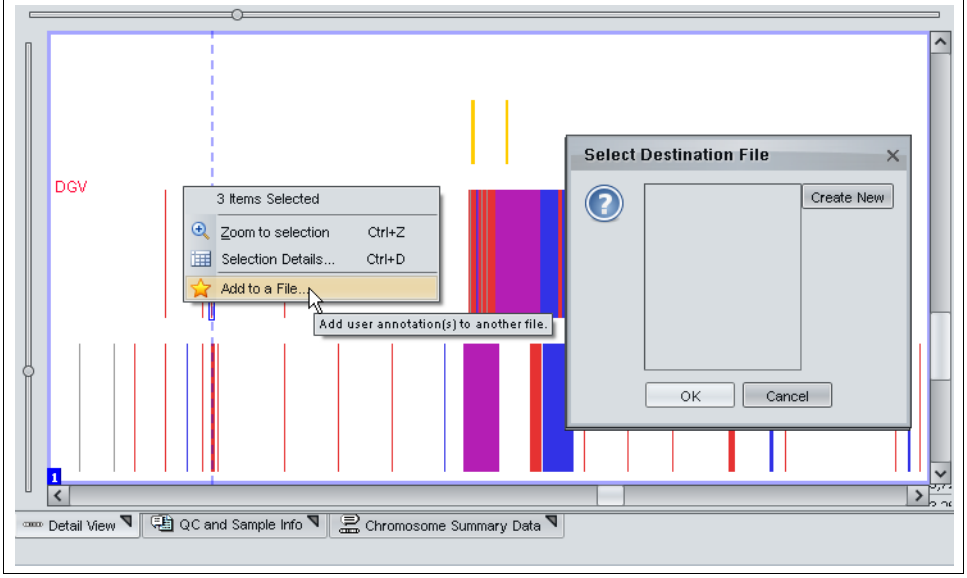
The AED file format is described in *Appendix B, AED File Format* (on page 298).

Creating an AED File of Annotations

1. In the Detail View, select the non-AED annotation(s) for the AED file:
   - Right-click an annotation and select **Add to a File** on the shortcut menu (Figure 12.1).

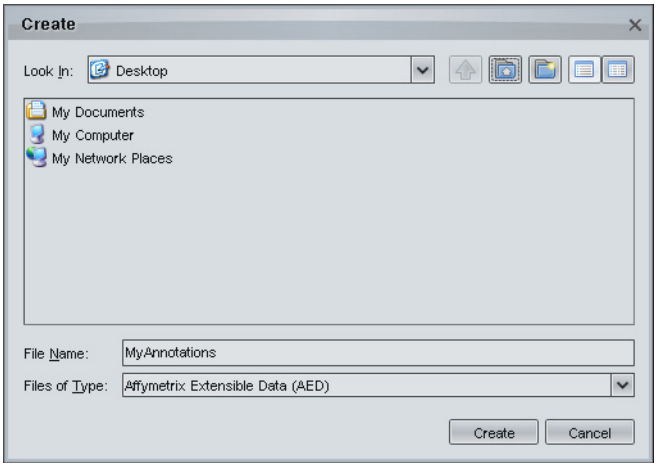
     Or
   - Draw a box around multiple annotations, right-click the selection, and select **Add to a File** on the shortcut menu.
2. In the dialog box that appears (Figure 12.1), click **Create New**.

Figure 12.1 Detail View, create a new AED annotation file



3. In the Create dialog box that appears next (Figure 12.2), select a folder and enter a file name.

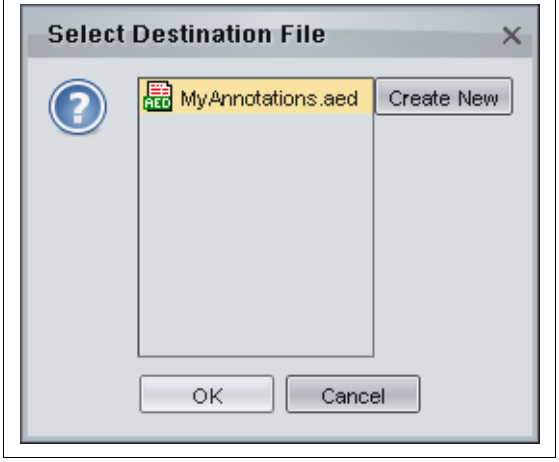
Figure 12.2 Create dialog box



4. Click **Create**.

   The Select Destination File dialog box (Figure 12.3) shows the name of the new AED file.

Figure 12.3 Select Destination File dialog box



5. Click **OK**.

   The Details View shows the new annotation (AED).

> **NOTE:** The AED file is automatically assigned the same genome assembly version (i.e., "hg18", "hg19", etc.) as the currently loaded NetAffx annotations.

Adding Annotations to an AED File

1. In the Detail View, select the annotation(s) that you want to add to an AED file:
   - Right-click an annotation and select **Add to a File** on the shortcut menu

     OR
   - Draw a box around multiple annotations, right-click the selection, and select **Add to a File** on the shortcut menu
2. In the Select Destination File dialog box (Figure 11.3), select an AED file and click **OK**.

> **NOTE:** Adding annotations to an AED file does not modify the genome assembly version. If the AED file does not specify a genome assembly version, none is automatically assigned to the AED file when annotations are added.

Creating an AED File of Regions

You may want to incorporate new segments of features from a set of samples in a new Regions file.

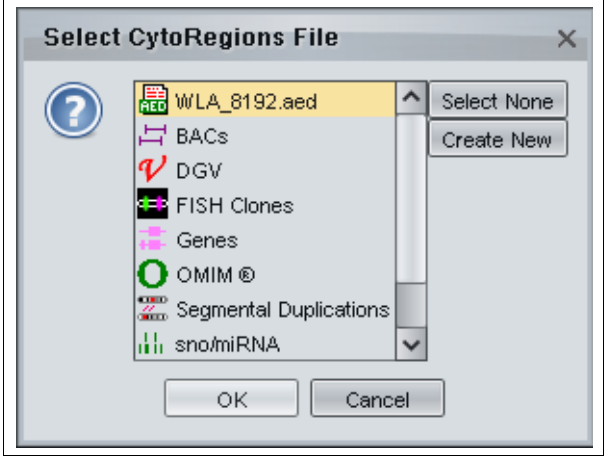
> **NOTE:** You can also use the Export feature to export data in existing files to an AED file. See *Exporting Information in AED or BED Format* (on page 224).

To create a new CytoRegions file in AED file format:

1. From the View menu, choose **Select CytoRegions** file or **Select Overlap Map**.

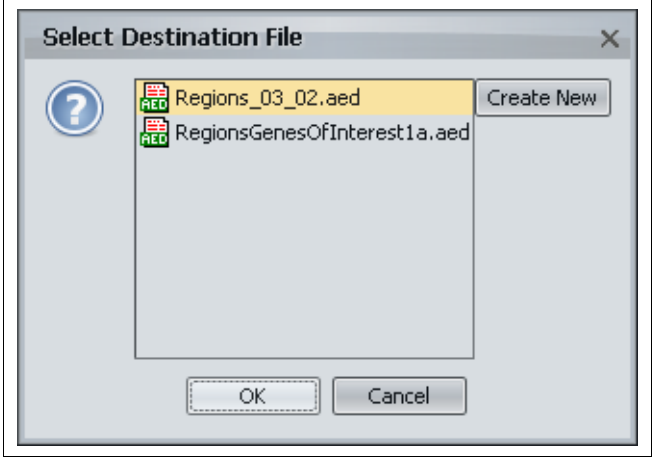
   The appropriate Select File dialog box opens (Figure 12.4).

Figure 12.4 Select CytoRegions file dialog box



NOTE: You can also create a new AED file when adding a region to an AED file.

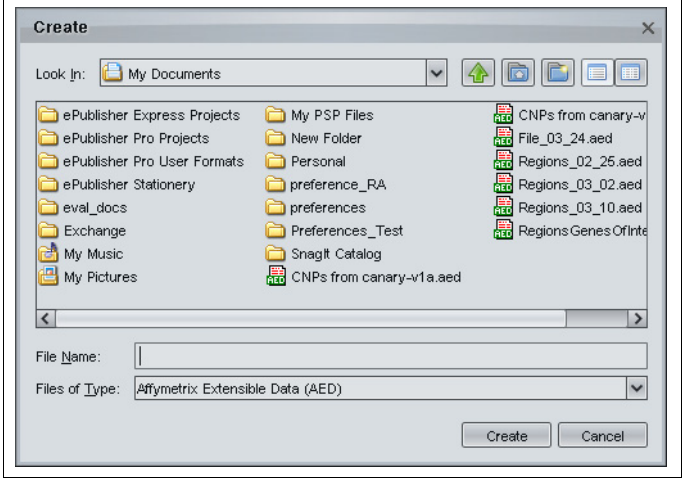
Figure 12.5 Select Destination File dialog box



2. Click **Create New** in the Select File dialog box.

   The Create dialog box opens (Figure 12.6).

Figure 12.6 Create dialog box

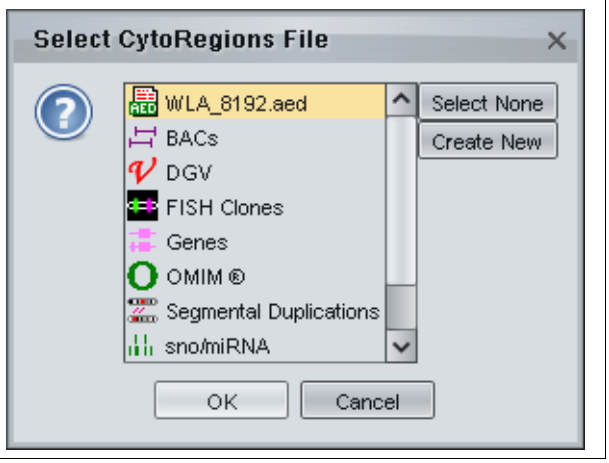


3. Use the dialog box controls to browse to a folder for the AED file.
4. Enter a file name.
5. Click **Create** in the Create dialog box.

   The Select File dialog box appears with the new file selected (Figure 12.7).

NOTE: The AED file is automatically assigned the same genome assembly version (i.e. "hg18", "hg19", etc.) as the currently loaded NetAffx annotations.

Figure 12.7 Select Destination File



You can select the new file or add regions to it, depending upon what function you were performing initially.

NOTE: To open an AED file, click the button or select File → Open on the menu bar.

Adding Regions to an Existing AED File

You can add a new region to an existing region information file in AED format by selecting the following features in the ChAS graphic views:

- Detected Segments
- Reference Annotations
- Regions in previously loaded files.

NOTE: You can edit the color of DGV annotations that have been added to an AED file by creating a color rule (see ***To create a color rule:*** **(on page 220)**). Alternatively, you can edit the color of a particular DGV annotation added to an AED file, in the Detail View (see ***Entering General Information*** **(on page 208)**).

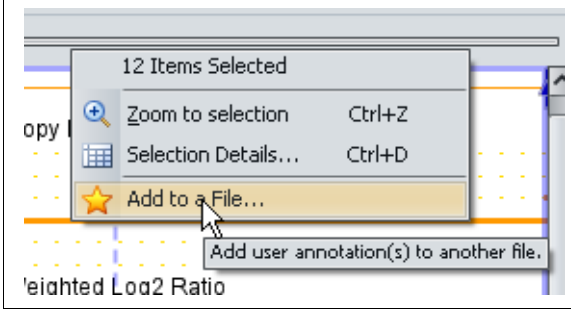
To add a section to a new region (AED) file:

1. Right-click on any of the following feature types in the graphic displays and select **Add to a File** from the right-click menu (Figure 12.8):
   - Segment
   - Reference Annotation (including Cytobands)

   NOTE: You should expand the reference annotations before selecting one to add to an AED file to avoid selecting multiple annotations (see ***Expanding and Contracting Annotations*** **(on page 124)**).

   - Region

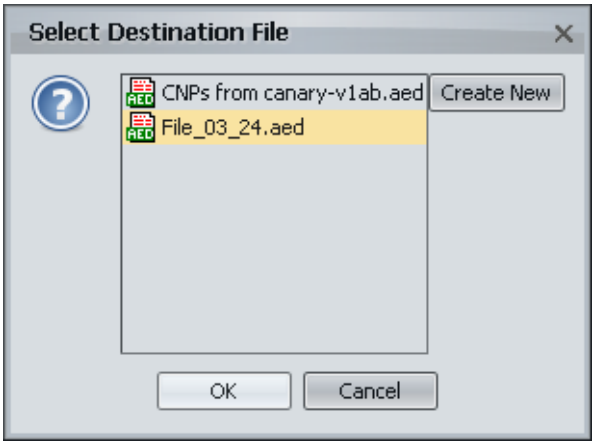
Figure 12.8 Right-click menu



The Select Destination File dialog box opens (Figure 12.9).

NOTE: Some options may not be available, depending upon the number of type of items you have selected.

Figure 12.9 Select Destination File dialog box

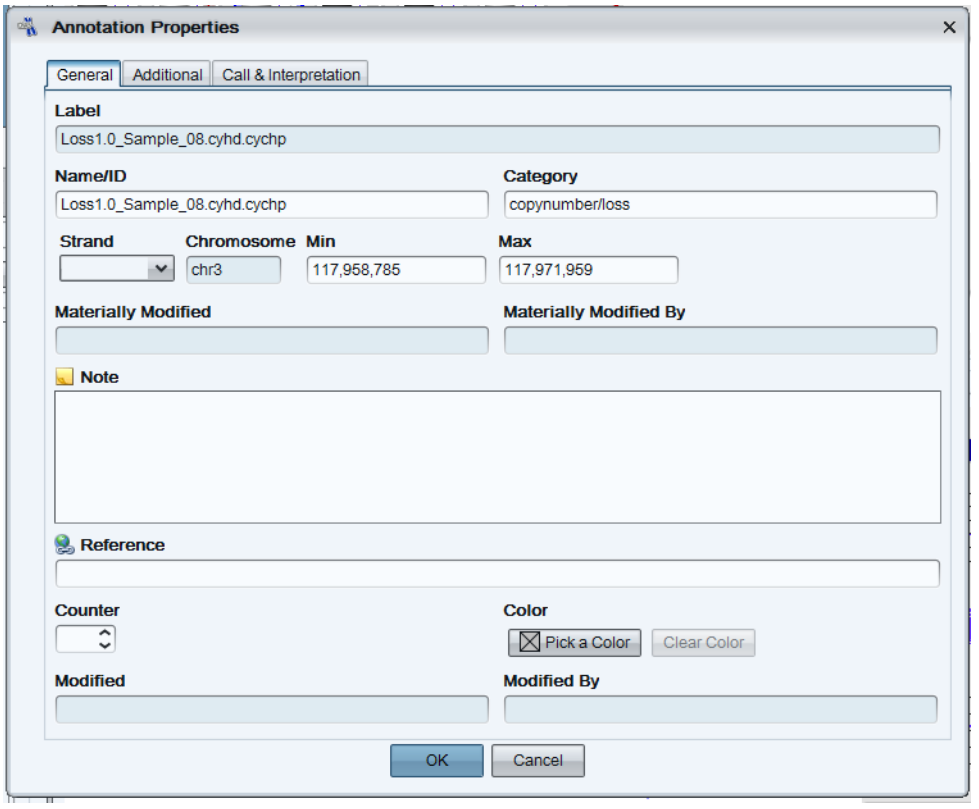


The Select Destination File dialog box displays a list of the currently existing AED files to which you may add the segment.

2. Select the region file you want to use and click **OK**.

   The Annotation Properties dialog box opens if you have selected a single item (Figure 12.10). If you have selected multiple items, the dialog box will not open.

Figure 12.10 Annotation Properties dialog box



See *Viewing and Editing Annotations* (on page 207) for information on editing the annotations.

3. After editing the annotation properties, click **OK** in the Annotation Properties dialog box.

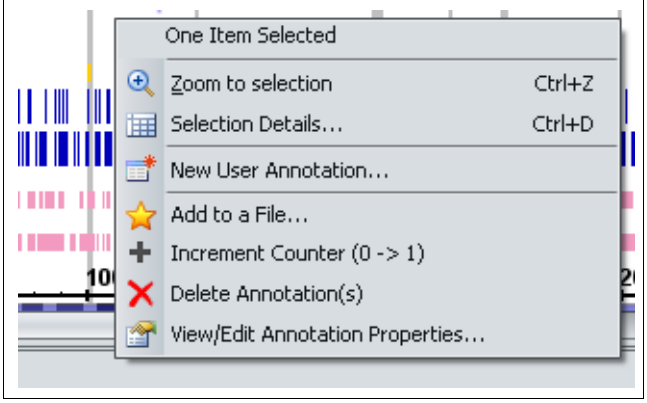
The section is saved in the Region file.

Deleting Regions from an AED File

To delete a region:

1. Right-click on a region in a region file and select **Delete Annotations(s)** (Figure 12.11).

Figure 12.11 Right-click menu for region



The region is deleted from the Region (AED) file.

Viewing or Editing AED File Properties

In ChAS, you can:

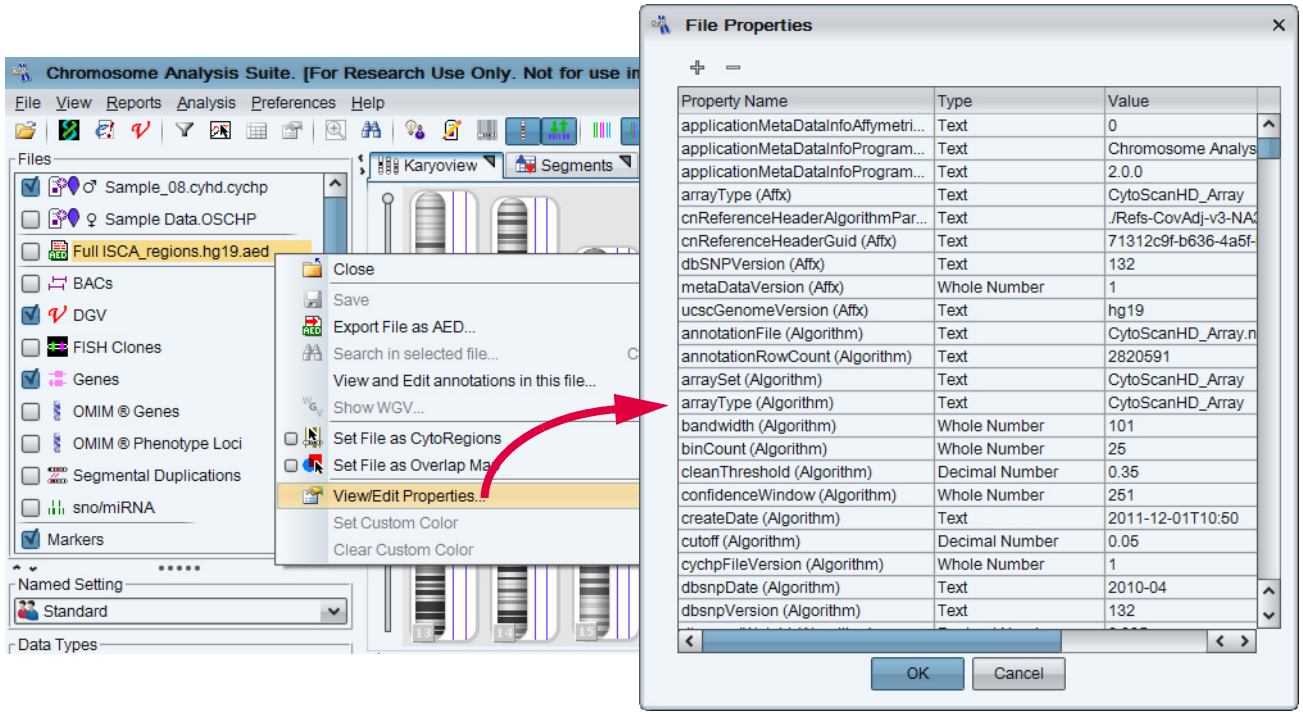
- View AED or BED file properties
- Edit AED file properties (modify, add, or delete properties)

NOTE: BED files cannot be edited in ChAS. For more information on editing BED files, see *Editing BED Files* (on page 315)

To view file properties:

1. Right-click the file and select **View/Edit Properties** on the shortcut menu.
   The Properties dialog box appears. (Figure 12.12 on page 204)

Figure 12.12 View AED or BED file properties



Viewing the Genome Assembly Version

The assigned genome assembly version of a loaded file can be viewed in the Properties box (Figure 12.12). The property, if it has been set for the file, is shown as "ucscGenomeVersion{Affx)". An AED file that is created within ChAS is automatically assigned the same genome assembly version as the loaded NetAffx annotations (for example, "hg19"). If you add annotations to an existing AED file, its genome assembly version will not be modified; and if no version is specified for the AED file, no version will be assigned to it.

NOTE: When you save a file as AED or BED, the current value of the genome assembly version property, if present, will be saved in the file. If two or more files are merged into an AED or BED file, the current value of the genome assembly version, if present in at least one of the files, will be saved in the merged file.

If an AED file does not include a genome assembly version, you can manually set it. To do this, in the Properties dialog box:

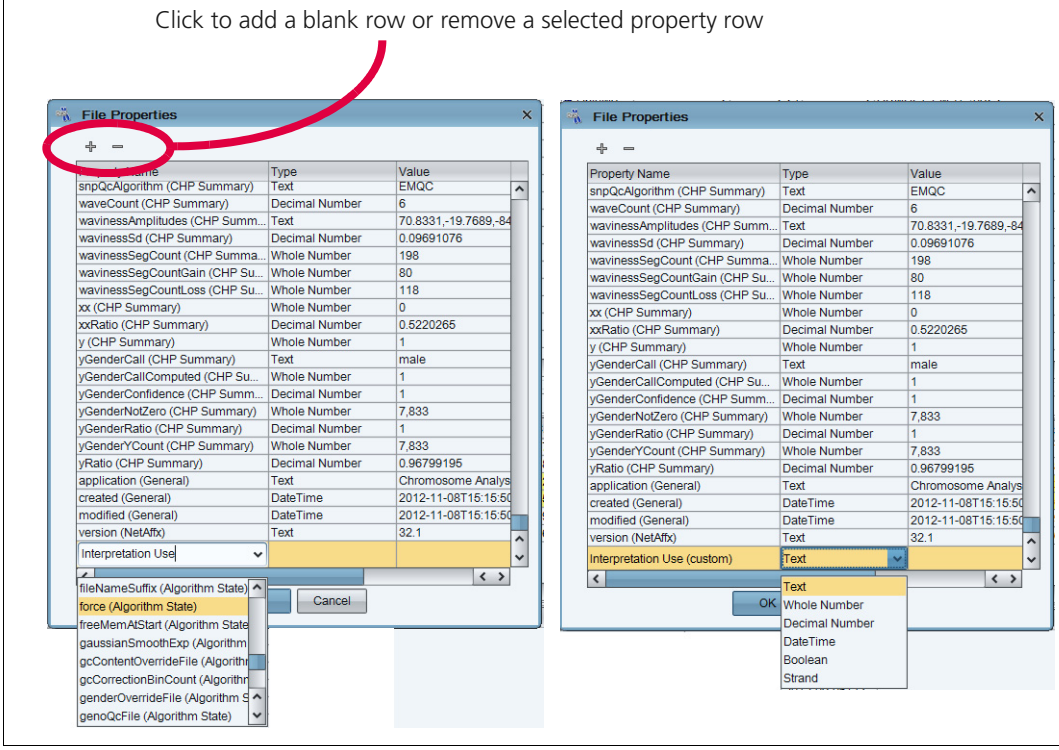
1. Add a new property row (see Figure 12.13).
2. Select the Property Name "ucscGenomeVersion(Affx)" from the drop-down list.
3. Select "Text" under the Type drop-down list.
4. In the "Value" column, enter the genome assembly version (for example, "Constitutional").

NOTE: You can manually set the genome version of an AED file by editing the "ucscGenomeVersion(Affx)" property. For more details on editing a property value, see *To edit a property value:* (on page 206)

To add a property:

1. In the Properties dialog box, click the ✚ symbol to create a blank row (Figure 12.13).
2. In the **Property Name** field, enter a property name or make a selection from the Property Name drop-down list (Figure 12.13).

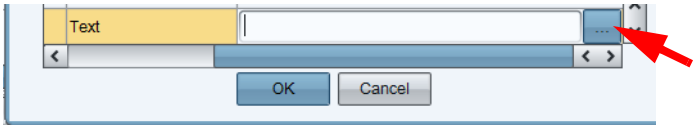
Figure 12.13 Specify a property name and type



3. Click **Type**, then make a selection from the drop-down list.
4. Click **Value**

   A cursor appears inside the Value field. (Figure 12.14).

Figure 12.14 Value field



5. Enter a value, then click **OK**. For significantly longer values, click ... to open a Value editing window. Enter your (longer) value in this window, then click **OK**.

   The new value is entered in the File Properties table.
6. Click **OK**.

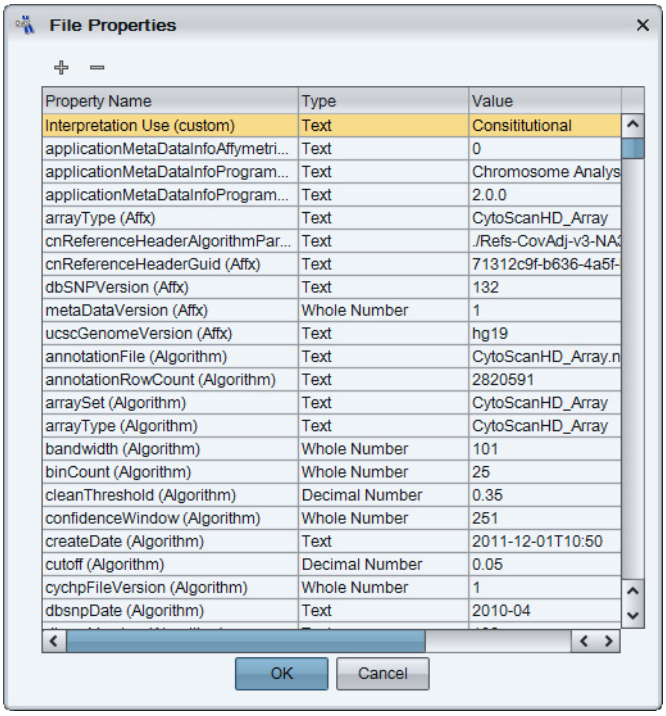
To remove a property:

1. In the Properties dialog box, select the row that you want to delete.
2. Click the ▬ icon.

To edit a property value:

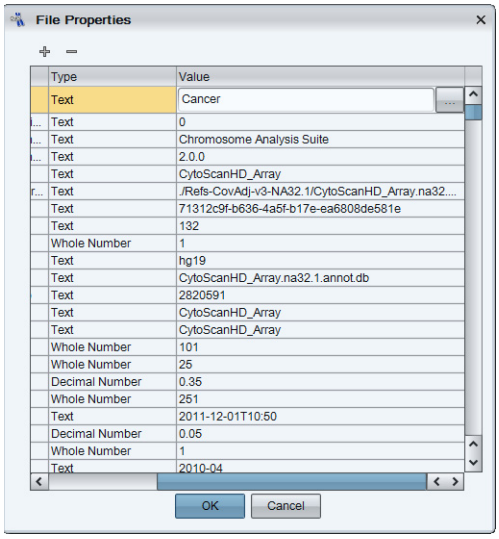
1. Right-click the file and select **View/Edit Properties** on the shortcut menu. The Properties dialog box appears.
2. Select the row in the Properties dialog box (Figure 12.15).

Figure 12.15 Select a property to edit



3. For example, click the **Value** field. (Figure 12.16).

Figure 12.16 Set a new property value



4. Enter a value, then click **OK**. For significantly longer values, click ... to open a Value editing window. Enter your (longer) value in this window, then click **OK**.

   The new value is entered in the File Properties table.

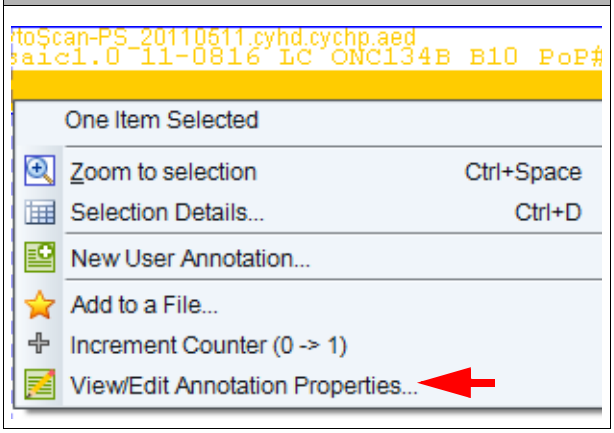
5. Click **OK**.

Viewing and Editing Annotations

The Annotation Properties box opens:

- When adding a region to a Region (AED) file
- When you select View/Edit Annotation Properties from the right-click menu (Figure 12.17) for the following types of features:
  - Detected Segments
  - Reference Annotations (view annotations only)
  - AED Annotations (view and edit annotations)

Figure 12.17 Right-click menu for View/Edit Annotation Properties in the Detail View



NOTE: The View/Edit Annotation Properties menu option is not available if you have more than one feature selected in the Detail View.

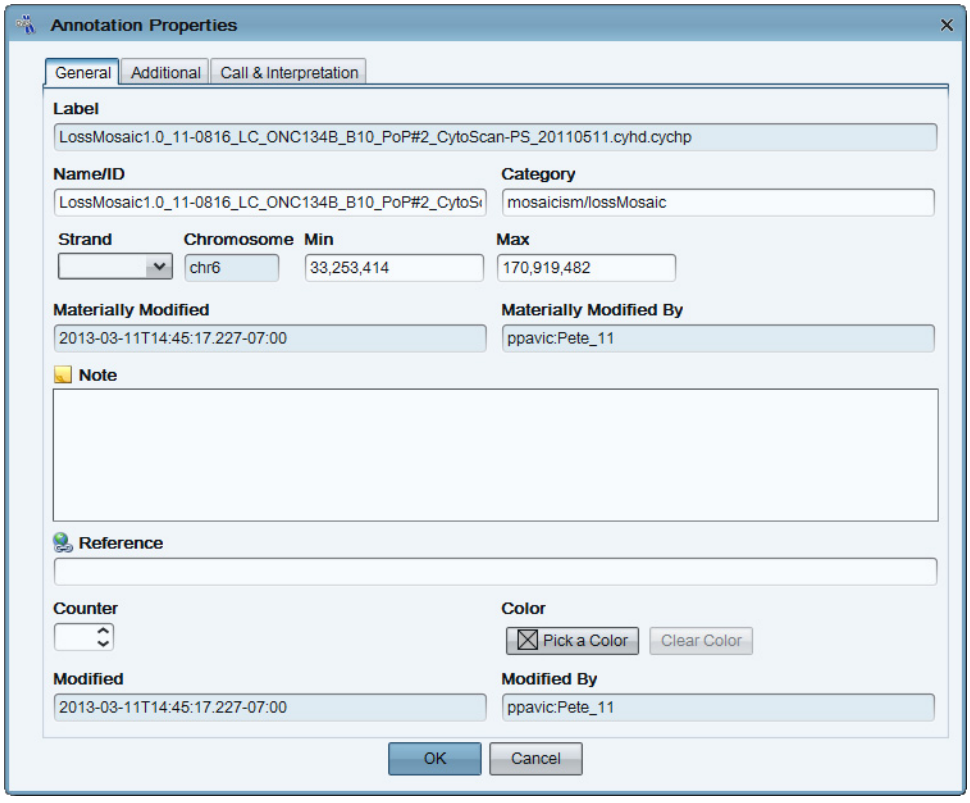
The Annotation Properties dialog box has 3 tabs:

- **General** (See *Entering General Information* (on page 208))
- **Additional** (See *Adding Properties* (on page 209))
- **Call & Interpretation** (See *Adding a Call and/or Interpretation* (on page 213))

You can also create new user annotations if you select an element. This feature allows you to create a region that is not based on a segment or reference annotation. See *New User Annotations* (on page 215).

Entering General Information

Figure 12.18 Annotation properties window of CHP Segment exported to AED



To edit General tab properties in AED annotations:

1. In the General Tab (Figure 12.18), enter, edit, or select:

Label	Label given to the AED, CHP, or other annotation. For annotations originating from CHP segments, can be the Type, State, and Filename of the CHP file. The Label is not editable.
Name/ID	Name/ID assigned to the annotation in the AED, CHP or other file.
Chromosome	Cannot be edited in Annotation Properties box. See *New User Annotations* (on page 215).
Min	The smallest of the annotation's chromosomal coordinates.
Max	The largest of the annotation's chromosomal coordinates.
Strand	The Sequence Strand of the item.
Materially Modified Time	Time stamp of when the start or end coordinate, type, or state of the segment or BED/AED annotation was last altered.
Materially Modified By	The Operating System user and ChAS User IDs of the Modifier who last changed a Material property (start or end coordinate, type, or state) of the segment or annotation.
Category	Information on the source of the region. If the region was added by selecting a CHP segment, the segment type is saved.
Note	Information and comments about the region.

Reference	URL/web address associated with the annotation.
Modified	The timestamp of the last modification of the annotation.
Modified By	The Operating System user and ChAS User IDs of the Modifier who last changed the annotation.
Counter	Allows you to track the number of times something has been seen.
Color	Allows assignment of a hard-coded color to the Annotation in ChAS's graphical views.

2. To change the color used for an individual annotation in the Detail View:

   A. Click **Choose Color.**
   The Choose Color window opens (Figure 12.19).

Figure 12.19 Choose Color window

   B. Choose a color in the Swatches tab, or click the HSB or RGB tab to define a color.
   C. Click **OK**.

NOTE: The color set in the Annotation Property dialog overrides the colors specified by a Color Rule created in the User Configuration window. For more details, see *AED/BED Color Rules* (on page 219).

Customizing Properties

Adding Properties

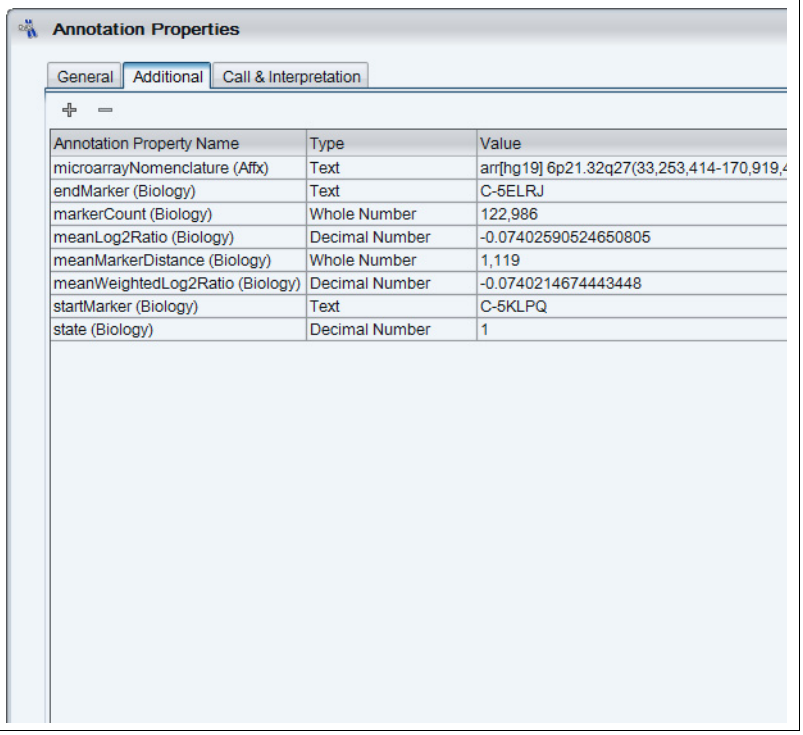
In the **Additional** tab you can enter new annotation information for an AED annotation. The information will be displayed in the:

- Tooltip when you position the mouse arrow over an annotation in the Details View
- Selection Details window

To add custom properties:

1. In the Annotation Properties window, click the **Additional** tab. (Figure 12.20 on page 210)

Figure 12.20 Additional Annotation Properties tab



Annotation Property Name Name assigned to the property.

Type Type assigned to the property.

Value Value assigned to the property.

Different default properties may already be assigned to the annotations added from reference annotations.

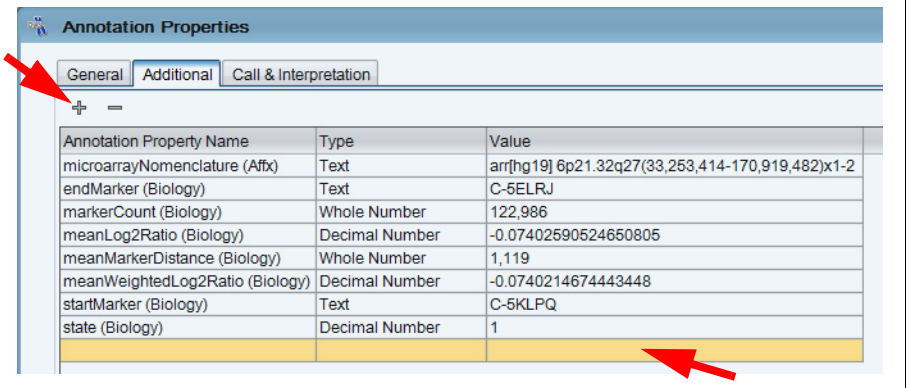
2. Click the **Add** button at the top of the table.

   A new row appears in the table. (Figure 12.21)

   If needed, click on a column's edge, then drag it (right) to make it wider.

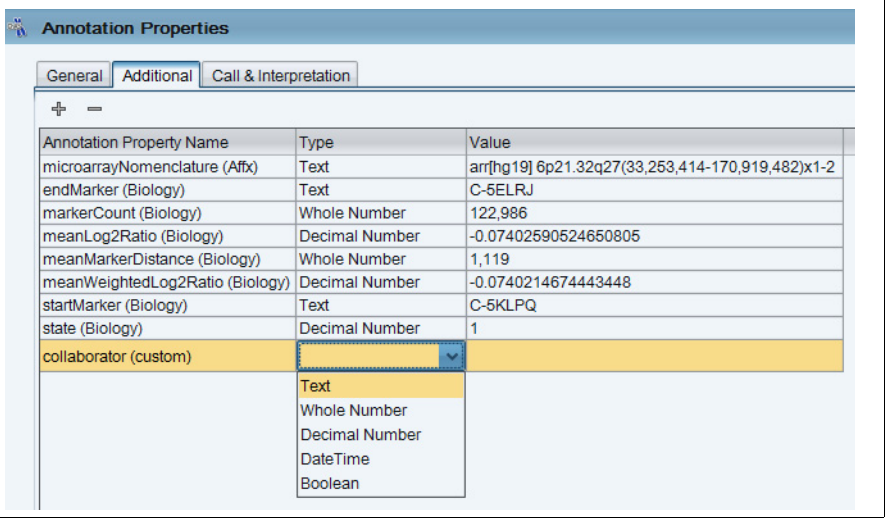
   You can delete a property by selecting a row and clicking the **Remove** button.

Figure 12.21 New property row added



3. Click in the row under the Property Name column and enter a name for the property. Note that your new entry is followed by the text "(custom)". (Figure 12.22) For more details, see *ChAS Properties and Types* (on page 307).

Figure 12.22 Entering a custom property name and selecting its Type



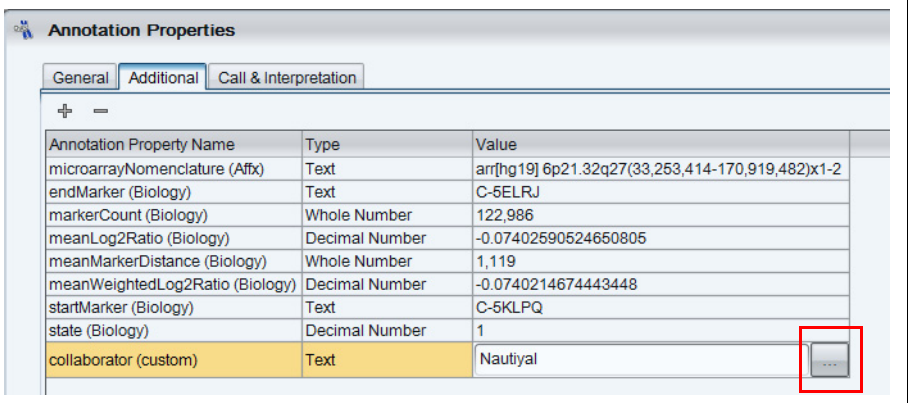
4. Click in the row under Type and select a property type from the drop-down list (Figure 12.23). For more details, see *ChAS Properties and Types* (on page 307).

Choose from:

- Text
- Whole Number
- Decimal Number
- Date Time
- Boolean

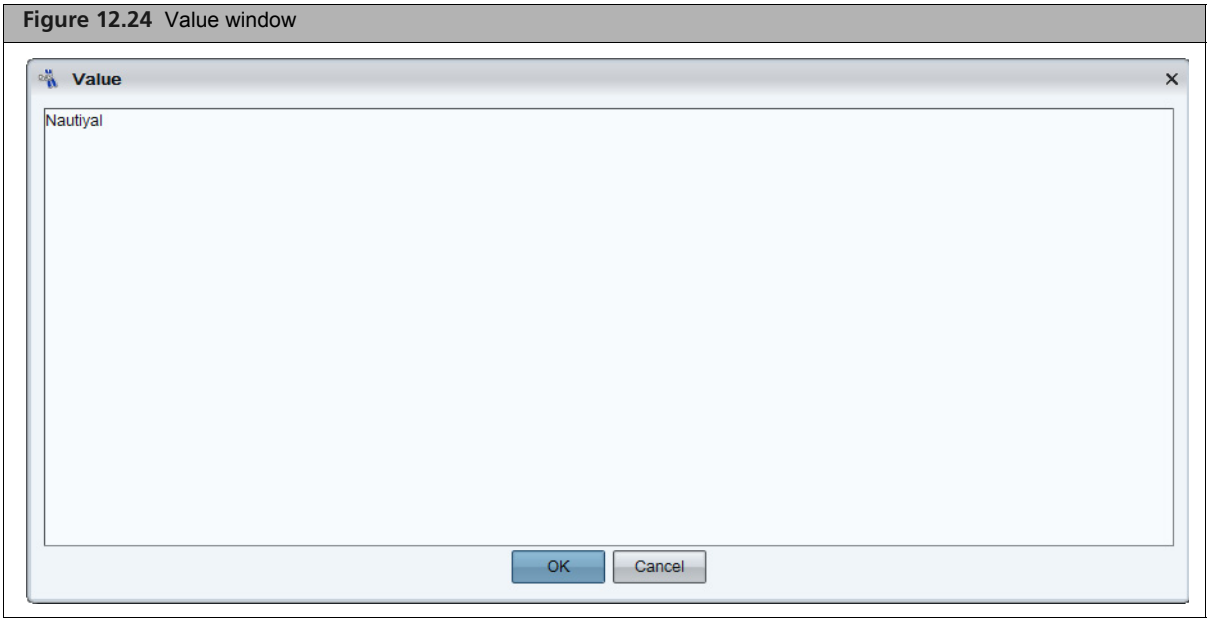
5. Click in the row under Value and enter the value (Figure 12.23).

Figure 12.23 Enter a value for the property type



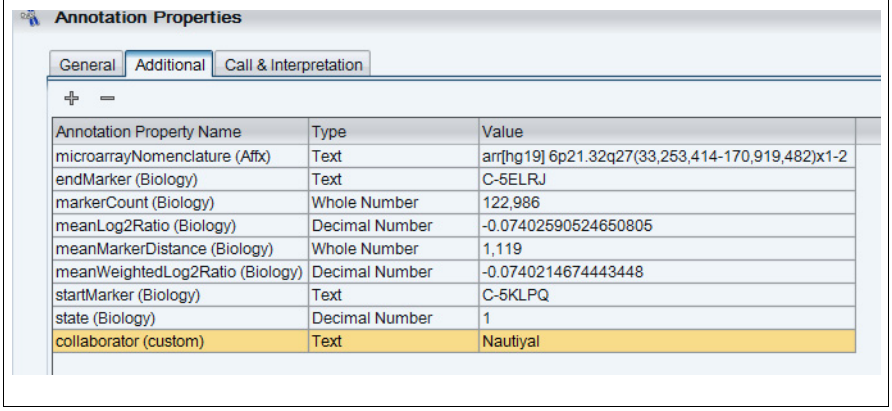
Alternatively, click the **Browse** button, and in the **Value** window that appears, enter the property value and click **OK.** (Figure 12.24 on page 212)

Figure 12.24 Value window



The property entry is completed (Figure 12.25).

Figure 12.25 Property entry completed



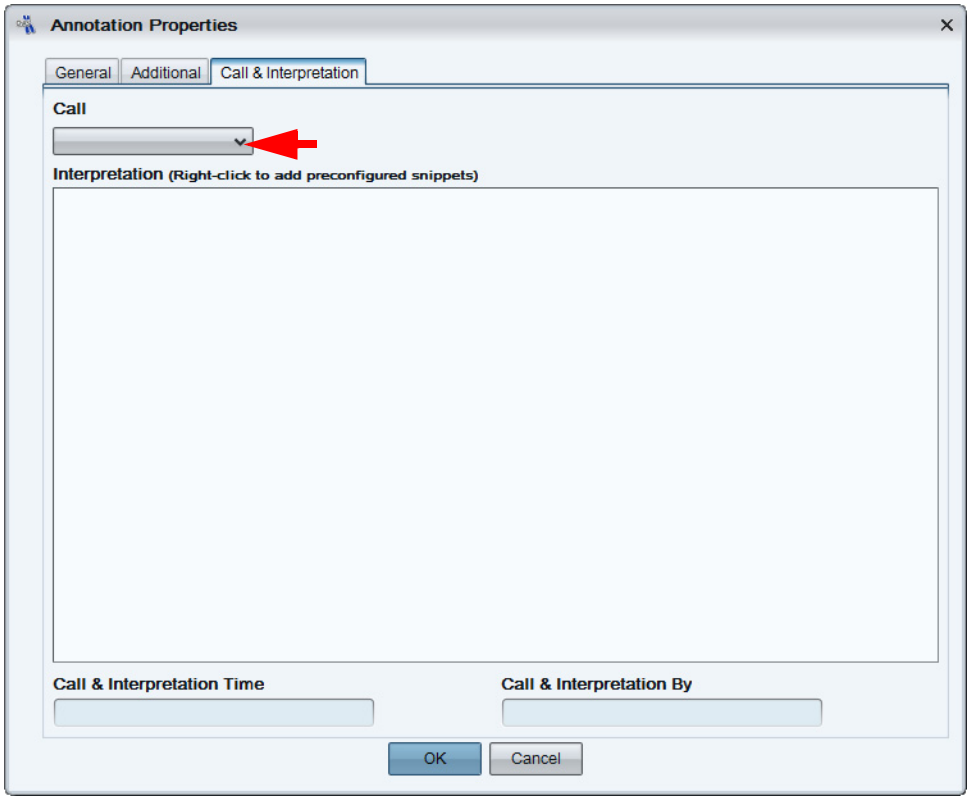
6. Click **OK** in the Annotation Properties window.

Adding a Call and/or Interpretation

To add a Call & Interpretation:

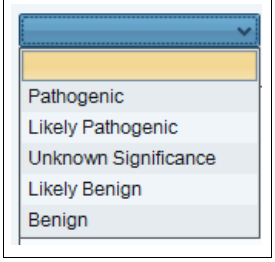
1. Click the **Call & Interpretation** tab.
   The Call & Interpretation tab window appears. (Figure 12.26 on page 213

Figure 12.26 Call & Interpretation tab



2. Click the **Call** drop-down to select a Call. (Figure 12.27)

Figure 12.27 Call drop-down menu



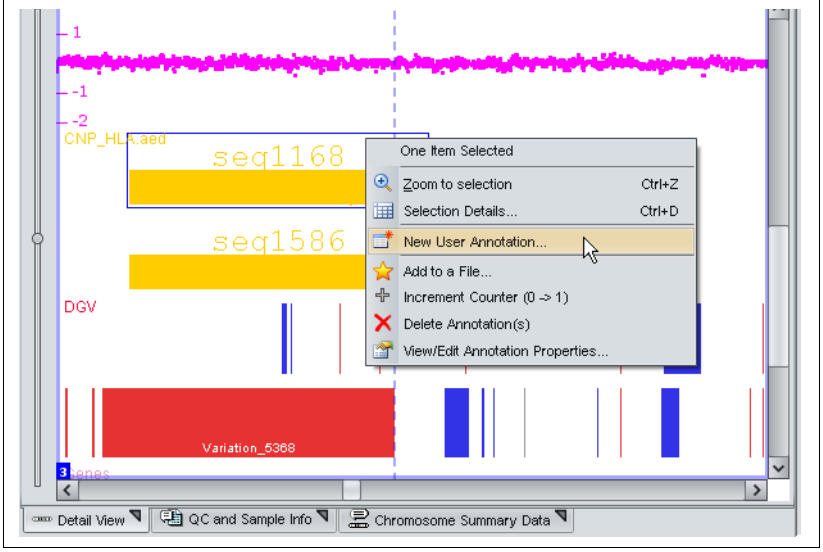
3. Click to select an appropriate Call. For information on how to set up and use Calls, see *Setting Up and Using the Calls Feature* (on page 164)
4. If needed, (in the Interpretation field) type in your interpretation or right-click to add a pre-configured snippet. For information on how to set up and use Snippets, see *Setting Up and Using the Interpretations and Snippets Feature* (on page 168).
5. Click **OK**.

New User Annotations

You can create a new annotation from an AED annotation.

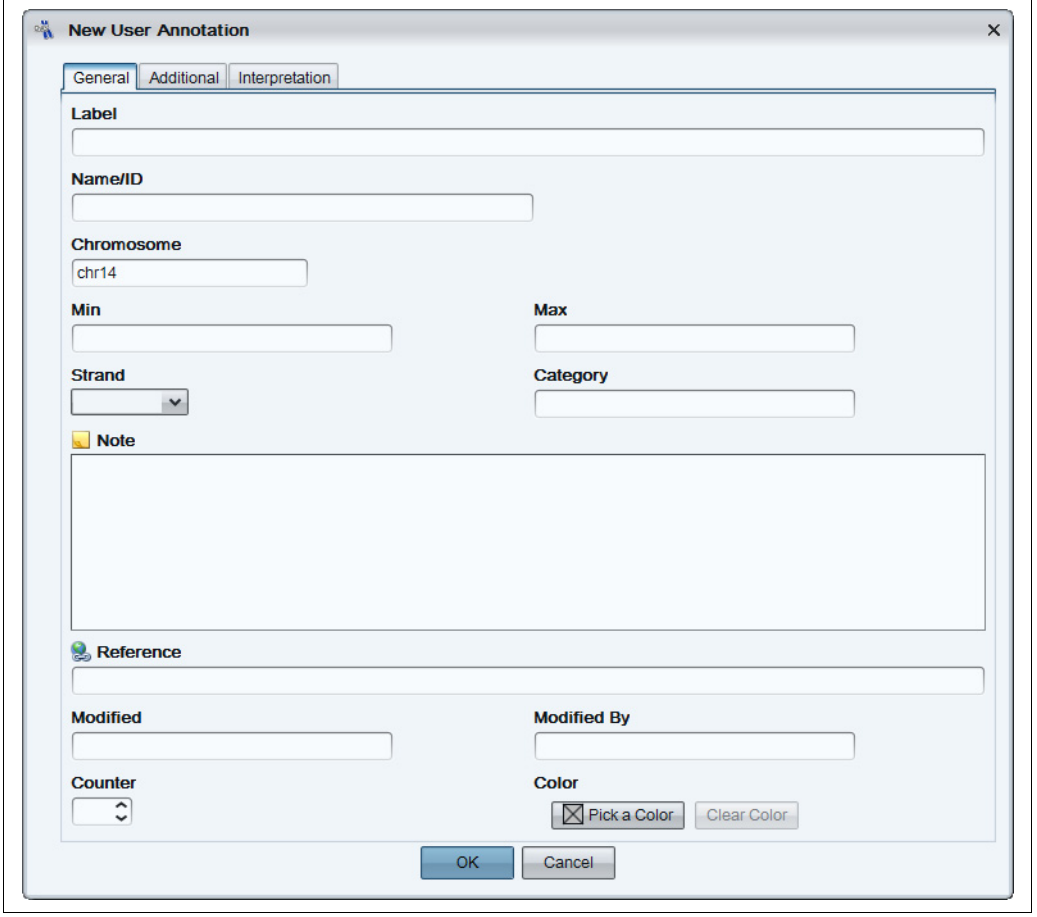
1. In the Details view, right-click an AED annotation and select New User Annotation on the shortcut menu (Figure 12.28).

Figure 12.28 Detail View



2. In the window that appears, enter the annotation information (Figure 12.29). For more details, see *Entering General Information* (on page 208) and *Customizing Properties* (on page 209).

Figure 12.29 Enter information for the new annotation



3. Click **OK**.

   The new user annotation is created and saved in the AED file.

NOTE: The default New User Annotation information includes only the chromosome number. It does not include any information or properties associated with the AED annotation.

Viewing and Batch Editing AED File Contents

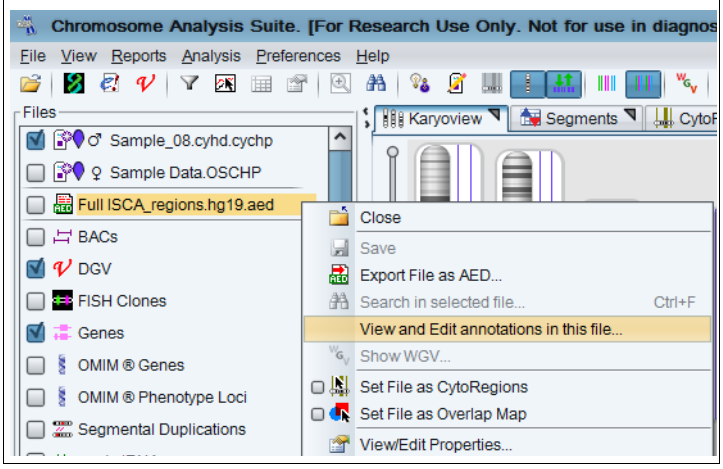
The AED Editor allows you to view multiple AED annotations and their properties in a table format and edit the values of properties easily - for single annotations or to edit a property for multiple annotations in a batch - to a new value.

To view and batch edit AED file contents:

1. At the Files pane (top left), right-click on the AED file you want to view/edit.

The following menu appears: (Figure 12.30)

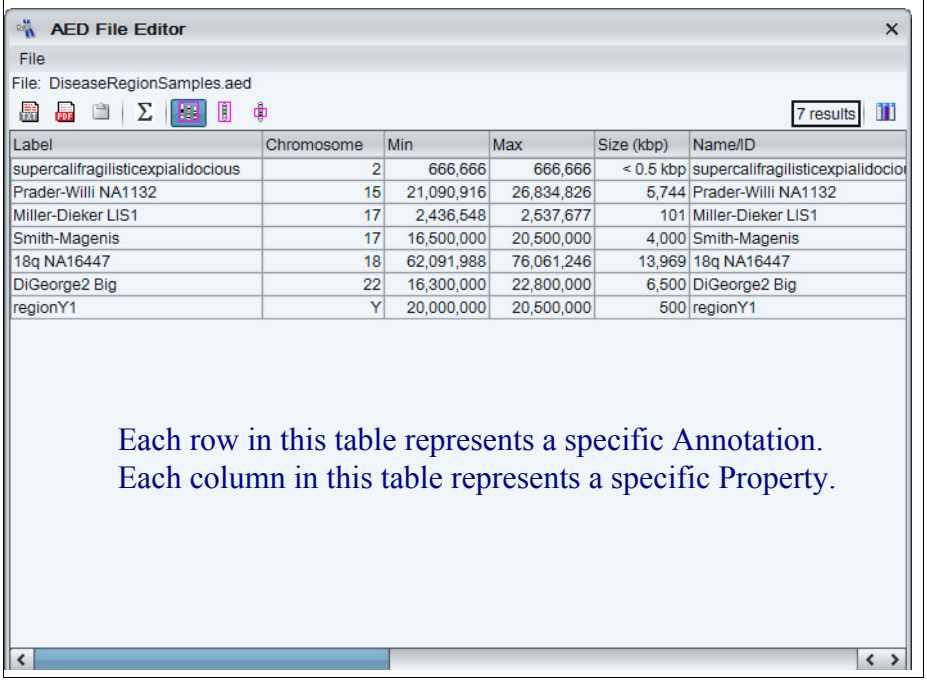
Figure 12.30 AED file - right-click menu



2. Click **View and Edit annotations in this file**.

   The AED File Editor table window appears. (Figure 12.31)

Figure 12.31 AED File Editor table



NOTE: The AED File Editor table displays ALL properties - of ALL the tabs - of ALL the annotations contained in an AED file.

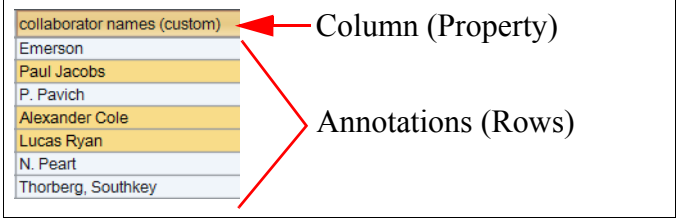
Use the scroll bar to move the table to the right to see more column entries and/or click, then drag the AED File Editor table window to make it wider.

TIP: To locate your table data faster, click on any appropriate column header to sort in a descending or ascending order.

NOTE: The number of columns displayed in the Select Columns window, varies with each AED file.

3. Locate the group of editable annotations (rows) from a column (property) you want to batch edit (combine together) as a single annotation.
4. Shift click or Ctrl click to select (highlight) each annotation (row) entry. (Figure 12.32).

Figure 12.32 Selection of multiple annotations



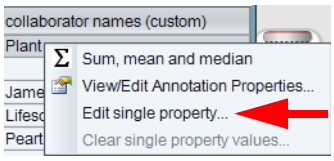
5. Right-click on a newly selected annotation.

   The following menu appears: (Figure 12.33)

IMPORTANT: Only editable fields can be edited. If a field is non-editable, its right-click menu selections are disabled (grayed out).

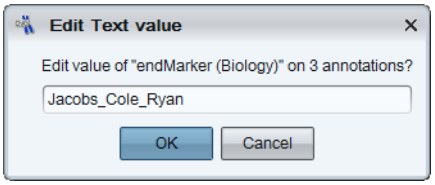
Not all user-editable AED file fields may currently be edited from within the AED Editor. Some basic values (start, stop, type) cannot be edited in the AED Editor table directly. You MUST use "View/Edit Annotation Properties..." for editing the particular field in the annotation of interest.

Figure 12.33 AED table field - right-click menu



6. Click **Edit single property**.
7. The following window appears: (Figure 12.34)

Figure 12.34 Edit Text Valu



8. Enter an appropriate label to distinguish your new batch annotation entry, then click **OK**.
9. Your batch (group) of annotations appear as follows: (Figure 12.35)

Figure 12.35 Batch edit results - Before and After

Before Batch Editing

collaborator names (custom)
Emerson
Paul Jacobs
P. Pavich
Alexander Cole
Lucas Ryan
N. Peart
Thorberg, Southkey

After Batch Editing

collaborator names (custom)
Emerson
Jacobs_Cole_Ryan
P. Pavich
Jacobs_Cole_Ryan
Jacobs_Cole_Ryan
N. Peart
Thorberg, Southkey

AED/BED Color Rules

The Color Rules feature enables you to set display colors for annotations in AED files by various annotation properties, depending upon the original source of the region (detected segment, reference annotation, etc.). You can color annotations using the properties of the annotations, including:

- name
- category
- markerCount
- confidence

Using the color rules, you can assign a different color to annotations with different properties, making it easier to track the different types of segment data stored in AED files. For example, you could assign different colors to different categories:

- GainMosaic
- LossMosaic
- Gain
- Loss
- LOH

You can also perform comparisons for numerical values, coloring only values above or below a certain level.

By default, ChAS displays the regions in an AED or BED file in a single color. There are several ways to edit annotation color:

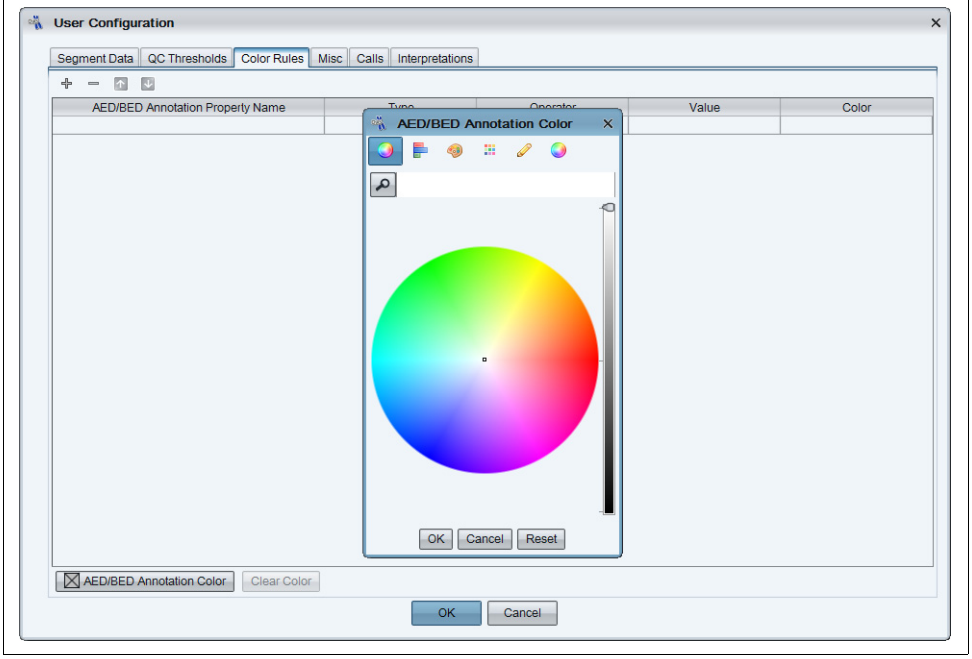
- Select a new color for all AED or BED files (page 220).
- Create a color rule that specifies a color for annotations which meet user-specified requirements (page 220).
- In the Detail View, choose a color for a particular annotation in the Annotation Property dialog box (page 220).

NOTE: **An AED annotation color set in the Annotation Properties dialog box (accessed in the Detail View) takes precedence over a color rule and the default AED/BED file color. A color rule can overwrite the AED or BED file default color.**

To select a default color for loaded AED or BED files:

1. Open the User Configuration dialog box (click the button or select **Preferences → Edit User Configuration** on the menu bar).
2. In the Color Rules tab (Figure 12.36), click the AED/BED Annotation Color button.
3. In the dialog box that appears, choose a color swatch or use the color controls to specify a color.
4. Click **OK** in the AED/BED Annotation Color dialog box.

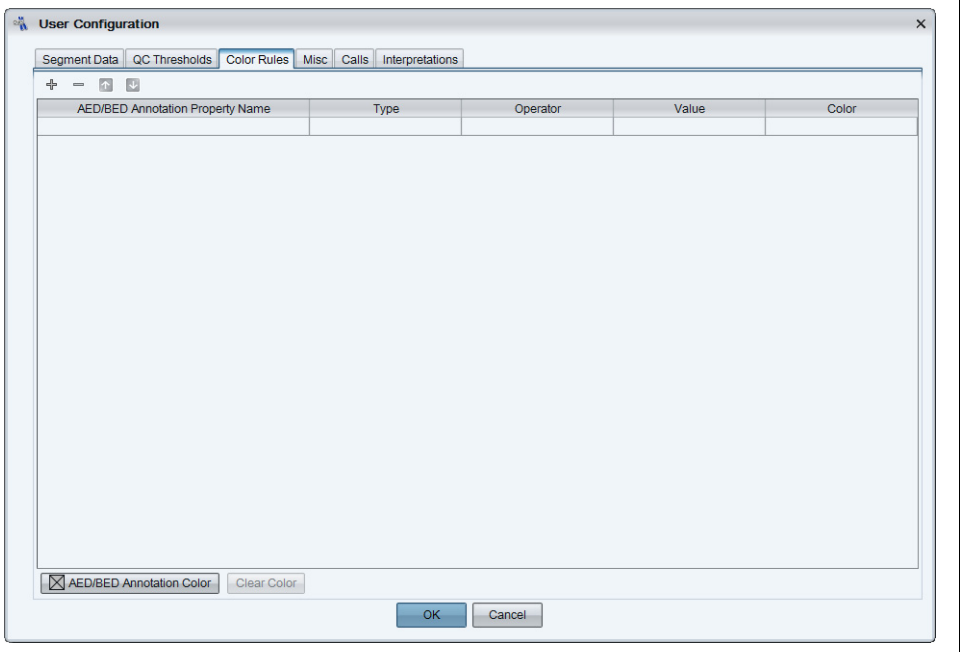
Figure 12.36 AED/BED Annotation Color dialog box



To create a color rule:

1. Open the User Configuration dialog box (click the button or select **Preferences → Edit User Configuration** on the menu bar).
2. Click the **Color Rules** tab (Figure 12.37).

Figure 12.37 User Configuration dialog box, Color Rules tab

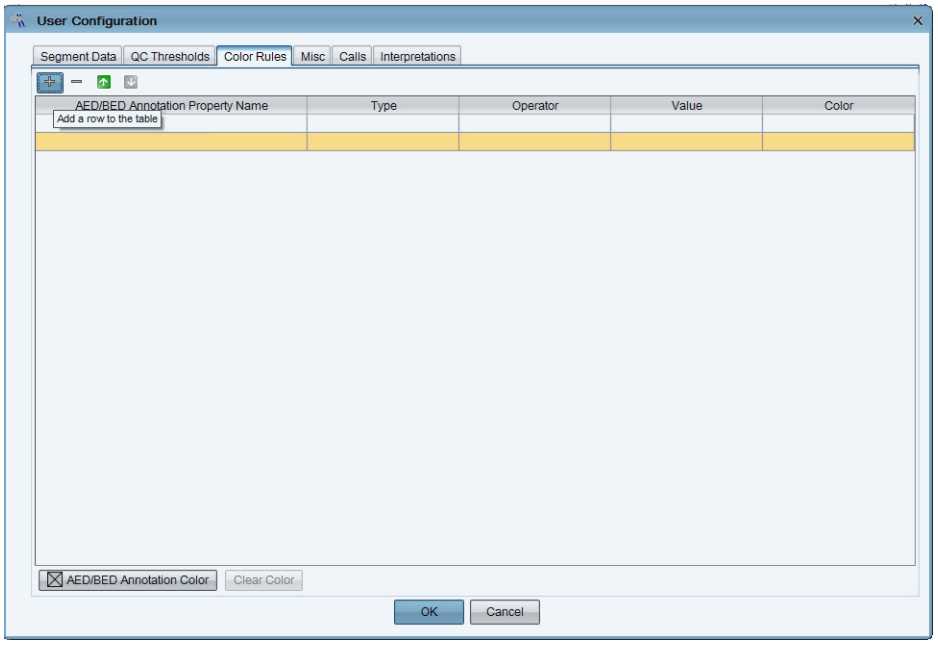


The tab has five columns:

AED/BED Annotation Property Name	Name assigned to the property.
Type	Type assigned to the property (for example, text).
Operator	Type of comparison with value performed .
Value	Value assigned to the property.
Color	Color assigned to the region property and value.

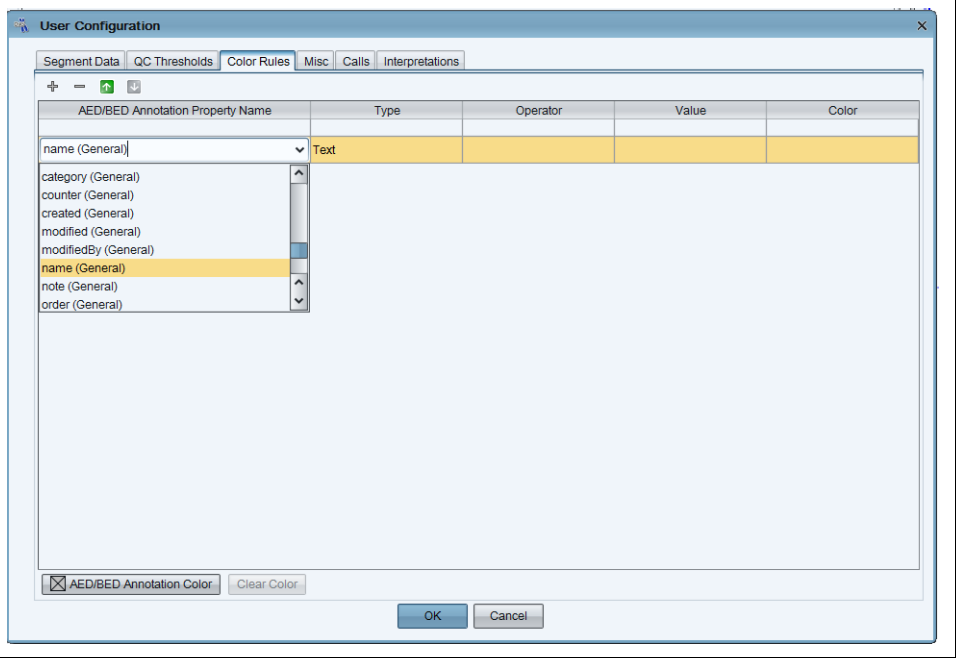
3. Click the **Add** button at the top of the table (Figure 12.38).
   A new row appears in the table.
   You can delete a property by selecting a row and clicking the **Remove** button.

Figure 12.38 New row added to the Color Rules table



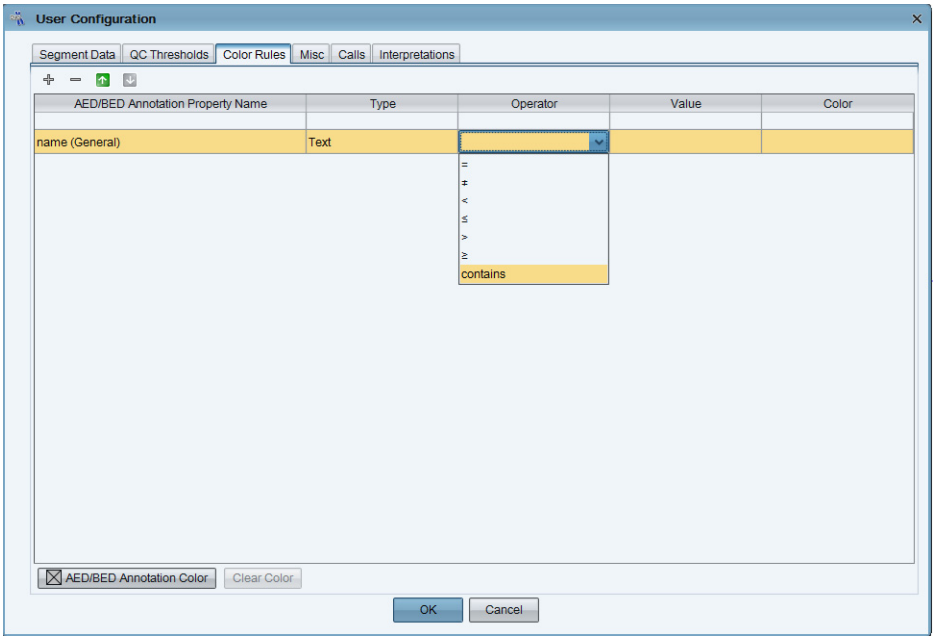
4. Click in the row under the Property Name column and enter a name for the property or select a property from the drop-down list (Figure 12.39). For more details, see *Appendix C, ChAS Properties and Types* (on page 307).

Figure 12.39 Enter or select a property name



5. Click in the row under Type and select a property type from the drop-down list (Figure 12.40).
6. Click in the row under Operator and select an operator from the drop-down list (Figure 12.40).

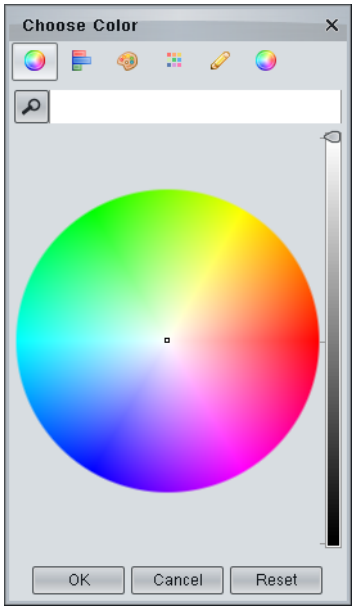
Figure 12.40 Selecting property type and operator for the comparison



7. Click in the row under Value and enter a value for the property.
8. Click in the row under Color.

   The Pick a Color dialog box opens (Figure 12.41).

Figure 12.41 Pick a Color dialog box

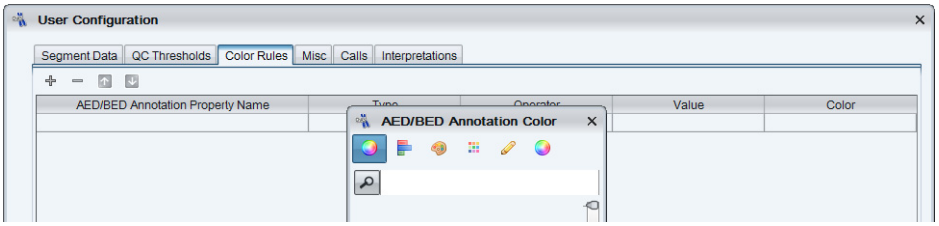


The dialog box provides several options for selecting a color.

9. Select a color and click **OK** in the Pick a Color dialog box.

   The property entry is completed (Figure 12.42).

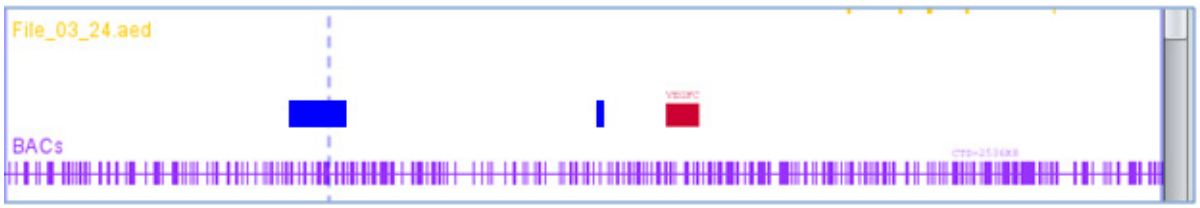
Figure 12.42 Property entry completed



10. Click **OK** in the Color Rules tab dialog box.

Regions that satisfy the property comparison are displayed in the selected color in all views (Figure 12.43).

Figure 12.43 Regions displayed in color



Exporting Information in AED or BED Format

You can export position data for the different features to AED or BED file format. The exported BED file contains only the names and locations of the detected segments. The exported AED file contains additional information, such as header information, feature ID, and hg version (which is the same as the NetAffx Genomic Annotations Database file loaded in the Browser at export).

NOTE: AED or BED files created in ChAS 1.0, 1.0.1, or 1.1 do not automatically include the hg version.

Table 12.1

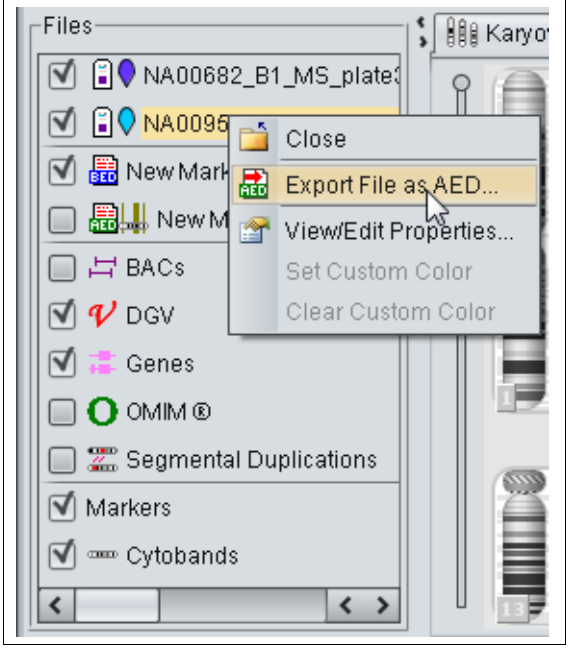
Position Data	Export to AED File	Export to BED File
Detected Segments for xxCHP files	Regions, names, and properties	Regions and names
Annotation Features in Reference Annotation files	Regions, names, and properties	Regions and names

To export position data as an AED or BED file:

1. From the File menu, select **Export as AED…**

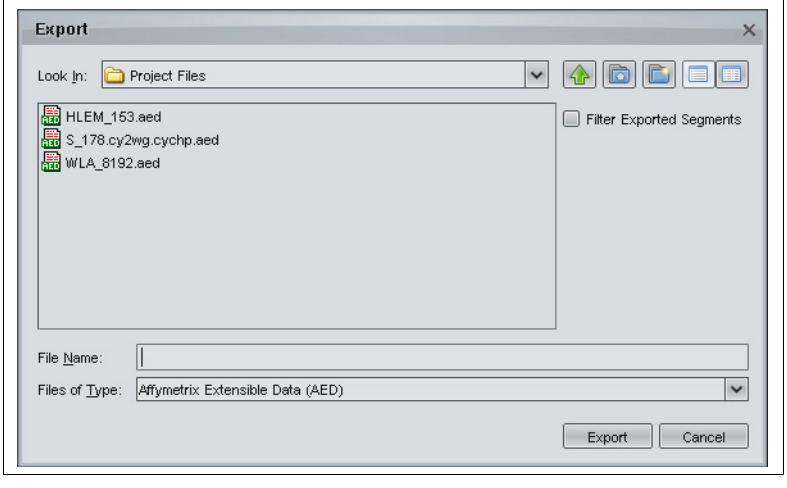
Alternatively, right-click the file in the Files list and select **Export File as AED** on the menu (Figure 12.44).

Figure 12.44 Right-click menu



The Export dialog box opens (Figure 12.45).

Figure 12.45 Export dialog box, AED file format selected



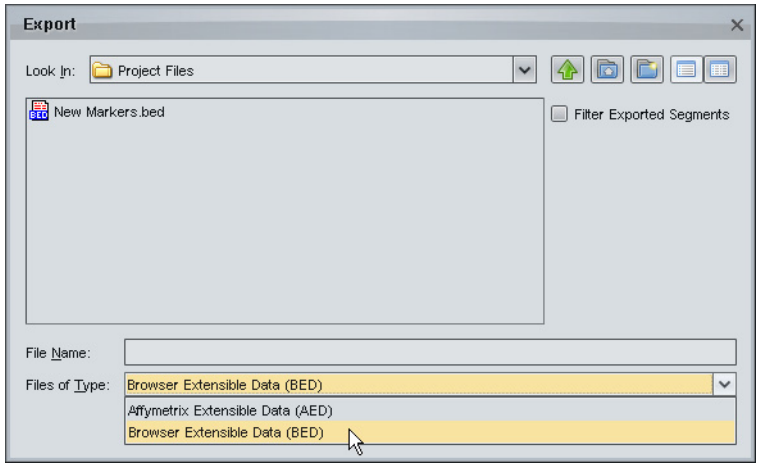
2. Choose the "Filter Exported Segments" option to restrict the export to the contents of the Segments table. If filters have been applied to the data, only the retained segments will be exported. Graph data and Chromosome Summary data will not be exported.

NOTE: If this option is not selected, all segments which were loaded with the CxCHP file will be exported along with header information, regardless of whether filters are applied. The export includes all segment data, regardless of check mark status (ON or OFF) in the Files windowpane.

3. Select a folder location for the file using the navigation tools.
4. To export to AED file format, enter a name for the file.

To export to BED file format (Figure 12.46), enter a name for the file, and select "Browser Extensible Data (BED)" from the Files of Type drop-down list.

Figure 12.46 Export dialog box, BED file format selected



5. (Optional) Select the Filter Exported Segments dialog box to export only segments that meet filter criteria.
6. Click **Export** in the Export dialog box.

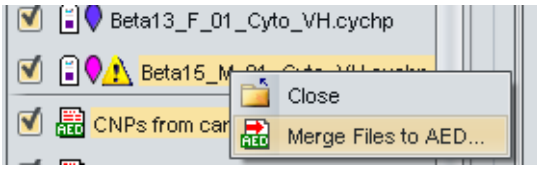
   The AED file is exported and can be loaded as a region information file or sent to another user.

You can also merge feature position data from multiple different files.

To merge and export feature position information for multiple files:

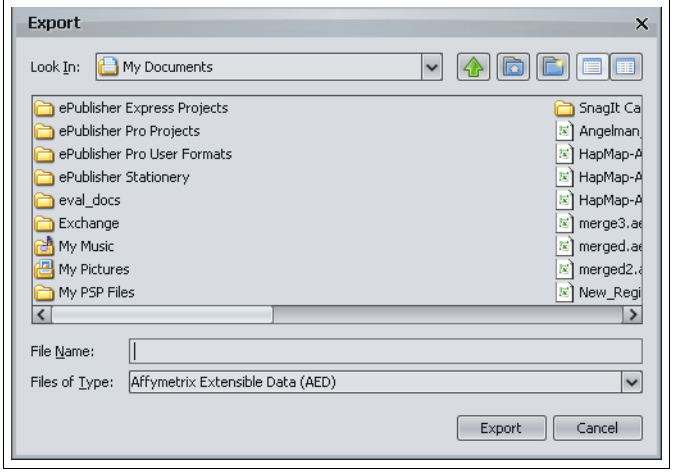
1. Select the files in the File List by clicking on them while pressing the CTRL key.
2. Right-click the selected files (Figure 12.47) and choose **Merge Files to AED** on the menu.

Figure 12.47 Files selected for merging AED outputs



The Export dialog box opens (Figure 12.48).

Figure 12.48 Export dialog box



3. Use the navigation controls to select a folder for the merged AED file and enter a file name for the file.
4. Click **Export**.

   The file with the merged AED region information is created and can be used as a region information file in ChAS.

IMPORTANT: After two AED files are merged, the original metadata in the header is not retained.

NOTE: When two or more files are merged into AED/BED format, the current value of the genome assembly version property, if present in at least one of the files, will be saved in the merged file.

Displaying Data in Table Views

The data in the xxCHP files can also be displayed and exported in tabular format, as well as the graphic representation in the Karyoview, Selected Chromosome View, and Detail View.

These tables include:

- *Selection Details Table* (on page 135)
- *CytoRegions Table* (on page 182)
- *Viewing the Overlap Map Table* (on page 192)
- *Segments Table* (on page 236)
- *Graphs Table* (on page 248)
- *Searching Results* (on page 260)
- *Finding Intersections* (on page 265)

This chapter covers:

- *Common Table Operations* (on page 228)
- *Segments Table* (on page 236)
- *Graphs Table* (on page 248)
- *Somatic Mutations Table* (on page 253)
- *QC and Sample Info Tab* (on page 255)
- *Chromosome Summary Data* (on page 258)
- *Auto-Generated Autosome LOH Percentage* (on page 260)
- *Searching Results* (on page 260)
- *Finding Intersections* (on page 265)

Common Table Operations

The controls that are common to all tables are described in this section. They include:

- *The Standard Tool Bar Controls* (on page 228)
- *Sorting by Columns* (on page 229)
- *Ordering Columns* (on page 230)
- *Selecting Columns for Display* (on page 230)
- *Sum, Mean, and Median Calculator* (on page 232)

The Standard Tool Bar Controls

The tool bar provides quick access to table functions.

Some of the tool bar controls are standard to all tables.

At Left:

Export as TXT file. See *Export Tables as TXT File* (on page 280).

Export as PDF Report. See *Export Table Data in a PDF File* (on page 275).

Copy selected cells to clipboard. See *Exporting a Segments Table with Modified Segments to a .TXT File* (on page 282).

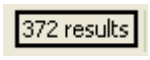
Calculate the sum, median, and mean of the selected values from a numeric column.

Display results for entire genome.

Display results for selected chromosome.

Display results for portion of chromosome displayed in Detail View.

At Far Right:



The number of rows in the table.

Opens the Select Columns dialog box that enables you to choose the column headers to show or hide.

The Export functions are described in *Exporting Table Data* (on page 274).

Sorting by Columns

You can sort a table by a single column's values, or by the values in up to three columns.

NOTE: You may sort on any column except, for reasons of efficiency, the marker name column in the graphs table.

Sorting on certain columns can cause a noticeable decrease in performance. For example, it is recommended that you do not sort a table using the columns in the Segments table that show the overlapping RefSeq, FISHClones, or other items. The data for these table cells is calculated only on an as-needed basis when it needs to be displayed. Using such a column for sorting would force the calculation of the data for all such cells. Since the sorting would be alphabetical, it is unlikely to be useful. Similarly, for reasons of efficiency, sorting based on the marker name column in the Graphs table is not allowed.

To sort a table by a single column:

- Click in the header of the column to sort the table by that column's values.

  A triangle appears in the header.

  - Triangle pointing upwards

indicates ascending sort order.
  - Triangle pointing downwards

indicates descending sort order.

  Click the header to toggle between Ascending, Descending, and no sort order.

NOTE: The "Type" column in the Segments table sorts segments based on the order that they appear in the Data Types windowpane, not in ascending/descending alphabetical order.

To perform a multi-column sort:

1. Click in the header of the first column you want to sort on.
2. Click in the header of the second column.
3. Click in the header of the third column, if desired.

   The last selected column has sort priority.

   Sort priority is indicated by the size of the triangle in the header (Figure 13.1).

Figure 13.1 Table sorted by descending order of Segment ID and ascending order of Size

CytoRegion Type	CytoRegion	Chromosome	Min	Max	▲ Size (kbp)	▼ Segment ID	▾ Segment File	Segment
Default User Annotation	CNP505	3	163,995,351	164,106,575	111	seg684	Beta10_F_01_Cyto_VH.cychp	▲ Gain
Default User Annotation	CNP2269	17	41,521,619	41,719,991	198	seg6392	Beta13_F_01_Cyto_VH.cychp	▲ Gain
Default User Annotation	CNP12506	17	26,040,482	26,043,732	3	seg6312	Beta13_F_01_Cyto_VH.cychp	▲ Gain
Default User Annotation	CNP2197	16	72,953,795	73,009,537	55	seg6151	Beta13_F_01_Cyto_VH.cychp	▲ Gain
Default User Annotation	CNP2156	16	21,422,575	21,498,841	76	seg5977	Beta13_F_01_Cyto_VH.cychp	▲ Gain
Default User Annotation	CNP147	1	194,997,658	195,068,695	71	seg587	Beta13_F_01_Cyto_VH.cychp	▲ Gain
Default User Annotation	CNP148	1	195,089,940	195,168,372	78	seg587	Beta13_F_01_Cyto_VH.cychp	▲ Gain
Default User Annotation	CNP933	6	32,539,530	32,681,749	142	seg572	Beta15_M_01_Cyto_VH.cychp	▼ Loss
Default User Annotation	CNP901	5	180,311,316	180,350,709	39	seg537	Beta15_M_01_Cyto_VH.cychp	▼ Loss
Default User Annotation	CNP12104	13	69,640,329	69,670,896	30	seg5218	Beta13_F_01_Cyto_VH.cychp	▼ Loss
Default User Annotation	CNP1952	13	68,149,981	68,166,243	16	seg5216	Beta13_F_01_Cyto_VH.cychp	▼ Loss
Default User Annotation	CNP12099	13	66,945,140	66,954,900	9	seg5214	Beta13_F_01_Cyto_VH.cychp	▼ Loss
Default User Annotation	CNP12094	13	65,094,528	65,103,708	9	seg5210	Beta13_F_01_Cyto_VH.cychp	▼ Loss
Default User Annotation	CNP1946	13	63,122,789	63,134,693	11	seg5208	Beta13_F_01_Cyto_VH.cychp	▼ Loss
Default User Annotation	CNP12093	13	64,198,454	64,209,619	11	seg5208	Beta13_F_01_Cyto_VH.cychp	▼ Loss
Default User Annotation	CNP12091	13	63,227,094	63,303,323	76	seg5208	Beta13_F_01_Cyto_VH.cychp	▼ Loss

Ordering Columns

To change the order of columns in a table:

- Click and drag in the column header to move the column to a new location.

Selecting Columns for Display

You can select columns using 2 methods:

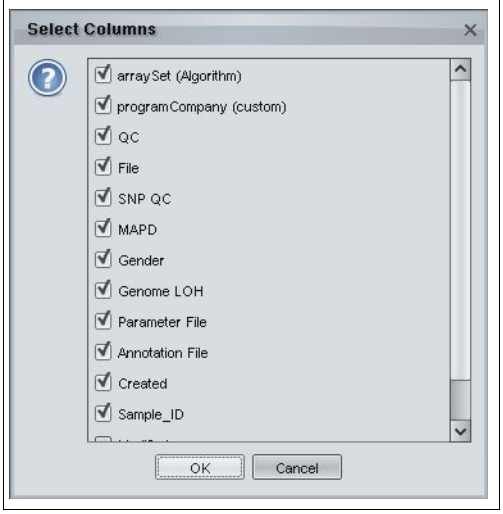
- Select Columns dialog box
- Right click on a column header, then choose **Select columns**.

To select columns to display or hide:

1. Click the Select Columns tool bar button .

   The Select Columns dialog box opens (Figure 13.2). The checkboxes for the columns are listed in the order currently displayed. Columns that are currently hidden in the table, appear at the end of the list.

Figure 13.2 Select Columns dialog box



NOTE: The specific items will vary, depending upon the type of table and data being displayed.

2. Select or deselect the checkboxes to display or hide the data and click **OK**.

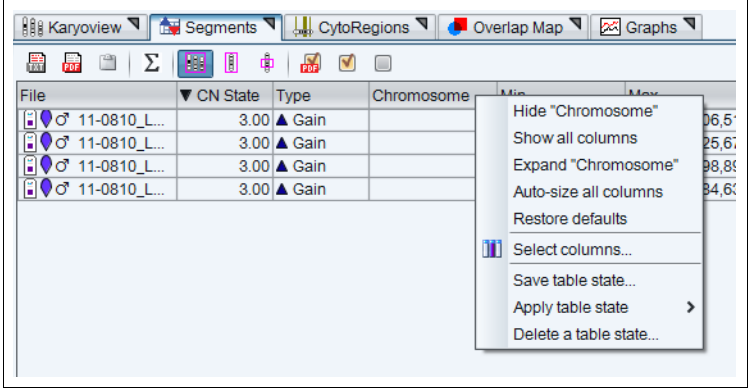
   The selected columns are displayed. These choices are remembered between sessions.

Right-click Menu

You can right-click in a column header to display a menu (Figure 13.2) to perform the following:

- Hide the selected column
- Show all columns, including hidden ones
- Expand selected column to display complete heading.
- Auto-size all columns.
- Restore the default selection of columns.
- Select other columns to hide or show.
- Create, select, save, and delete saved table states.

Figure 13.3 Column heading right-click menu



To use a heading's right-click menu:

1. Right-click in the heading.
   The menu appears (Figure 13.3).

> **NOTE: Specific right-click menu items will vary, depending upon the type of table and data being displayed.**

2. Select the desired option from the menu.

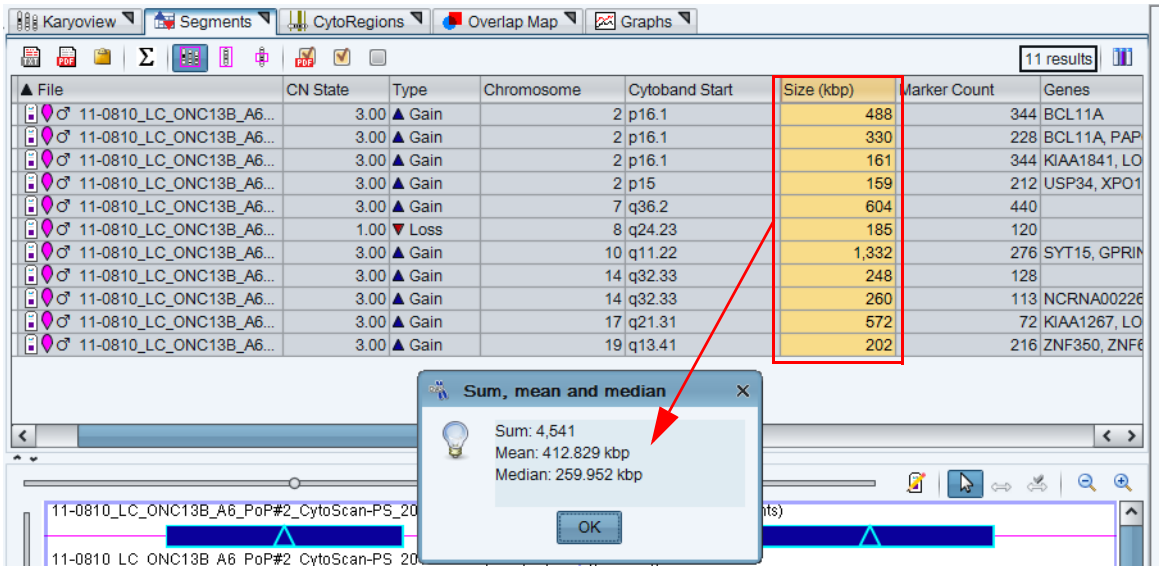
Sum, Mean, and Median Calculator

Use this calculator to calculate the sum, mean, and median of the selected or multiple numeric fields.

To calculate a multiple numeric values in a column:

1. Ctrl click or Shift click to highlight multiple numeric fields (MUST BE in the same column).
2. Click Σ.
   Your multiple numeric values are calculated and summarized. (Figure 13.4 on page 232)
3. Click **OK**.:

Figure 13.4 Calculator



Saved Table States

Saving multiple customized tables is a very useful tool, as it allows you to use a large number of columns in a table for interpreting the biology of a particular sample. On the other hand, saving a table state with a smaller subset of columns in a different order - may be more appropriate for including in a report.

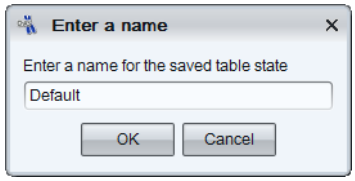
In either case, having various Saved Table States allows you to switch back and forth between multiple sets of tables/columns without having to painstakingly recreate interpreting and reporting columns for each of your sample selections.

IMPORTANT: You can restore your default settings at any time by right-clicking on a table column header and selecting "Restore Defaults" from the menu.

Do the following to save the current segment table state to Default:

1. Right-click on any column header
2. Click **Save table state**.
3. The following dialog window appears: (Figure 13.5)

Figure 13.5 Save Table State - Enter a name



4. Type **Default**.
5. Click **OK**.

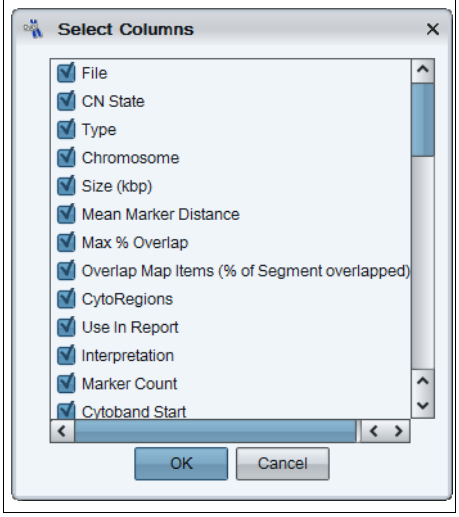
   The table is now saved (as Default) to your User Profile for future use/reference.

Modifying Table States

To add more columns into a table (Example: For use in interpreting the biology of a particular sample):

1. Click the Select Columns icon. (Top right of Segments table)
   A list of available columns appears. (Figure 13.6)

Figure 13.6 Save Table State - Available Columns



2. Move the scroll bar downward to reveal more available columns.

3. Click a column's checkbox to select it.
4. Click **OK**.

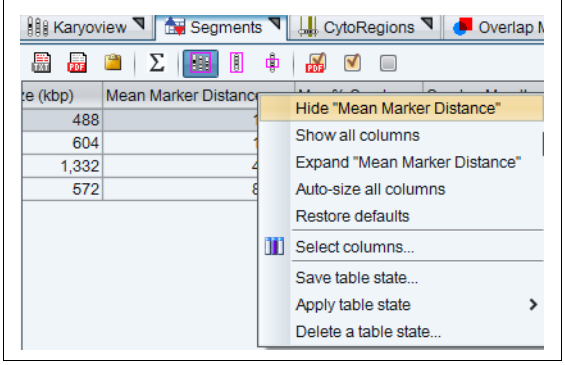
   The additional column(s) is now added to the **FAR RIGHT** of your table.
5. Click on the new column, then drag and drop it into your desired spot/order within the table.
6. You may need to use the left/right scroll bar (in combination with) "drag and drop" to place your new column(s) farther left.

To remove (hide) columns in a table (Example: For use in a report):

1. Right-click on a column header you want to remove (hide) from the table.

   The following menu appears: (Figure 13.7)

Figure 13.7 Save Table State - Remove (Hide) Column



2. Click to select **Hide "Column Name"**.

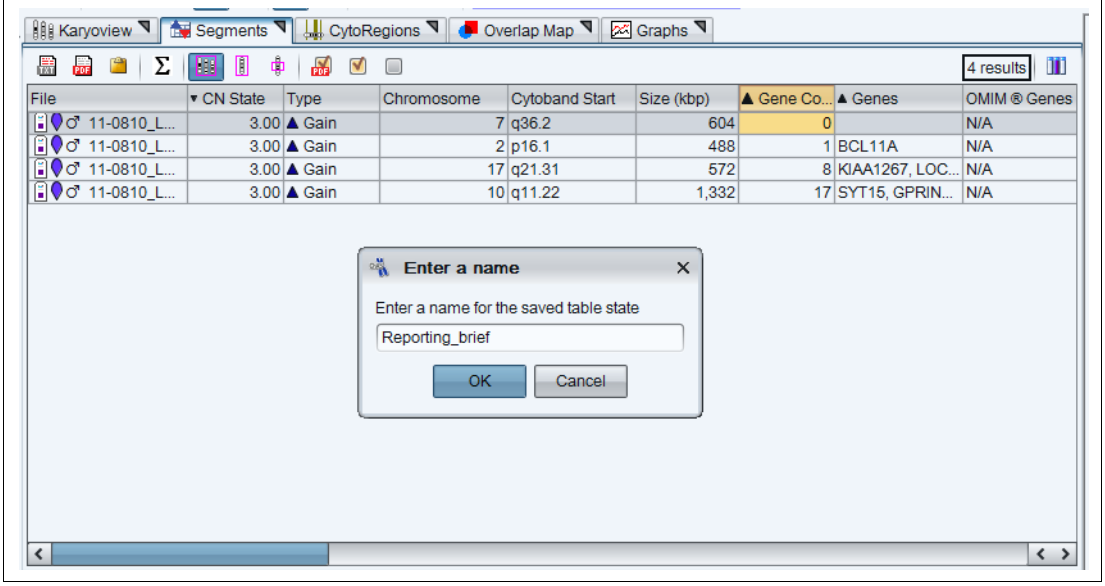
   The column is now removed (hidden) from your table.
3. Repeat steps 1-2 to remove (hide) additional columns that are not required for a report.

Do the following to save your newly edited segment table state:

1. Right-click on any column header.
2. Click **Save table state**.

3. The following dialog window appears: (Figure 13.5)

Figure 13.8 Saved Table State - Example: Table used for reporting.



4. Type a name for your new table state. (Example: Reporting_brief)

   The saved table state remembers the columns that you selected, their order, their widths, and which ones were used for sorting.

5. Click **OK**.

   The edited table is now saved to your User Profile for future use/reference.

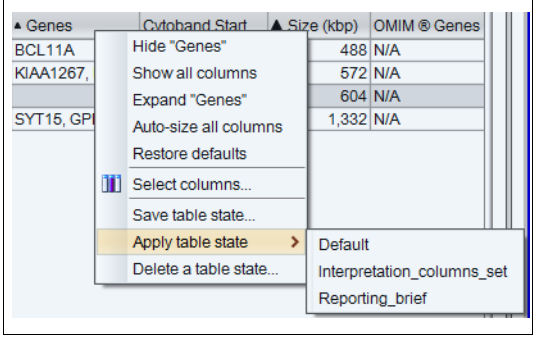
Applying Previously Saved Table States

To apply (open) a Saved Table State:

1. Right-click on any column header
2. Click **Apply table state**.

   The following sub-menu appears displaying your list of saved tables: (Figure 13.9)

Figure 13.9 Apply Table State - Saved tables list



3. Click on the table you want to display.

   The table is now displayed.

4. Repeat steps 1-3 to apply (open) other previously saved tables as needed.

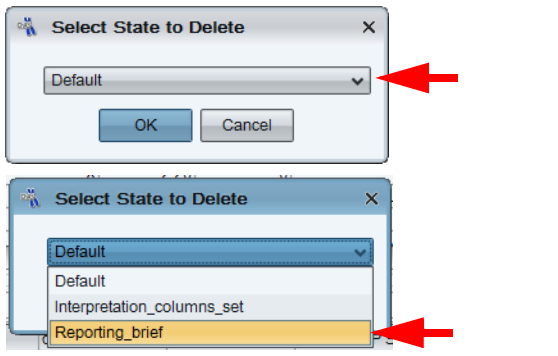
Deleting Previously Saved Table States

To delete a previously Saved Table State:

1. Right-click on any column header
2. Click **Delete table state**.

   The following dialog window appears: (Figure 13.10)

Figure 13.10 Apply Table State - Deleting saved tables

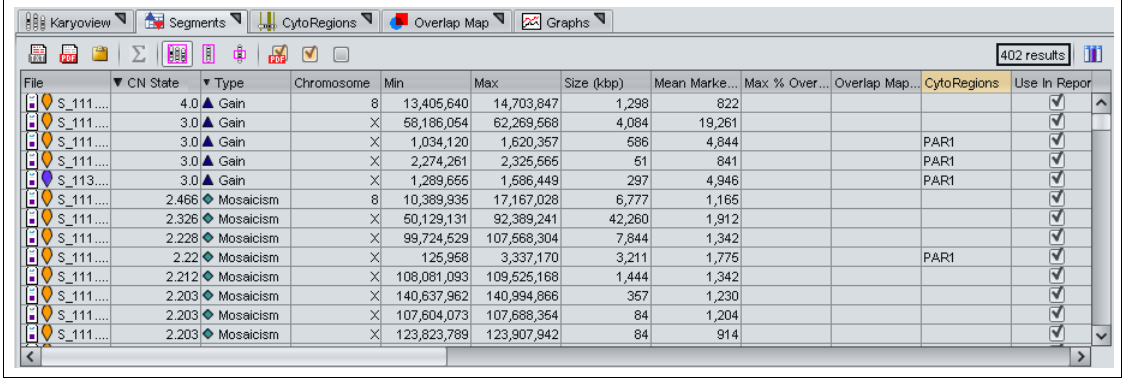


3. Click the drop-down button to reveal your saved tables.
4. Click to select the table you want to delete.
5. Click **OK**
6. Repeat steps 1-5 to delete other saved tables as needed.

Segments Table

The Segments table (Figure 13.11) displays a list of the detected segments in the loaded sample data files.

Figure 13.11 Segments Table



File	CN State	Type	Chromosome	Min	Max	Size (kbp)	Mean Marke...	Max % Over...	Overlap Map...	CytoRegions	Use In Repor
S_111....	4.0	Gain	8	13,405,640	14,703,847	1,298	822				
S_111....	3.0	Gain	X	58,186,054	62,269,568	4,084	19,261				
S_111....	3.0	Gain	X	1,034,120	1,620,357	586	4,844			PAR1	
S_111....	3.0	Gain	X	2,274,261	2,325,565	51	841			PAR1	
S_113....	3.0	Gain	X	1,289,655	1,586,449	297	4,946			PAR1	
S_111....	2.466	Mosaicism	8	10,389,936	17,167,028	6,777	1,165				
S_111....	2.326	Mosaicism	X	50,129,131	92,389,241	42,260	1,912				
S_111....	2.228	Mosaicism	X	99,724,529	107,568,304	7,844	1,342				
S_111....	2.22	Mosaicism	X	125,958	3,337,170	3,211	1,775			PAR1	
S_111....	2.212	Mosaicism	X	108,081,093	109,525,168	1,444	1,342				
S_111....	2.203	Mosaicism	X	140,637,962	140,994,866	357	1,230				
S_111....	2.203	Mosaicism	X	107,604,073	107,688,354	84	1,204				
S_111....	2.203	Mosaicism	X	123,823,789	123,907,942	84	914				

The Segment Table components:

- *Tool Bar* (on page 237)
- *Segments Table* (on page 237)

There are certain columns in the Segments table which dynamically compute intersections of reference annotations with the segments. The data in these columns is computed on an as-needed basis for each cell. You may see text such as "<Working…>" in these cells while the data is being calculated. The results for all cells will be calculated when exporting to PDF or TXT, or copying to the clipboard. Hiding columns that are not needed may improve performance, particularly during export operations.

Sorting the table based on the dynamically computed columns may be slow.

You can also create a PDF report for selected segments only, using the specialized export controls of the Segments table. See *Exporting Selected Segments to a PDF File* (on page 243).

To highlight segments in the views or the table:

- Double-click in a row of the table to zoom to the segment in the Karyoview, Selected Chromosome and Detail Views.
- Click on a segment in the Karyoview, Selected Chromosome or Detail View to highlight the segment in the Segments table.

Tool Bar

The tool bar (Figure 13.12) provides quick access to table functions. The standard functions are described in *The Standard Tool Bar Controls* (on page 228).

Figure 13.12 Segments Table tool bar

The tool bar has three specialized buttons.

Export Checked Rows to a PDF File

Check all in column "Use in Report"

Uncheck all in column "Use in Report".

These buttons are described in *Exporting Selected Segments to a PDF File* (on page 243).

The Results box displays the number of detected segments in the displayed sample files.

Segments Table

In the Segments table, "N/A" can mean that the information (for example, FISH Clones or sno/mRNA) is not available in the NetAffx database because the information has not yet been mapped. For example, FISH Clones or sno/mRNA files will not appear in the Files list for the NA31 (hg19) ChAS Browser NetAffx Genomic Annotations file. "N/A" can also mean that a column which has been persisted to appear from a previous user profile, no longer has data in the current NetAffxGenomic Annotations file that is loaded. For instance the "Old OMIM" column will have "N/A" values when Browser annotations (version NA32.3 or higher) are loaded in ChAS 2.1 (or higher) software. *See ChAS Browser NetAffxGenomic Annotations (on page 311).*

Annotations which share genomic coordinates with a segment are listed in order of start coordinate value, smallest values first (i.e. from left to right in the Details View). For annotations with the same start coordinate (for example, isoforms of a single gene), the one with the smallest end coordinate is listed before others with larger stop coordinates.

If a column in the Segments table contains more than 10 items, "…" is displayed after the 10th item to indicate that some data are not displayed in order to save calculation time. For example, "…" will follow the 10th name in the Genes column. However, a complete list of the genes will be included when the information is copied to the system clipboard or exported to reports. For gene isoforms with identical names, only one instance of the gene locus will be listed in the Segment table to reduce duplicate gene names.

The table can display each segment with the following information (the default set of columns in a new user profile may include only a subset of the total columns listed. *See Saved Table States (on page 232)* for instructions on how to switch quickly between table column sets for a particular table.

Materially Modified Segments (merged, created *de novo*, segments with edited start or end coordinates) and deleted segment have a different appearance in the Segments Table and export to TXT differently depending on the software settings. For more information. *See Exporting a Segments Table with Modified Segments to a .TXT File (on page 282)*.

Segment Table Components

File	File the segment was detected in
CN State	Copy Number State (not displayed for LOH segment types). The expected Copy Number State on the X chromosome in normal males is not constant over its entire length. This is due to the structure of the sex chromosomes. See *LOH Segments on the X and Y Chromosomes* (on page 47) for more information.
Type	Type of segment, for example, LOH. When sorting by this column, the segments of a particular sample are listed in the same order that they appear in the Data Types window pane.
Chromosome	Chromosome on which the segment was found.
Min	Start position of segment.
Max	End position of segment. For all segments, the segment start coordinates are always lower by one bp from the coordinate for the starting probe of the segment as reported in the graphs table while the end coordinate matches the coordinate for the ending probe as reported in the graphs table (see Appendix E, *Genomic Position Coordinates* on page 313).
Size (kbp)	Size of the segment.
Mean Marker Distance	Length of the segment in base pairs divided by the number of markers in the segment.
Interpretation	User-editable field for free-text interpretation on the segment
Call	User-editable field populated by a user-configurable drop list of Calls
Call & Interpretations By	The current computer Operting System login ID and ChAS user profile name at the time that the Call or Interpretation field was last edited.
Call & Interpretation Time	The time and date when the Call or Interpretation field was last edited.
Materially Modified Segment	Indication that segment was previously merged, deleted, or had its start or end boundary, type, or state altered by a ChAS user. (ChAS-based processes of Smoothing and Joining are not "Modifications", nor are making Calls or Interpretations, in this context).

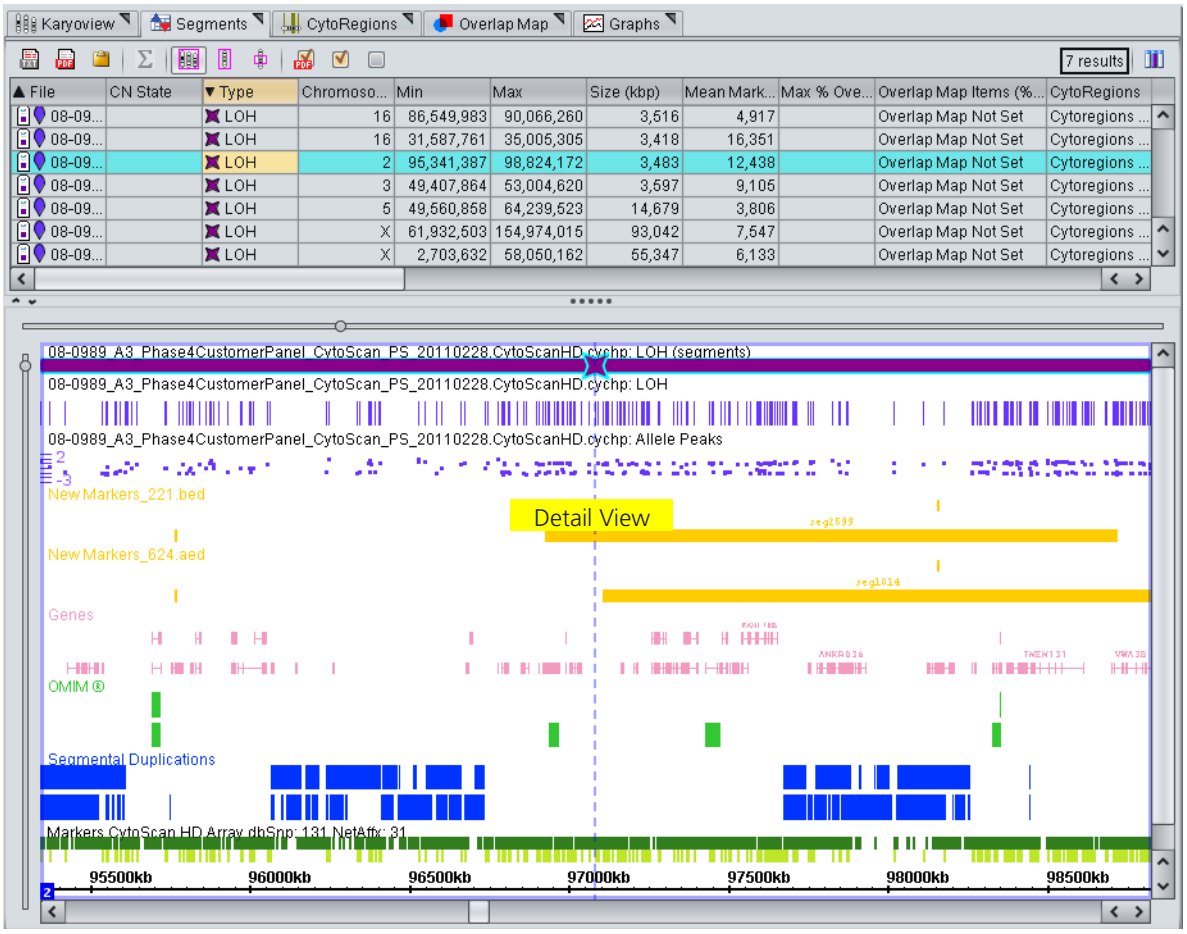
Materially Modified By	The current computer Operating System login ID and ChAS user profile name at the time that the segment was last materially modified.
Materially Modified Time	The time and date when the Segment was last materially modified.
Max % Overlap	The highest percentage by which some item(s) in the Overlap Map overlaps the segment. Segments completely overlapped by an Overlap Map item are 100% overlapped. This number is used for Filtering Segments out by "Overlap".
Overlap Map Items (% of Segment overlapped)	Item(s) in the Overlap Map which overlap the segment, followed by the percentage by which the segment is overlapped by that Item.
CytoRegions	Names of the CytoRegions with which the segment shares coordinates.
Use in Report	Allows manual selection of Segments for export to a Segments Table PDF, rather than all segments in the table.
Marker Count	Number of markers in the segment.
Cytoband Start	Cytoband in which the segment begins.
Cytoband End	Cytoband in which the segment ends.
Genes	List of RefSeq genes from the Genes track that share coordinates with the segment. Identically named gene isoforms are NOT repeated.
Gene Count	A count of the gene names listed in the Genes column
DGV	List of DGV variations that share coordinates with the segment.
FISH Clones	List of FISH clones that share coordinates with the segment.
sno/miRNA	List of sno/miRNA features that share coordinates with the segment.
BACs	List of BACs that share coordinates with the segment.
Old OMIM®	The column formerly known as "OMIM" in the ChAS software when NetAffxGenomicAnnotation Browser files version NA32.1 and earlier are loaded. This column's information is now out of date and has been superceded by the newer OMIM Genes and OMIM Phenotype Loci columns present in NA32.3 and above browser annotation files.
OMIM® Genes	List of OMIM Genes that share coordinates with the segment.
OMIM® Gene Count	A count of the OMIM Gene names listed in the OMIM Genes column.
OMIM® Phenotype Loci	List of OMIM Phenotype Loci that share coordinates with the segment.
Segmental Duplications	List of Segmental Duplications that share coordinates with the segment.
Smoothed/Joined	Indication that segment was created by smoothing or joining two or more segments in the initial segment detection.
Segment Label	A label comprised of the segment's Type, State, and Filename.
Segment Name/ID	File-specific identifier assigned to the detected segment.
Start Marker	The array marker name which marks the beginning of the segment.
End Marker	The array marker name which marks the end of the segment.
Preceding Marker	The array marker just above the segment in the data track used as input for the segment. Note: This column is only applicable to CNState Gain and Loss segments.
Preceding Marker Location	The coordinate location of the array marker just above the segment in the data track used as input for the segment. Note: This column is only applicable to CNState Gain and Loss segments.

Following Marker	The array marker just below the segment in the data track used as input for the segment. Note: This column is only applicable to CNState Gain and Loss segments.
Following Marker Location	The coordinate location of the array marker just below the segment in the data track used as input for the segment. Note: This column is only applicable to CNState Gain and Loss segments.
Mean Log2 Ratio	The mean of all the Log2 Ratio values contained in the segment.
Mean Weighted Log2 Ratio	The mean of all the Weighted Log2 Ratio values contained in the segment.
Median Log2 Ratio	The median of all the Log2 Ratio values contained in the segment.
Microarray Nomenclature	An ISCN-based description of the segment.
Sample UUID	Unique identifier for the CHP file.
Max % Coverage	The highest percentage by which a segment covers some item(s) in the Overlap Map.
Number of Overlap Map Items	Number of Overlap Map items which share genomic coordinates with the segment.
% of Overlaps Map Item covered by Segment	Overlap Map Item and the percentage by which it is covered by the segment.
Full Location	Chromosome Start and Stop in a user-friendly format for use in external databases.

Obtaining All Annotations Associated With a Segment

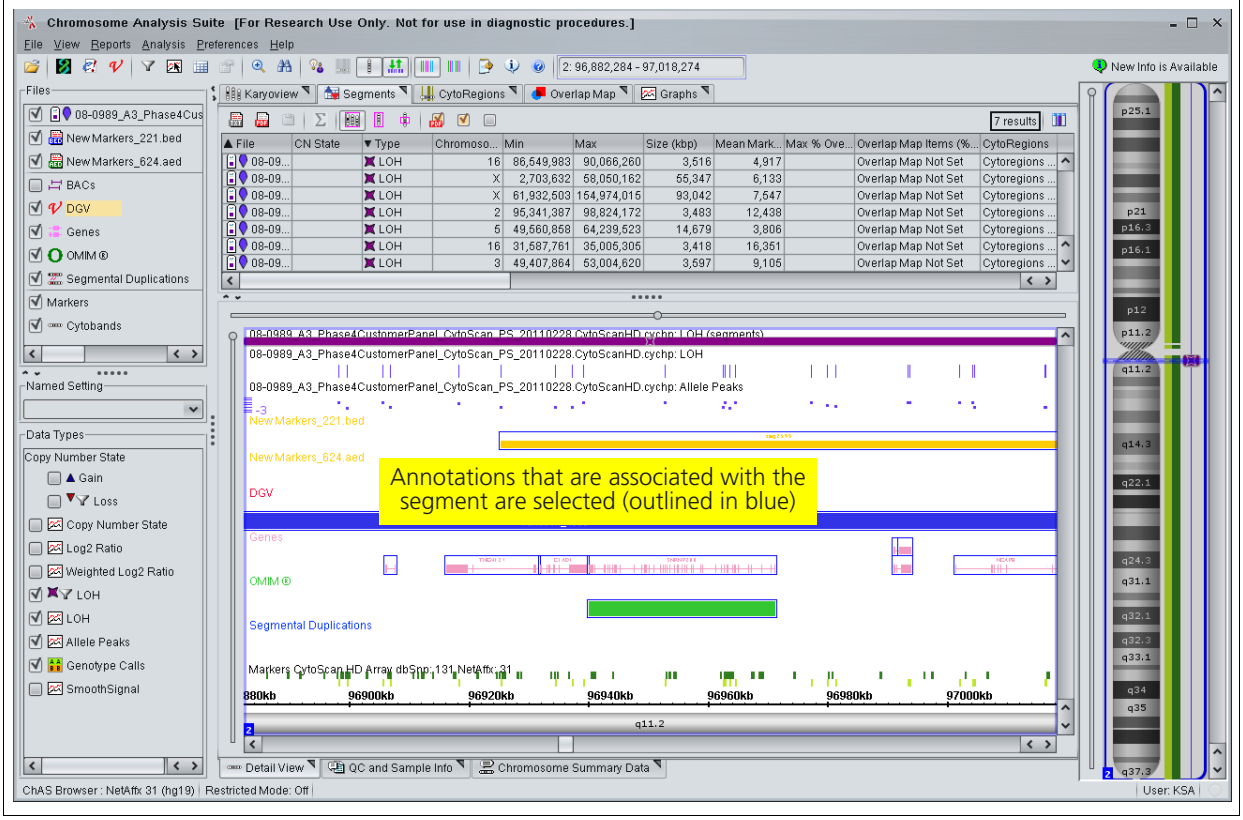
1. Select a segment in the Segments table, Karyoview, Detail View, or Selected Chromosome View, then press **Ctrl+Space** or right-click on the segment to select **Zoom to selection** from the menu. The exact length of the selected segment fills the entire width of the Detail View (Figure 13.13).

Figure 13.13 Detail View zoomed in on the segment selected in the Segments table



2. Click the tool bar button to expand all annotation tracks.

Figure 13.14 Detail View, annotations have been expanded and selected

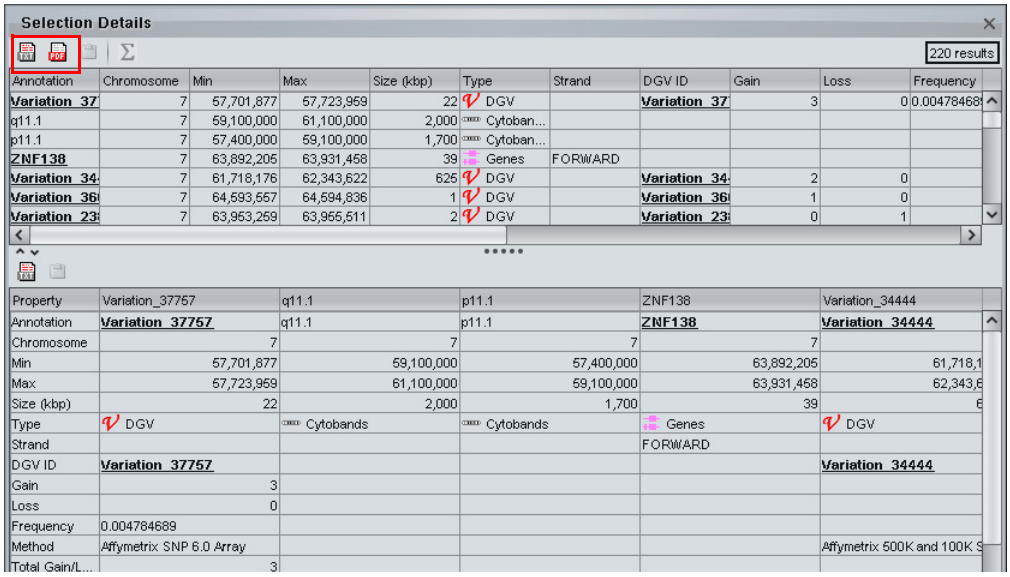


3. Using the mouse, draw a box around all of the genes and annotations of interest. When you release the mouse button, a blue box outlines the selected items in the Detail View. (Figure 13.14)
4. Click the tool bar button.

   The Selection Details table appears (Figure 13.15). It includes all of the items selected in the Detail View. For more details on the table, see *Selection Details Table* (on page 135).

Figure 13.15 Selection Details table

Click a button to export the information



Exporting Selected Segments to a PDF File

In the Segments table, you can select segments to include in a PDF report.

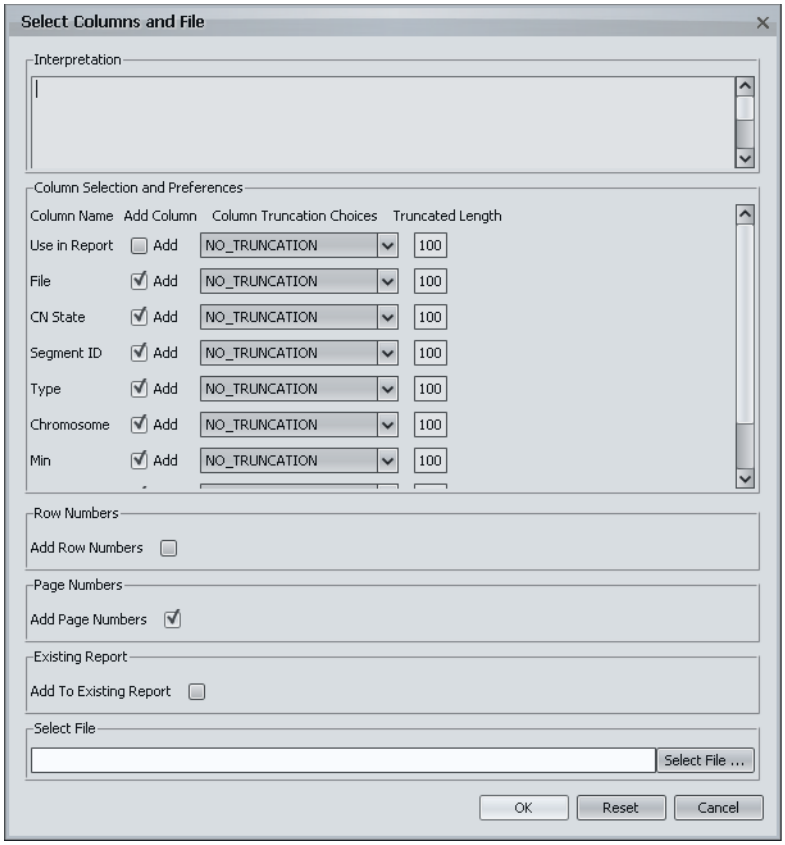
IMPORTANT: The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.

NOTE: You can also use the standard table export functions, described in *Export Table Data in a PDF File* (on page 275), to export data from the Segment table, but you CANNOT select specific segments for export using the standard export functions.

To export selected segments in a PDF file:

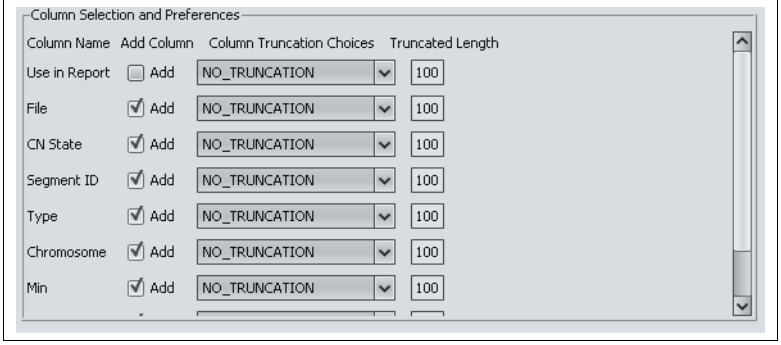
1. Select the columns for display. *See Saved Table States (on page 232)*.
   Limiting the columns of information will make the PDF easier to read.
2. Select and de-select the check boxes in the Use in Report column to select segments for export.
   You can:
   - Click the Check All button in the table tool bar to select all segments.
   - Click the Uncheck All button in the table tool bar to deselect all segments.
3. Click the Export Checked Rows button to export the checked segments to a PDF file.
   The Select Columns and File dialog box opens (Figure 13.16).

Figure 13.16 Select Columns and File dialog box



4. Select the columns to be displayed and the truncation rules for the content of the columns (Figure 13.17).

Figure 13.17 Column Selection and Preferences

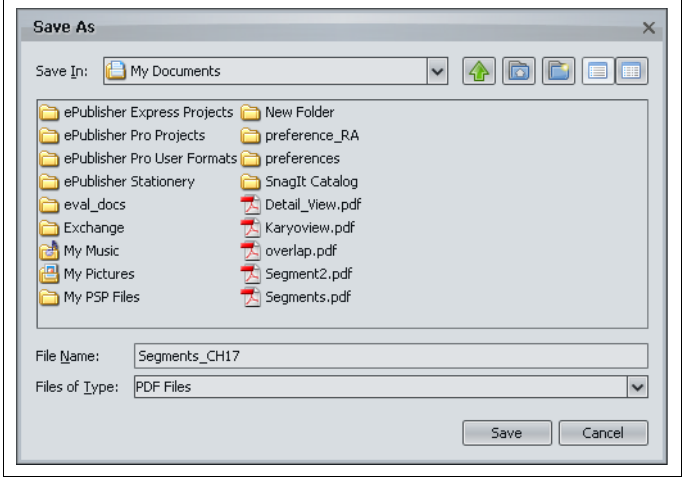


Column Name	Header of column in table.
Add Column	Select checkbox to display column in PDF.
Column Truncation Choices	Select Truncation Options: • No_Truncation: field is exported as is, using wrap-around if necessary. • Truncate_Beginning: Truncates content at the beginning of the field, leaving as many characters as specified in Truncated Length box. • Truncate_Middle: Truncates content in the middle of the field, leaving characters at the beginning and end, with ellipses (…) to mark the truncated characters. • Truncate_End: Truncates content at the end of the field, leaving as many characters as specified in Truncated Length box.
Truncated length	Number of characters displayed after truncating the data.

5. Select the options for adding row numbers and page numbers, if desired.
6. Select the option for adding to existing report, if desired. See *Combining PDFs into a Single Report* (on page 284).
7. Click **Select File…**

   The Save As dialog box opens (Figure 13.18).

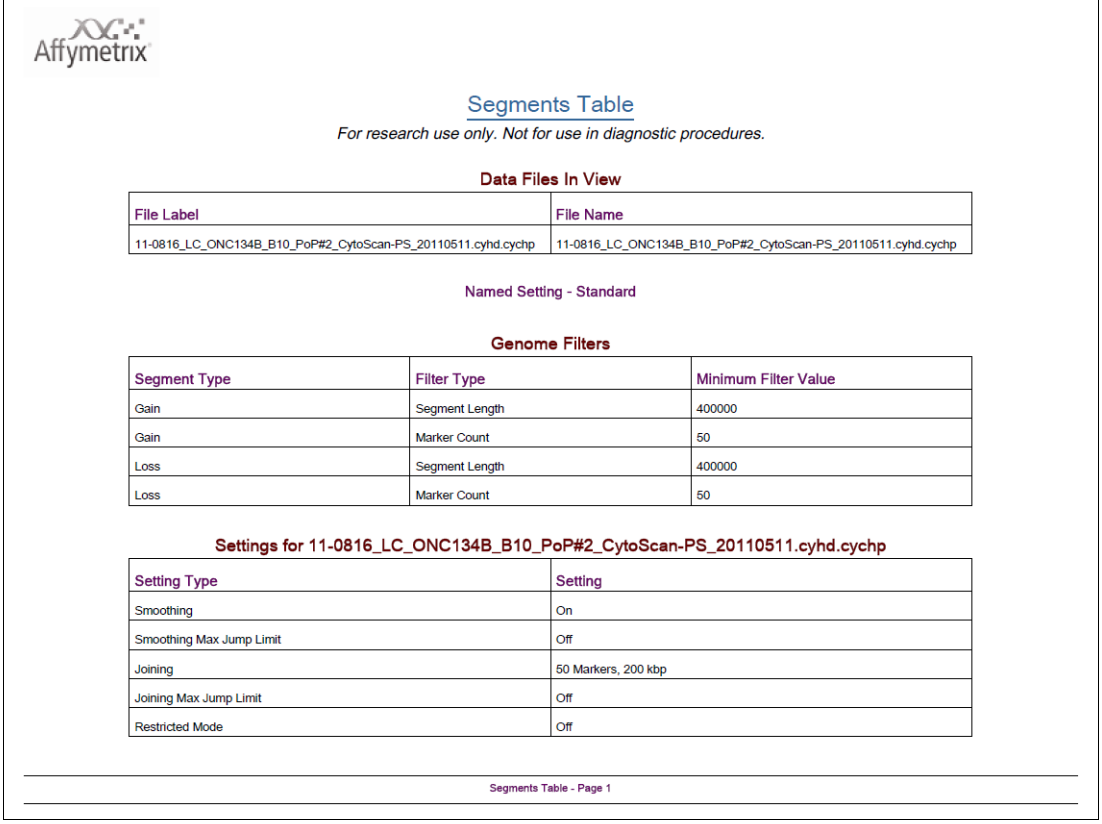
Figure 13.18 Save As dialog box



8. Select a folder location for the file using the navigation tools.
9. Enter a name for the PDF file, or select a file for the information to be appended to.
10. Click **Save** in the Save As dialog box.
11. Click **OK** in the Select Columns and File dialog box.

A PDF file is created with the selected segment data (Figure 13.19, Figure 13.20, Figure 13.21)

Figure 13.19 PDF file, page 1



Affymetrix

Segments Table

For research use only. Not for use in diagnostic procedures.

Data Files In View

File Label	File Name
11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_20110511.cyhd.cychp	11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_20110511.cyhd.cychp

Named Setting - Standard

Genome Filters

Segment Type	Filter Type	Minimum Filter Value
Gain	Segment Length	400000
Gain	Marker Count	50
Loss	Segment Length	400000
Loss	Marker Count	50

Settings for 11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_20110511.cyhd.cychp

Setting Type	Setting
Smoothing	On
Smoothing Max Jump Limit	Off
Joining	50 Markers, 200 kbp
Joining Max Jump Limit	Off
Restricted Mode	Off

Segments Table - Page 1

Figure 13.20 PDF file, page 2

Setting Type	Setting
QC Metrics	mapd ≤ 0.25; snpQC ≥ 15; wavinessSd ≤ 0.12

Microarray Nomenclature

11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_20110511.cyhd.cychp:
arr[hg19] 5p15.33p15.31(408,191-9,147,251)x3, 5p15.31p15.1(9,153,447-17,457,890)x3, 5p15.1p14.3(17,478,490-19,427,622)x3, 5p14.3p13.2(19,573,864-34,799,067)x3, 5p13.2(35,095,856-37,121,657)x3, 5p13.2(37,373,774-37,867,140)x3, 8p23.3p23.2(2,095,440-3,238,755)x1, 8p23.2p23.1(5,515,516-8,657,463)x1, 8p23.1(9,799,952-10,428,420)x1, 8p22p21.3(18,632,189-19,150,880)x1, 8p21.3(20,871,989-21,381,652)x1, 8p21.3(21,453,224-22,037,634)x1, 8p21.3(22,489,199-23,142,234)x1, 8p21.1p12(28,608,428-29,153,037)x1, 8p11.23p11.22(36,596,983-38,556,532)x3, 8p11.22(38,712,599-39,246,760)x3, 8p11.21(39,843,294-40,701,486)x3, 8p11.21(40,708,420-41,128,710)x3, 8p11.21(41,546,091-42,359,484)x3, 8p11.21(42,615,579-43,026,713)x3, 8p11.1q12.1(43,308,594-57,006,476)x3, 8q12.1(57,079,698-61,538,423)x3, 8q12.2q12.3(61,633,287-63,981,227)x3, 8q12.3q13.1(64,174,027-66,162,323)x3, 8q13.1q21.11(66,832,680-74,564,645)x3, 8q21.11(74,678,627-75,100,051)x3, 8q21.11q21.12(75,200,634-79,831,539)x3, 8q21.12q21.13(80,073,939-80,655,778)x3, 8q21.13(80,745,550-81,253,112)x3, 8q21.13q21.3(81,413,406-88,202,737)x3, 8q21.3(88,349,344-90,639,110)x3, 8q21.3q22.1(91,017,130-95,394,662)x3, 8q22.1q22.2(95,927,685-100,366,870)x3, 8q22.2(100,556,943-101,505,174)x3, 8q22.3(101,653,351-103,058,024)x3, 8q22.3(103,318,496-104,712,568)x3, 8q22.3q23.1(105,565,117-110,426,607)x3, 8q23.2q23.3(110,551,033-114,048,383)x3, 8q23.3(114,065,943-115,632,043)x3, 8q23.3(115,650,683-116,541,758)x3, 8q23.3q24.12(116,594,211-120,882,163)x3, 8q24.12(120,885,308-121,317,090)x3, 8q24.12q24.13(121,490,686-124,103,255)x3, 8q24.13q24.21(124,437,925-129,442,671)x3, 8q24.21q24.3(129,458,472-140,761,415)x3, 8q24.3(140,809,247-141,802,431)x3, 8q24.3(142,224,672-146,295,771)x3, 11q22.1q22.2(101,950,388-102,575,754)x1, 11q23.1(110,945,179-111,874,060)x1, 14q32.33(104,410,114-107,053,817)x3, 16p11.2p11.1(32,524,764-35,283,214)x1, 17q21.31(44,212,823-44,784,639)x3, 21q21.3(27,447,808-28,032,671)x3, 21q21.3q22.11(31,424,401-31,852,180)x3, 22q11.21(20,184,050-20,761,384)x3, 22q11.22(22,673,638-23,258,438)x3, 22q11.23q12.1(25,569,645-26,249,277)x3, 22q12.3(34,397,452-35,048,300)x3

Segments Details

Figure 13.21 PDF file, page 3

File	CN State	Type	Chromosome	Cytoband Start	Size (kbp)	Gene Count	Genes	OMIM ® Genes Count	OMIM ® Genes	CytoRegions	Call	Interpretation
11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_20110511.cyhd.cycp	3.0	Gain	5	p15.33	8739.06	45	AHRR, C5orf55, EXOC3, LOC25845, SLC9A3, CEP72, TPPP, ZDHHC11, BRD9, TRIP13, NKD2, SLC12A7, SLC6A19, SLC6A18, TERT, CLPTM1L, SLC6A3, LPCAT1, SDHAP3, LOC728613, MIR4277, MRPL36, NDUFS6, IRX4, IRX2, C5orf38, LOC285577, IRX1, LOC340094, ADAMTS16, KIAA0947, FLJ33360, MED10, UBE2QL1, LOC255167, NSUN2, SRD5A1, PAPD7, MIR4278, LOC442132, ADCY2, C5orf49, FASTKD3, MTRR, SEMA5A	0	N/A	Cytoregions Not Set		

Segments Table - Page 3

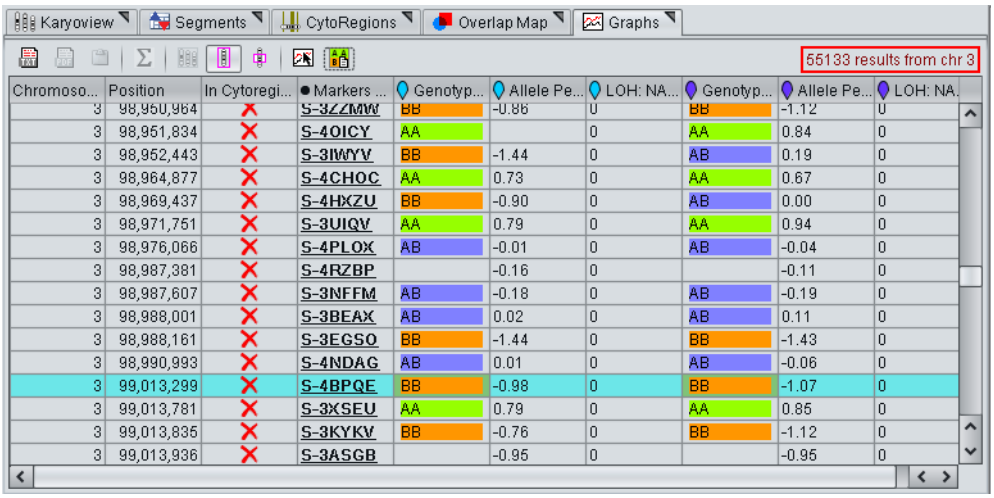
Graphs Table

The Graphs table displays the marker data used to create the graphs in the Detail view. Markers that are not used for the graphs currently displayed do not appear in this table. As in the Detail View, only markers from a single chromosome are displayed.

The column headings are colored according to the tracks used for the Karyoview, Selected Chromosome View, and Details View.

The Graphs table includes genotype SNP calls for CytoScan® HD and 750K Array results (Figure 13.22).

Figure 13.22 Example Graphs table with Genotype Calls for CytoScan® HD results



The table displays each data point in the displayed graphs.

To highlight markers in the views or the table:

- Click in a row of the table to place the cursor on the marker in the Chromosome and Detail Views.
- Click on a marker in the Selected Chromosome or Detail View to highlight the marker in the Graphs table.

The Graphs tab has the following elements:

- Tool bar
- Graph Data table

Tool bar

Figure 13.23 Graphs tool bar

The tool bar (Figure 13.23) allows you to:

- Export data in TXT format only. See *Export Tables as TXT File* (on page 280)
- View data for chromosome and selected chromosome region only. You cannot display data from the whole genome in Graphs tab. You can export this data for the whole genome using the *Exporting QC Tables and Probe-Level Data Using Generate Report* (on page 63).

- Export genotype results

The Graphs Settings button opens the Graph Settings panel, enabling you to change the style of graph, scale, and other features for the data graphs. See *Changing Graph Appearance* (on page 125).

Graphs Table

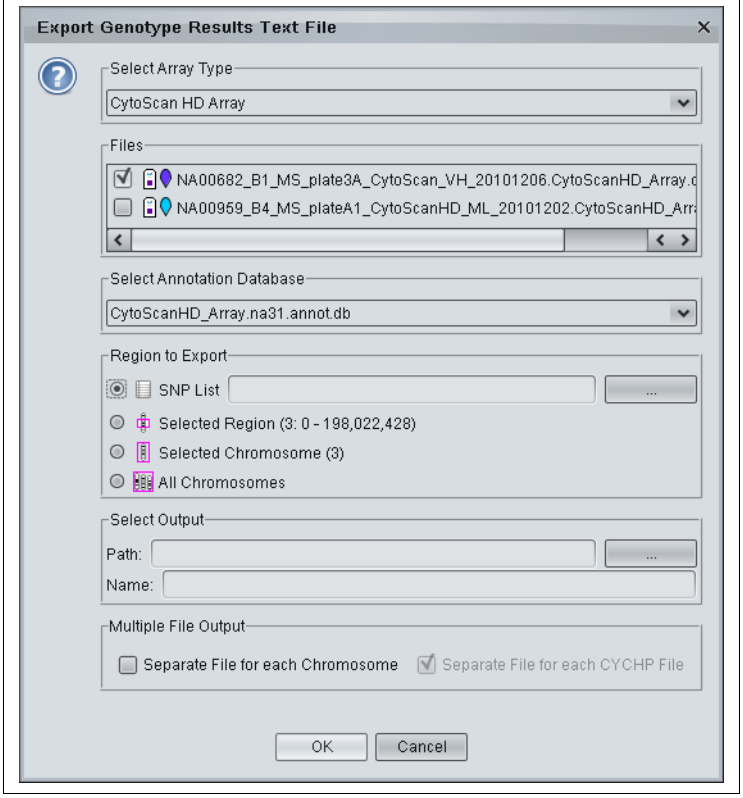
The Graphs table displays:

Chromosome	Chromosome the marker is located on.
Position	Position where the marker is located in the chromosome.
In Cytoregion	Whether marker is located in a cytoregion or not: In Cytoregion Not in Cytoregion. See *Using CytoRegions* (on page 178) for more information.
Markers	Marker ID. Right-click to link to NetAffx information about the marker. **Note: For efficiency reasons, it is not possible to sort the table on this column.**
Genotype	The SNP genotype call.
Graph Data types	 The Column header displays: • Color Nib • Data Type • Name of sample file The table cell displays the value for the marker.

To export genotype calls:

1. Click the tool bar button. Alternatively, select **Reports → Export Genotype Results Text File** from the menu bar.
2. In the dialog box that appears (Figure 13.24), select the array type, results file(s) (CYCHP), and annotation database to use for the export.

Figure 13.24 Export Genotype Results

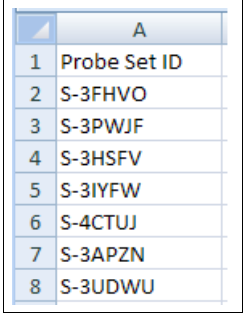


3. Specify the region to export:

Export Option	Exports Genotypes for:
SNP List	SNPs specified in a user-created SNP list (.txt).
Selected Region	SNPs included in the chromosome region selected in the Karyoview.
Selected Chromosome	SNPs on the chromosome selected in the Karyoview.
All Chromosomes	SNPs on all 24 chromosomes.

NOTE: The SNP list (Figure 13.25) should include one column header named "Probe Set ID" or "ProbeSet_ID" and one probe set name per row. The export will not proceed without one of these column headers.

Figure 13.25 Example SNP list



	A
1	Probe Set ID
2	S-3FHVO
3	S-3PWJF
4	S-3HSFV
5	S-3IYFW
6	S-4CTUJ
7	S-3APZN
8	S-3UDWU

4. Enter the path name or click the Browse button to select a folder for the output.
5. Enter a file name prefix. If only one output file is created (see below), this will be the file name. If multiple files are created, a suffix will be added to this string to create the file name. Do not include the file extension here.
6. Select a Multiple File Output option which determines if a separate file will be created for each chromosome and/or CYCHP file.

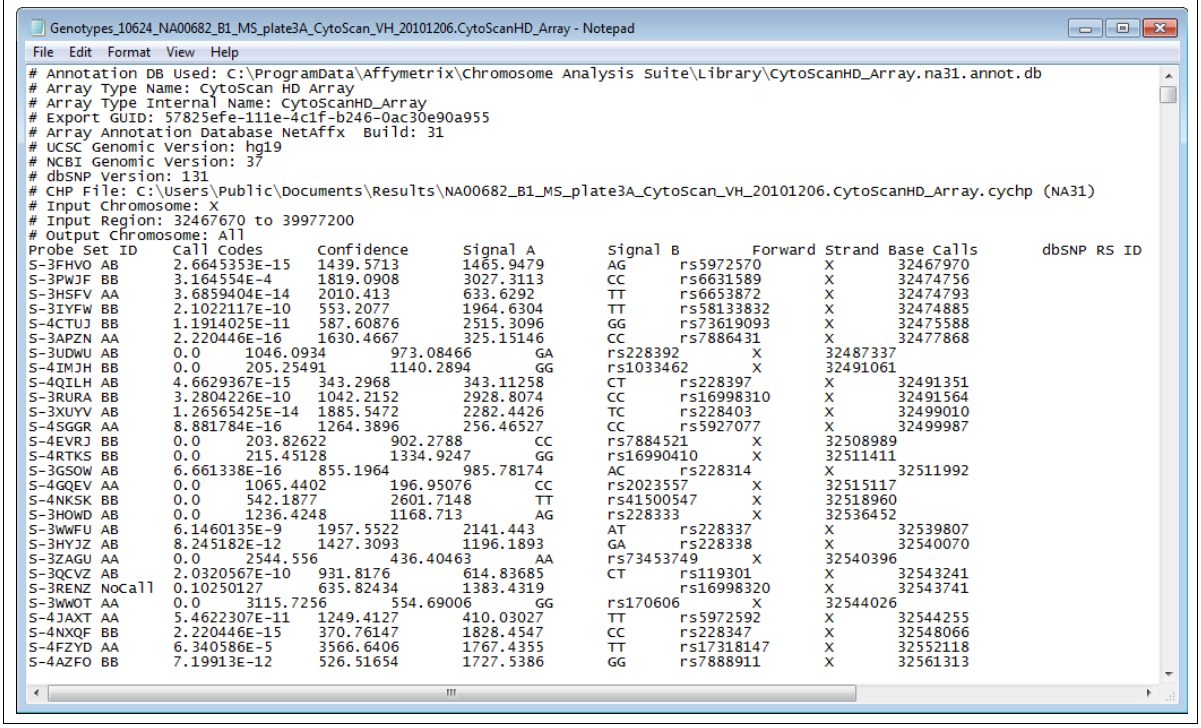
Table 13.1

Selected Output Option(s)	Files Created
None	One output file will be exported that contains all chromosome and all CYCHP file data. There will be separate data columns for each CYCHP file in the exported file.
Separate File for each Chromosome	Creates a separate file for each chromosome in the output data. If all chromosomes are selected, 24 files will be created. There will be separate data columns for each CYCHP file in the exported file.
Separate File for each CYCHP File	Creates one text file per CYCHP file. Each file contains genotype calls for all chromosomes.
Separate File for each Chromosome and Separate File for each CYCHP File	Create a separate file for each CYCHP file and for each chromosome. If three CYCHP files are selected and all chromosomes are reported on, this will create 72 files.

7. Click **OK**.

NOTE: Exporting of Genotypes may take several minutes to complete, as this process is dependant on the total number of SNPs selected for export.

Figure 13.26 Example of exported genotypes from a selected region of chromosome X



Genotypes_10624_NA00682_B1_MS_plate3A_CytoScan_VH_20101206.CytoScanHD_Array - Notepad

File Edit Format View Help

```
# Annotation DB Used: C:\ProgramData\Affymetrix\Chromosome Analysis Suite\Library\CytoScanHD_Array.na31.annot.db
# Array Type Name: CytoScan HD Array
# Array Type Internal Name: CytoScanHD_Array
# Export GUID: 57825efe-111e-4c1f-b246-0ac30e90a955
# Array Annotation Database NetAffx  Build: 31
# UCSC Genomic Version: hg19
# NCBI Genomic Version: 37
# dbSNP Version: 131
# CHP File: C:\Users\Public\Documents\Results\NA00682_B1_MS_plate3A_CytoScan_VH_20101206.CytoScanHD_Array.cychp (NA31)
# Input Chromosome: X
# Input Region: 32467670 to 39977200
# Output Chromosome: All
Probe Set ID    Call Codes      Confidence      Signal A        Signal B        Forward Strand Base Calls       dbSNP RS ID
S-3FHVO AB      2.6645353E-15   1439.5713       1465.9479       AG      rs5972570       X       32467970
S-3PWJF BB      3.164554E-4     1819.0908       3027.3113       CC      rs6631589       X       32474756
S-3HSFV AA      3.6859404E-14   2010.413        633.6292        TT      rs6653872       X       32474793
S-3IYFW BB      2.1022117E-10   553.2077        1964.6304       TT      rs58133832      X       32474885
S-4CTUJ BB      1.1914025E-11   587.60876       2515.3096       GG      rs73619093      X       32475588
S-3APZN AA      2.220446E-16    1630.4667       325.15146       CC      rs7886431       X       32477868
S-3UDWU AB      0.0     1046.0934       973.08466       GA      rs228392        X       32487337
S-4IMJH BB      0.0     205.25491       1140.2894       GG      rs1033462       X       32491061
S-4QILH AB      4.6629367E-15   343.2968        343.11258       CT      rs228397        X       32491351
S-3RURA BB      3.2804226E-10   1042.2152       2928.8074       CC      rs16998310      X       32491564
S-3XUYV AB      1.26565425E-14  1885.5472       2282.4426       TC      rs228403        X       32499010
S-4SGGR AA      8.881784E-16    1264.3896       256.46527       CC      rs5927077       X       32499987
S-4EVRJ BB      0.0     203.82622       902.2788        CC      rs7884521       X       32508989
S-4RTKS BB      0.0     215.45128       1334.9247       GG      rs16990410      X       32511411
S-3GSOW AB      6.661338E-16    855.1964        985.78174       AC      rs228314        X       32511992
S-4GQEV AA      0.0     1065.4402       196.95076       CC      rs2023557       X       32515117
S-4NKSK BB      0.0     542.1877        2601.7148       TT      rs41500547      X       32518960
S-3HOWD AB      0.0     1236.4248       1168.713        AG      rs228333        X       32536452
S-3WWFU AB      6.1460135E-9    1957.5522       2141.443        AT      rs228337        X       32539807
S-3HYJZ AB      8.245182E-12    1427.3093       1196.1893       GA      rs228338        X       32540070
S-3ZAGU AA      0.0     2544.556        436.40463       AA      rs73453749      X       32540396
S-3QCVZ AB      2.0320567E-10   931.8176        614.83685       CT      rs119301        X       32543241
S-3RENZ NoCall  0.10250127      635.82434       1383.4319               rs16998320      X       32543741
S-3WWOT AA      0.0     3115.7256       554.69006       GG      rs170606        X       32544026
S-4JAXT AA      5.4622307E-11   1249.4127       410.03027       TT      rs5972592       X       32544255
S-4NXQF BB      2.220446E-15    370.76147       1828.4547       CC      rs228347        X       32548066
S-4FZYD AA      6.340586E-5     3566.6406       1767.4355       TT      rs17318147      X       32552118
S-4AZFO BB      7.19913E-12     526.51654       1727.5386       GG      rs7888911       X       32561313
```

The exported text file (Figure 13.26) includes information about the analysis (for example, array type, NetAffx annotation database, hg version, and chromosome).

NOTE: If the option "Separate File for each CYCHP File" was not selected, many of the headers will be repeated for each CYCHP file. The header titles will be appended with the CYCHP file name to indicate which file the column belongs to.

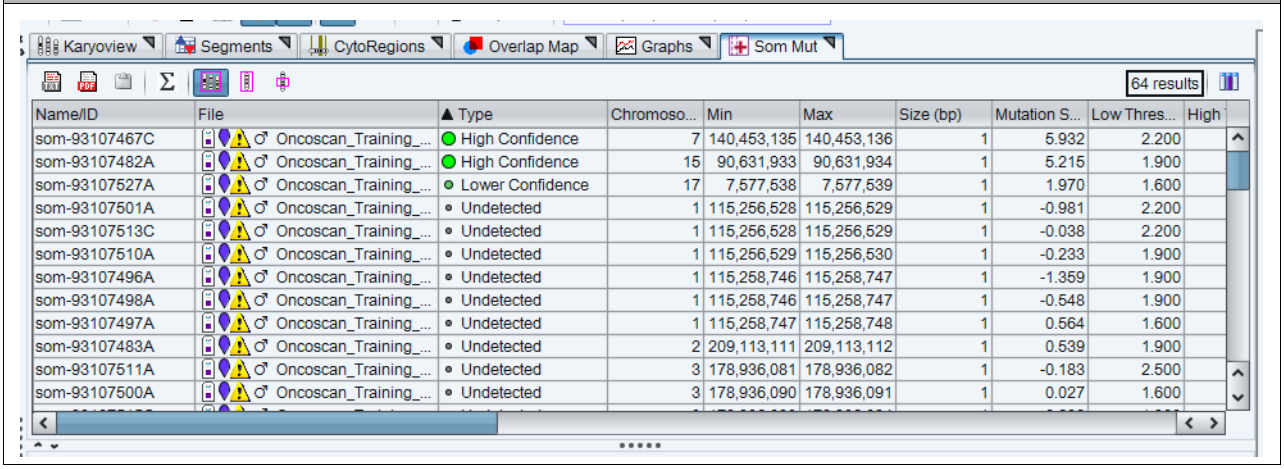
The column headers report the following information:

Column Header	Description
Probe Set ID	Probe set identifier
Call Codes	Genotype call for the SNP.
Confidence	Confidence value for the call.
Signal A	Raw signal value for Signal A on the probe set.
Signal B	Raw signal value for Signal B on the probe set.
Forward Strand Base Calls	Base calls for the forward strand.
dbSNP RS ID	dbSNP RS ID value
Chromosome	Chromosome associated with the probe sets.
Chromosomal Position	Chromosome position of the SNP.

Somatic Mutations Table

The Somatic Mutations table displays the somatic mutation data and annotations from OncoScan FFPE Assay OSCHPs only. (Figure 13.27)

Figure 13.27 Somatic Mutations Table



Name/ID	File	Type	Chromoso...	Min	Max	Size (bp)	Mutation S...	Low Thres...	High
som-93107467C	Oncoscan_Training_...	High Confidence	7	140,453,135	140,453,136	1	5.932	2.200	
som-93107482A	Oncoscan_Training_...	High Confidence	15	90,631,933	90,631,934	1	5.215	1.900	
som-93107527A	Oncoscan_Training_...	Lower Confidence	17	7,577,538	7,577,539	1	1.970	1.600	
som-93107501A	Oncoscan_Training_...	Undetected	1	115,256,528	115,256,529	1	-0.981	2.200	
som-93107513C	Oncoscan_Training_...	Undetected	1	115,256,528	115,256,529	1	-0.038	2.200	
som-93107510A	Oncoscan_Training_...	Undetected	1	115,256,529	115,256,530	1	-0.233	1.900	
som-93107496A	Oncoscan_Training_...	Undetected	1	115,258,746	115,258,747	1	-1.359	1.900	
som-93107498A	Oncoscan_Training_...	Undetected	1	115,258,746	115,258,747	1	-0.548	1.900	
som-93107497A	Oncoscan_Training_...	Undetected	1	115,258,747	115,258,748	1	0.564	1.600	
som-93107483A	Oncoscan_Training_...	Undetected	2	209,113,111	209,113,112	1	0.539	1.900	
som-93107511A	Oncoscan_Training_...	Undetected	3	178,936,081	178,936,082	1	-0.183	2.500	
som-93107500A	Oncoscan_Training_...	Undetected	3	178,936,090	178,936,091	1	0.027	1.600	

The Somatic Mutation Table Components include:

- Tool bar (below)
- Somatic Mutation Table on page 254

To highlight segments in the views or in the table:

- Click the row of the table to zoom to the somatic mutation in the Karyoview, Selected Chromosome View and Detail View.
- Click on the somatic mutation in the Karyoview, Selected Chromosome View or Detail View to highlight the somatic mutation in the Somatic Mutations Table.
  - Somatic Mutations with a High Condifidence call will appear as a large, bright green circle in the graphic views.
  - Somatic Mutations with a Lower Condifidence call will appear as a smaller, dark green circle in the graphic views.
  - Undetected Somatic Mutation calls remain small black dots in the grapic views.

NOTE: Calls can be removed from the table (and graph view) by deselecting the Undetected, Lower Confidence, High Confidence check box(es) in the Data Types window.

Tool Bar

The Tool bar (Figure 13.28) provides quick access to table functions. Its standard functions are described in *The Standard Tool Bar Controls* (on page 228).

Figure 13.28 Tool bar

Somatic Mutation Table

The table can display each somatic mutation with the following information (the default set of columns in a new user profile may include only a subset of the total columns listed).

File	Name of the OSCHP file the somatic mutation is in.
High Threshold	High confidence MutScore threshold. Measurements equal to or greater than this threshold are called "High confidence," describing the likelihood that the mutation is present.
Low Threshold	Lower confidence MutScore threshold. Measurements with a MutScore below this value are called "Undetected". Measurements equal to or greater than this threshold but less than the High Threshold are called "Lower confidence," describing the likelihood that the mutation is present.
Size (bp)	The size of the Somatic mutation in base pairs.
Type	HighConfidence, LowerConfidence or Undetected call as to the presence of the mutation.
MutScore	(measured quantile normalized signal - median signal for this marker in the reference model file) / (95th percentile signal for this marker in the reference model file - median signal for this marker in the reference
Name/ID	Affymetrix identifier for the marker.
Channel	CEL file from which the signal is measured. "A" is the AT CEL, "C" is the GC CEL.
ProbeSetType	ProbesetType is Somatic Mutation.
CommonName	Abbreviated description of the mutations to which this ProbeSet is known to respond. The name has the form [Gene]:[amino acid change for mutation]:[cDNA change for mutation]. In the event that the ProbeSet cannot differentiate among multiple mutations to which it can respond, the slash (/) delimits the multiple known mutations.
Event	Describes if the probeset is a point mutation, deletion, insertion or sequence variant.
Event Type	Describes if the event is missense, frame-shift, in-frame insertion or deletion.
COSMIC ID	The identifier of the mutation as listed in the COSMIC database, which is a catalogue of somatic mutations in cancer. More information on these mutations can be found at http://cancer.sanger.ac.uk
Genes	RefSeq gene that shares coordinates with the somatic mutation.
Mutation Syntax (aa)	Encoding of which nucleotide was changed and its location in the CDS.
Mutation Syntax (cds)	Encoding of which amino acid was changed and the location of the codon.
Chromosome	Chromosome on which the somatic mutation is found.
Min	Start position of the somatic mutation.
Max	Stop position of the somatic mutation.
Mutation (nt)	Wild type > mutant nucleotide change on the coding strand.
Mutation (aa)	Wild type > mutant amino acid change on the coding strand.

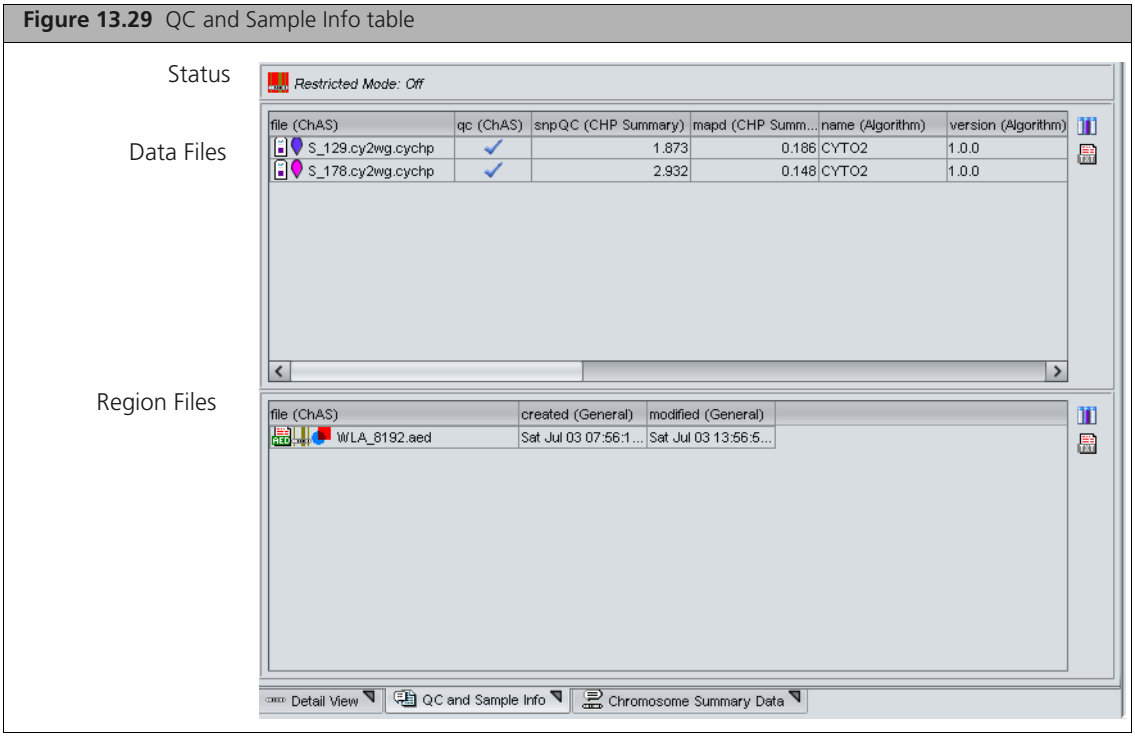
Strand	Coding strand of the associated gene (Forward or Reverse).
Source DB	Cosmic database version.
accession number	RefSeq Gene accession number.
Fwd 5' flank	Sequence flanking the mutation at the 5' end.
Fwd 3' flank	Sequence flanking the mutation at the 3' end.

NOTE: Changes made to an OSCHP file in the Somatic Mutation Viewer Application will require the OSCHP file to be reloaded into the ChAS Browser to reflect the change made to the sample.

QC and Sample Info Tab

The QC and Sample Info tab (Figure 13.29) in the lower pane displays information about the loaded Data and Region files.

Figure 13.29 QC and Sample Info table



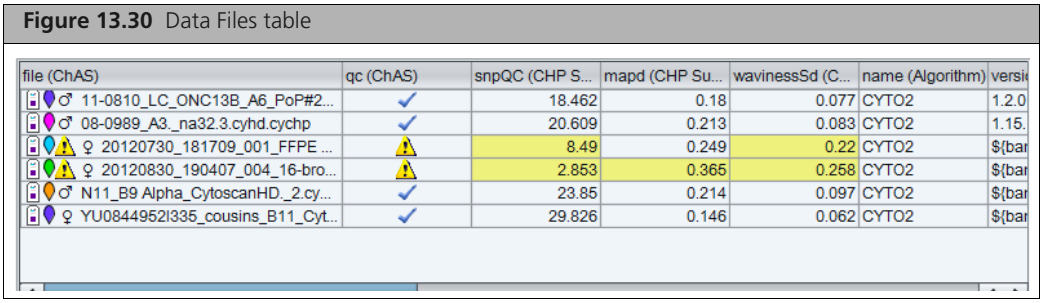
The top section displays Status for Restricted Mode (see *Using Restricted Mode* (on page 186))

The tables display information on:

- Loaded data files. *See QC and Sample Information Table (on page 256)*
- Loaded region files. *See Loaded AED/BED Files Table (on page 257)*

QC and Sample Information Table

Figure 13.30 Data Files table



file (ChAS)	qc (ChAS)	snpQC (CHP S...	mapd (CHP Su...	wavinessSd (C...	name (Algorithm)	versi
♂ 11-0810_LC_ONC13B_A6_PoP#2...	✓	18.462	0.18	0.077	CYTO2	1.2.0
♂ 08-0989_A3._na32.3.cyhd.cychp	✓	20.609	0.213	0.083	CYTO2	1.15.
♀ 20120730_181709_001_FFPE ...	⚠	8.49	0.249	0.22	CYTO2	${bar
♀ 20120830_190407_004_16-bro...	⚠	2.853	0.365	0.258	CYTO2	${bar
♂ N11_B9 Alpha_CytoscanHD_.2.cy...	✓	23.85	0.214	0.097	CYTO2	${bar
♀ YU0844952I335_cousins_B11_Cyt...	✓	29.826	0.146	0.062	CYTO2	${bar

The table (Figure 13.30 on page 256) displays the following data by default.

File	File Name.
QC	In or out of QC bounds.
Failed QC	Specific metrics that have failed QC within a failing sample row are identified with a yellow background.
SNP QC	SNP QC score
MAPD	Median Absolute Pairwise Difference score.
wavinessSd	Waviness standard deviation (Waviness SD) is a global measure of variation of microarray probes that is insensitive to short-range variation and focuses on long-range variation.
Name	Algorithm name
Version	Algorithm version
Sex Determination	Gender call for the sample. See *Gender Call Algorithms* (on page 256)).
AutosomeGenomeLOH	The proportion of LOH on chromosome 1 to 22.
GenomeLOH	The proportion of LOH on chromosomes 1 to 22, chromosome X, and chromosome Y.
Array type	Type of array used in analysis.
Parameter File	Name of the chasparam file used to create the CYCHP file
Annotation File	Name of the Annotation file used to create the CYCHP file
Reference File	Reference file used in the single sample analysis
Created	Date file was created.
Modified	Date file was last modified.
Sample ID	Sample identifier in the ARR sample file.

Other data from the header of the Sample Data file or the ARR file can also be selected for display in the Select Columns dialog box.

You can only hide or display columns by using the Column Select dialog box, at the right of the table.

Gender Call Algorithms

The table below explains which algorithm is used to make the gender calls for the different arrays.

The CytoScan® HD Array uses the call “Y-gender” which gives a male/female call.

Depending on the version they were created under, various GTC 2.x and 3.x SNP6 CNCHP files use other gender calls present in their CNCHP file header.

These calls used from the CNCHP file header are NOT the same gender calls used for those files in GTC, since the GTC-displayed gender calls were stored in GQC or CN_SEGMENTS files which are not supported in ChAS.

NOTE: For more details how the array-specific algorithms call LOH segments for the X or Y chromosome, see *LOH Segments on the X and Y Chromosomes* (on page 47).

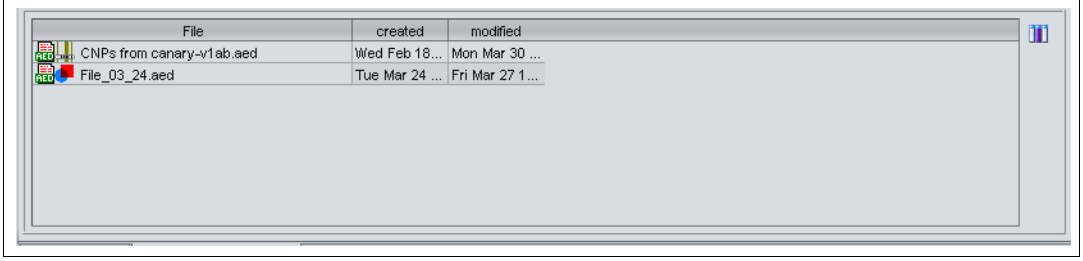
Table 13.2 Gender Call Algorithims

Software/Array Type	Gender Call Algorithm	Call	Gender Call Confidence
ChAS 1.0x, 1.1x, ChAS 1.2, 1.2x ChAS 2.0, 2.1/ CytoScan® HD Array	Y-gender	male/female	yes
GTC 3.0 to GTC 4.1/ Genome-Wide SNP Array 6.0	affymetrix-chipsummary-Gender	male/female/unknown	no
GTC 2.1/ Genome-Wide SNP Array 6.0	affymetrix-chipsummary-hasY	male/female	no
GTC 2.0/ Genome-Wide SNP Array 6.0	affymetrix-chipsummary-hasY	male/female	no
OncoScan Console 1.0/ Oncoscan Assay Array	Y-gender	male/female	yes

Loaded AED/BED Files Table

This table (Figure 13.31) displays information about the loaded Region (AED or BED) files.

Figure 13.31 AED/BED Files information



The Region information files section displays information on:

File File Name with Icons displayed if selected as Overlap File or CytoRegions File).

Created Date and time file was created.

Modified Date and time file was last modified.

You can only hide or display columns by using the Column Select dialog box, at the right of the table.

Chromosome Summary Data

The Chromosome Summary Data Tab has 2 components:

- Chromosome Summary Data
- Auto-Generated Autosome LOH Percentage

The Chromosome Summary Data table (Figure 13.32) summarizes particular data across each chromosome in the loaded sample data files.

The available data types are:

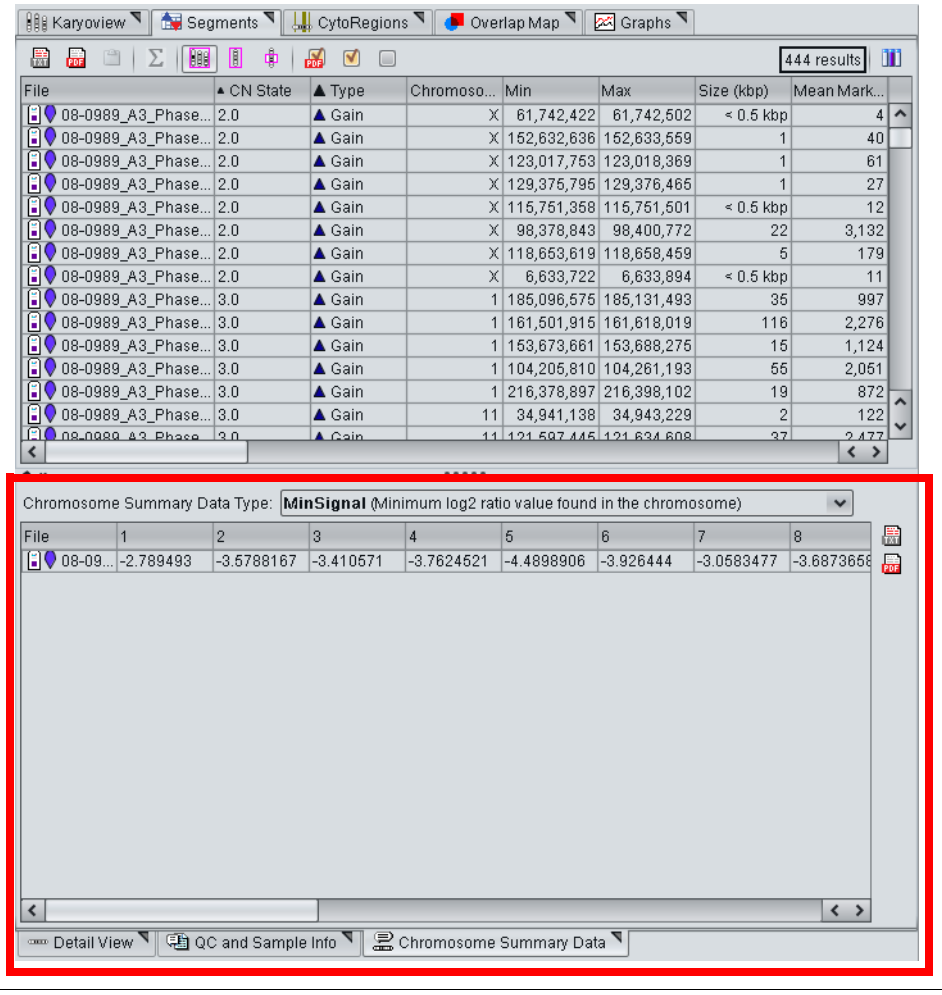
- Min Signal – minimum Log2 Ratio value found in the chromosome
- Median Signal – median Log2 Ratio value found in the chromosome
- Max Signal – maximum Log2 Ratio value found in the chromosome
- Median CN State – median calibrated Log2 Ratio
- Mosaicism – median mosaicism mixture value

NOTE: The mosaicism CN state value is not an integer due to cell populations with different CN state values. In the Chromosome Summary Data table, the mosaicism value indicates how much the median CN state value is above or below two, the normal CN state value for autosomes and X in females. For example, a median mosaicism mixture CN state value of 2.48 is displayed as 0.48 in the Chromosome Summary table.

NOTE: Mosaicism (median mosaicism mixture value) for normal males is -1.0 for chromosome X and -1.0 for chromosome Y. Mosaicism for normal females is 0.0 for chromosome X and -2.0 for chromosome Y. A mosaic XO female is treated the same as a mosaic autosomal monosomy (i.e., the mosaicism level of chromosome X will be between -1.0 and 0). A mosaic XXY male is also given a mosaicism level between -1 and 0 for chromosome X.

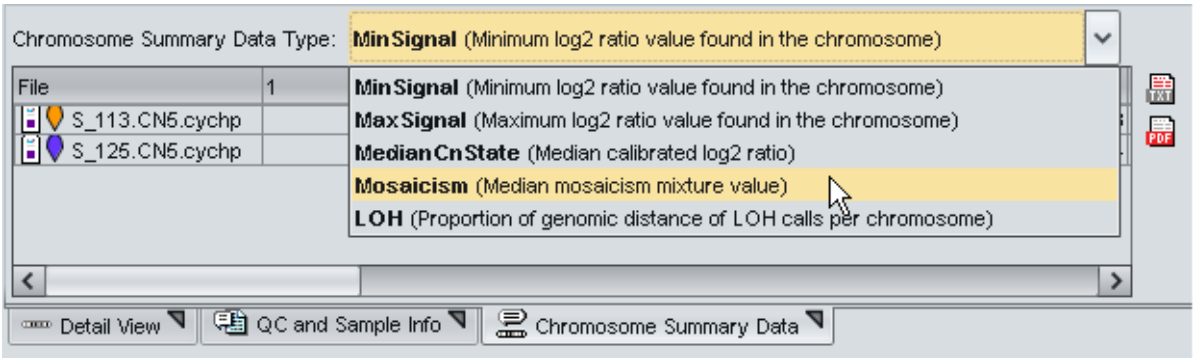
- LOH – proportion of genomic distance of LOH calls per chromosome

Figure 13.32 Chromosome Summary Data, Min Signal in each chromosome



To choose the data type, make a selection from the drop-down list. (Figure 13.33)

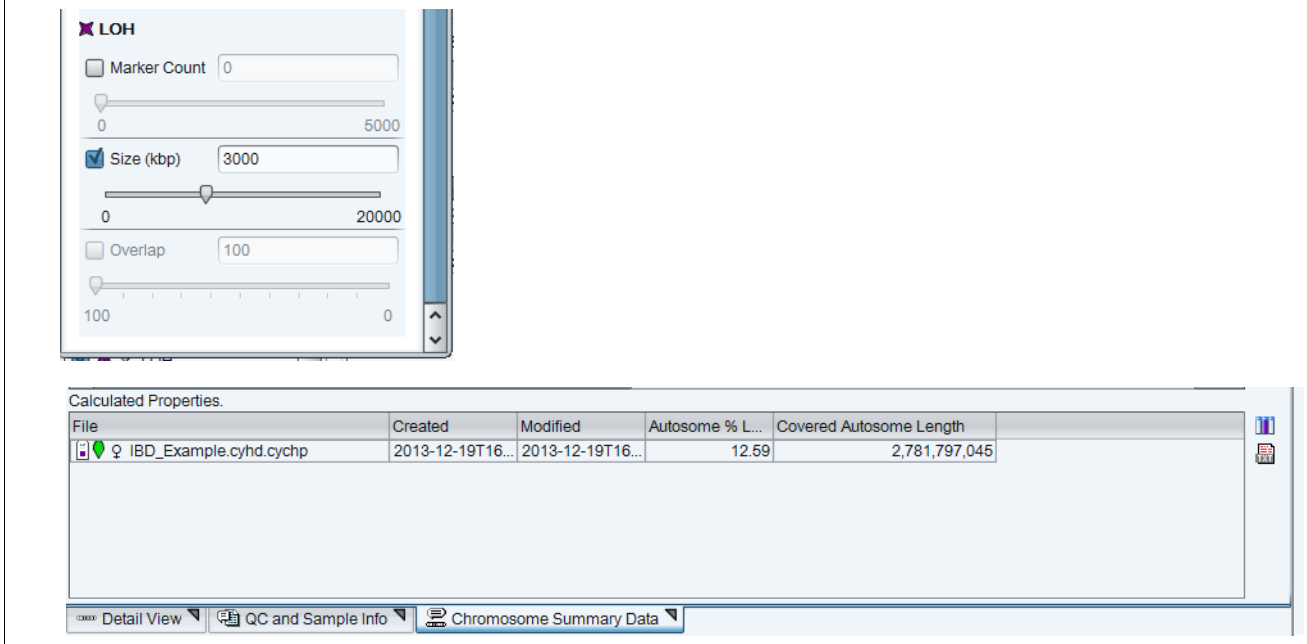
Figure 13.33 Selecting the data type for the Chromosome Summary table



Auto-Generated Autosome LOH Percentage

The percentage LOH displayed is calculated for the autosomes based on the filter size set for LOH in the Filters Window. (Figure 13.34).

Figure 13.34 Custom Autosomal Genome LOH Percentage



NOTE: You must check the LOH Segment Data type to view the sample's percent LOH.

Searching Results

The Search function allows you to search:

- Detected Segments
- Reference Annotations
- Loaded Region Information Files

The search can find:

- Names of Reference Annotations
- BED and AED file elements, including those in files designated as CytoRegions or Overlap Maps
- Loaded and displayed segments

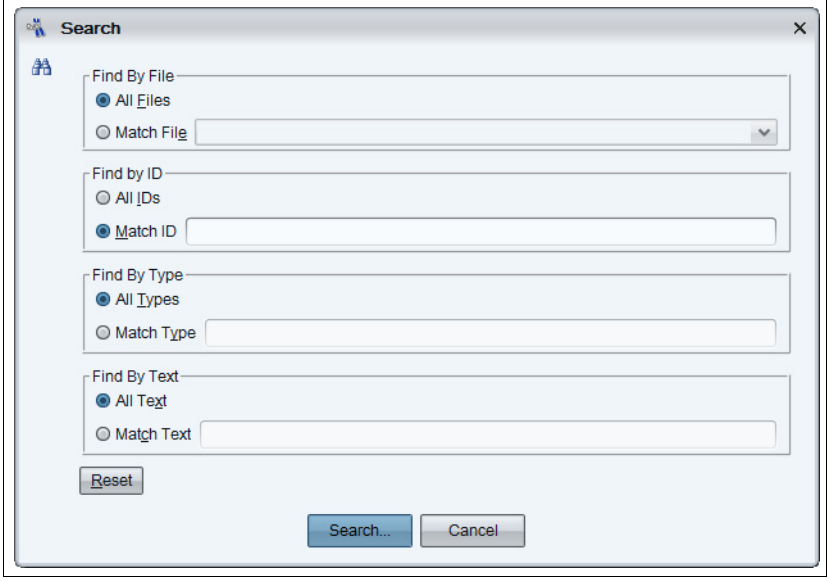
You can search by:

- File (select the files to be searched)
- ID Label
- Type

To perform a search:

1. From the View menu, select **Search** or click the upper tool bar's icon.
   The Search window opens. (Figure 13.35 on page 261)

Figure 13.35 Search dialog box



2. Search all files in the File List.
   Alternatively, click **Match File** and select a file to be searched from the drop-down list.

NOTE: Only files that are check marked in the Files List appear in the Match File drop-down list. Only files that are check marked in the Files list will be searched.

3. Select Match ID and enter an ID (not case sensitive) for the search.
   You can use the "*" character as a wild card for the search.
4. Select Match Type and enter a type name (not case sensitive) for the search.

NOTE: You can use the asterisk symbol * as a wildcard. Using the wildcard is very useful when you want to search for the type "OMIM ®" because the symbol ® does not need to be entered. Another convenient way is to simply choose the OMIM file from the "Find By" File list.

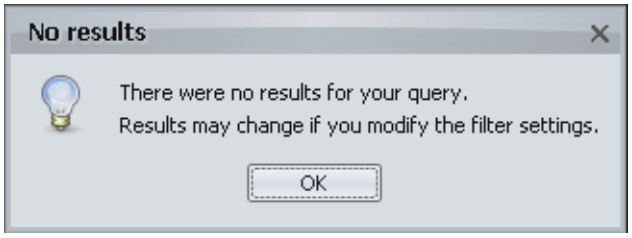
You can enter:

- Names for types of reference annotation features (Genes, DGV, etc.)
- Names for types of segments (Loss, Gain, etc.)

Click **Reset** to clear the search entries.

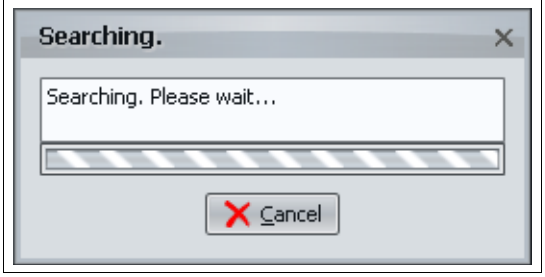
5. Click **Search…** to begin the search.
   If no results are found, the following notice appears (Figure 13.36).

Figure 13.36 No results notice



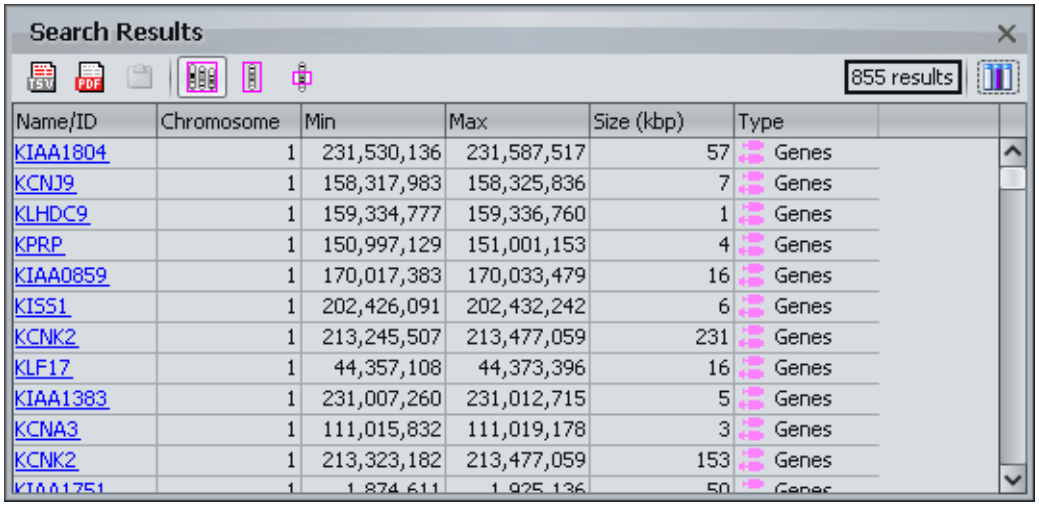
If the search takes more than a few seconds, an In Progress notice appears. (Figure 13.37)

Figure 13.37 Searching notice



If results are found, the Search Results table opens. (Figure 13.38)

Figure 13.38 Search Results

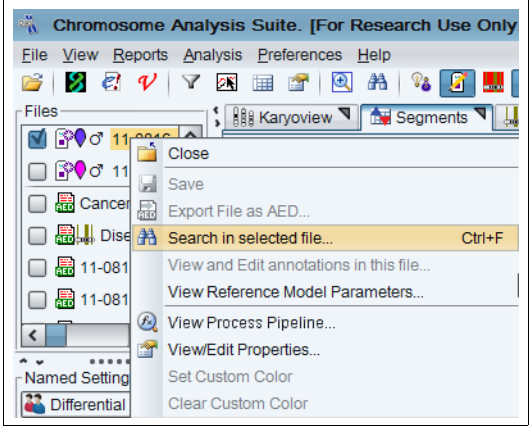


Name/ID	Chromosome	Min	Max	Size (kbp)	Type
KIAA1804	1	231,530,136	231,587,517	57	Genes
KCNJ9	1	158,317,983	158,325,836	7	Genes
KLHDC9	1	159,334,777	159,336,760	1	Genes
KPRP	1	150,997,129	151,001,153	4	Genes
KIAA0859	1	170,017,383	170,033,479	16	Genes
KISS1	1	202,426,091	202,432,242	6	Genes
KCNK2	1	213,245,507	213,477,059	231	Genes
KLF17	1	44,357,108	44,373,396	16	Genes
KIAA1383	1	231,007,260	231,012,715	5	Genes
KCNA3	1	111,015,832	111,019,178	3	Genes
KCNK2	1	213,323,182	213,477,059	153	Genes
KIAA1751	1	1,874,611	1,925,136	50	Genes

To perform a search in a selected file:

1. Right-click on a selected (checked) file inside the **Files** list pane.

Figure 13.39 Searching notice

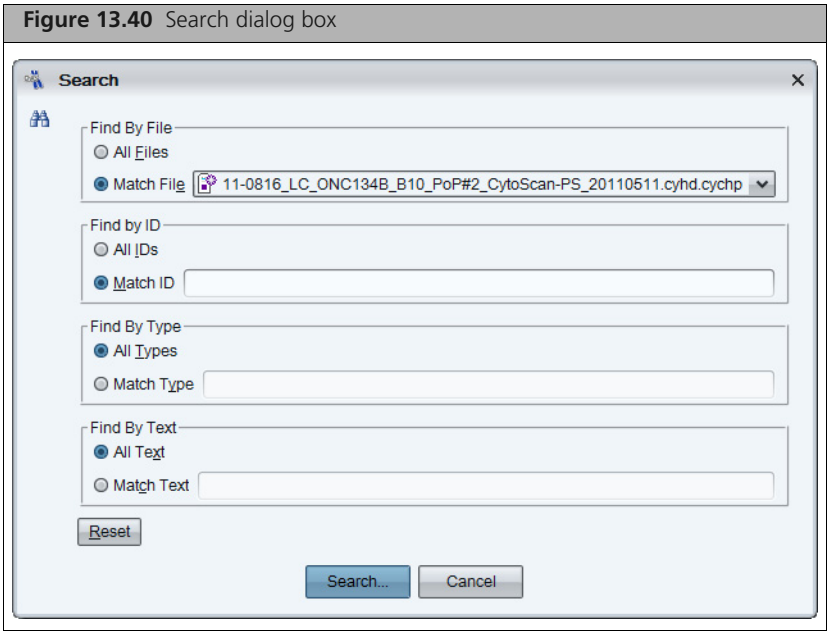


2. Click to select **Search in selected file...** (Figure 13.39)

   The Search window opens.

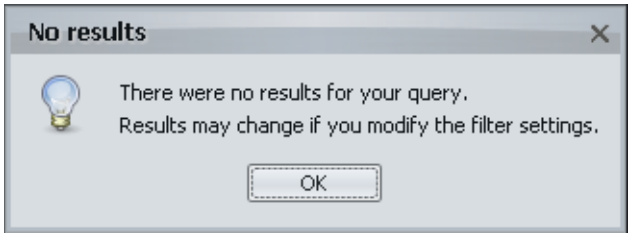
3. Click the appropriate radio button(s), enter your search criteria, then click **Search**.

Figure 13.40 Search dialog box



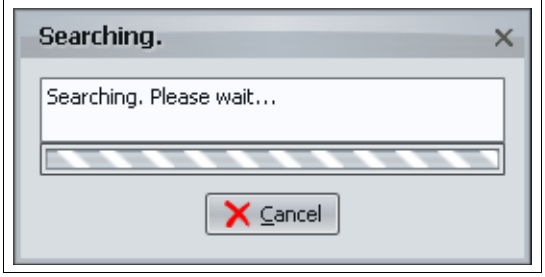
If no results are found, the following notice appears (Figure 13.41).

Figure 13.41 No results notice



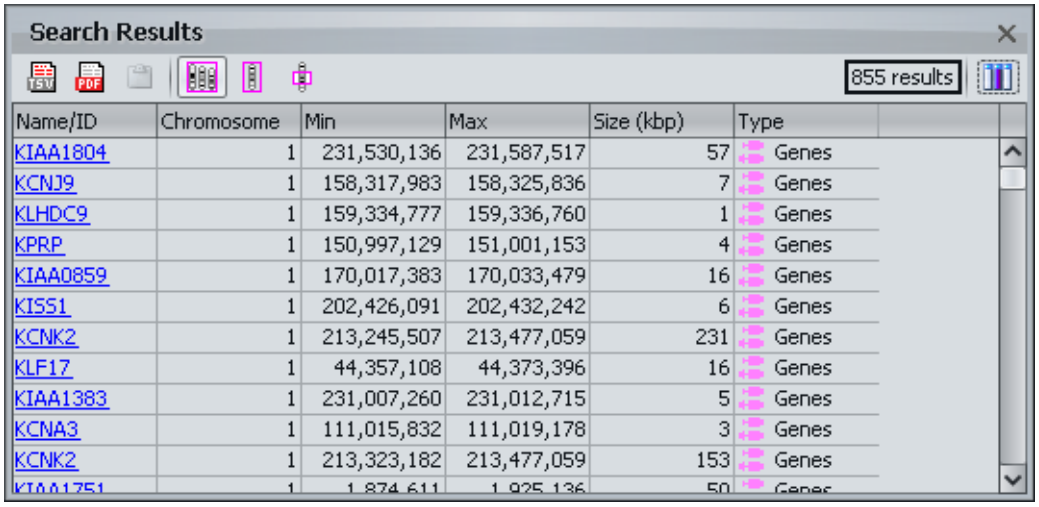
If the search takes more than a few seconds, an In Progress notice appears. (Figure 13.42)

Figure 13.42 Searching notice



If results are found, the Search Results table opens. (Figure 13.43 on page 264)

Figure 13.43 Search Results



Name/ID	Chromosome	Min	Max	Size (kbp)	Type
KIAA1804	1	231,530,136	231,587,517	57	Genes
KCNJ9	1	158,317,983	158,325,836	7	Genes
KLHDC9	1	159,334,777	159,336,760	1	Genes
KPRP	1	150,997,129	151,001,153	4	Genes
KIAA0859	1	170,017,383	170,033,479	16	Genes
KISS1	1	202,426,091	202,432,242	6	Genes
KCNK2	1	213,245,507	213,477,059	231	Genes
KLF17	1	44,357,108	44,373,396	16	Genes
KIAA1383	1	231,007,260	231,012,715	5	Genes
KCNA3	1	111,015,832	111,019,178	3	Genes
KCNK2	1	213,323,182	213,477,059	153	Genes
KIAA1751	1	1,874,611	1,925,136	50	Genes

To highlight features in the views or the table:

- Double-click in a row of the table to zoom to the feature in the Selected Chromosome and Detail Views.
- Click on a feature in the Selected Chromosome or Detail View to highlight the feature in the Search Results table (the feature must be listed in the table to be highlighted).

You can perform the common table operations in the Search Results table (see *Common Table Operations* (on page 228)).

The table displays the following information for each row in the Search Results:

Label	Name or ID of the Item.
Chromosome	Chromosome the items are located in.
Min	Starting position of the item.
Max	Ending position of the item. For all segments, the segment start coordinates are always lower by one bp from the coordinate for the starting probe of the segment as reported in the graphs table while the end coordinate matches the coordinate for the ending probe as reported in the graphs table (see Appendix E, *Genomic Position Coordinates* on page 313).
Size (kbp)	Size of the item.
Type	Type of Item.

Finding Intersections

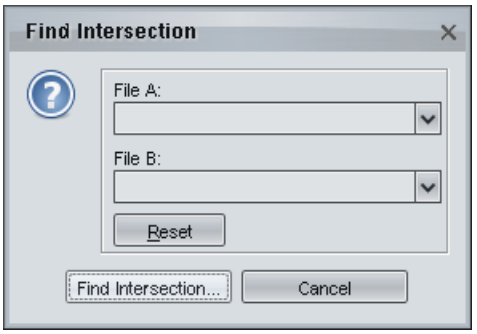
The Find Intersection function enables you to find segments and regions that overlap for different:

- Detected Segments
- Reference Annotations
- Loaded Region Information Files

To find the intersection of items in selected files:

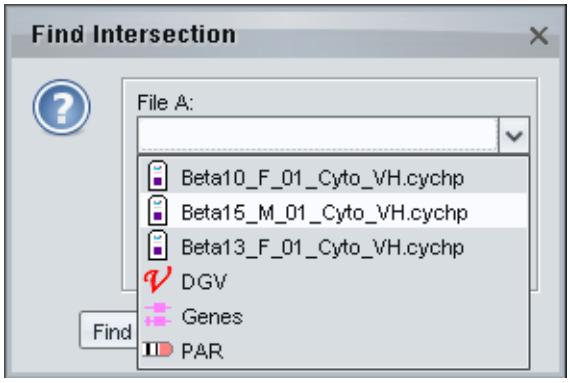
1. From the View menu, select **Find Intersections**…
   The Find Intersection dialog box opens (Figure 13.44).

Figure 13.44 Find Intersection dialog box



2. Select the first file for the comparison from the File A drop-down list (Figure 13.45).

Figure 13.45 Drop-down list



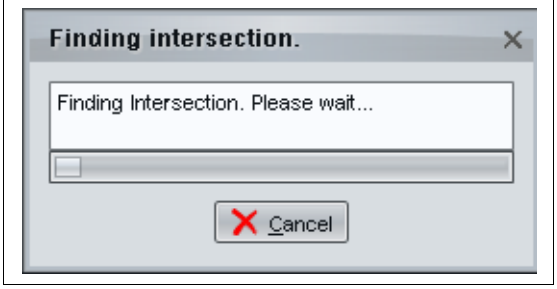
The list shows the available Sample files, Region Information Files, and Reference Annotations.

NOTE: Only files that are check marked in the Files List appear in the Match File drop-down list.

3. Select the second file from the File B drop-down list.
4. Click **Find Intersection…**

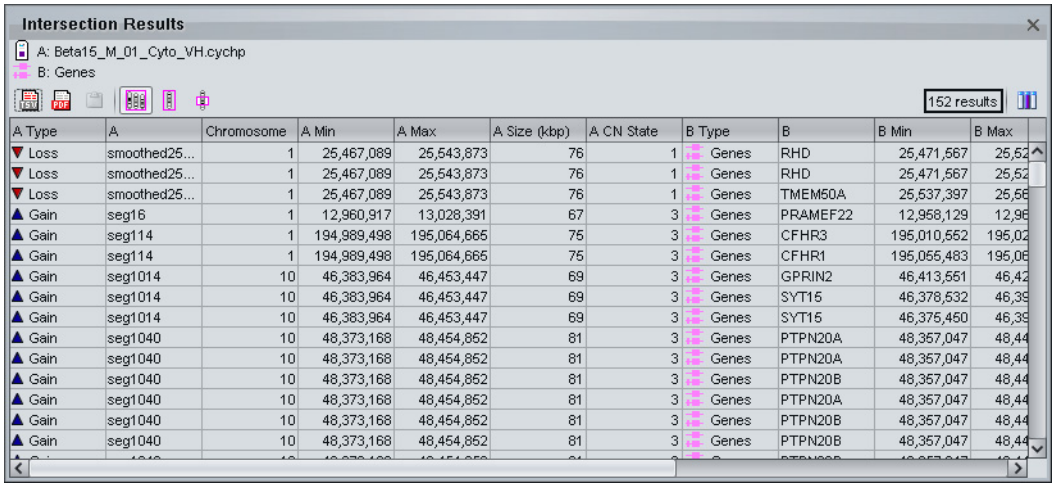
   The Finding Intersection notice opens (Figure 13.46).

Figure 13.46 Finding Intersection notice



When the comparison is finished, the Intersection Results table opens (Figure 13.47).

Figure 13.47 Intersection Results table



A Type	A	Chromosome	A Min	A Max	A Size (kbp)	A CN State	B Type	B	B Min	B Max
Loss	smoothed25...	1	25,467,089	25,543,873	76	1	Genes	RHD	25,471,567	25,52
Loss	smoothed25...	1	25,467,089	25,543,873	76	1	Genes	RHD	25,471,567	25,52
Loss	smoothed25...	1	25,467,089	25,543,873	76	1	Genes	TMEM50A	25,537,397	25,56
Gain	seg16	1	12,960,917	13,028,391	67	3	Genes	PRAMEF22	12,958,129	12,96
Gain	seg114	1	194,989,498	195,064,665	75	3	Genes	CFHR3	195,010,552	195,02
Gain	seg114	1	194,989,498	195,064,665	75	3	Genes	CFHR1	195,055,483	195,06
Gain	seg1014	10	46,383,964	46,453,447	69	3	Genes	GPRIN2	46,413,551	46,42
Gain	seg1014	10	46,383,964	46,453,447	69	3	Genes	SYT15	46,378,532	46,39
Gain	seg1014	10	46,383,964	46,453,447	69	3	Genes	SYT15	46,375,450	46,39
Gain	seg1040	10	48,373,168	48,454,852	81	3	Genes	PTPN20A	48,357,047	48,44
Gain	seg1040	10	48,373,168	48,454,852	81	3	Genes	PTPN20A	48,357,047	48,44
Gain	seg1040	10	48,373,168	48,454,852	81	3	Genes	PTPN20B	48,357,047	48,44
Gain	seg1040	10	48,373,168	48,454,852	81	3	Genes	PTPN20A	48,357,047	48,44
Gain	seg1040	10	48,373,168	48,454,852	81	3	Genes	PTPN20B	48,357,047	48,44
Gain	seg1040	10	48,373,168	48,454,852	81	3	Genes	PTPN20B	48,357,047	48,44

The table displays the names of the A and B files above the tool bar.

To highlight features in the views or the table:

- Double-click in a row of the table to zoom to the feature for File A in the Selected Chromosome and Detail Views.
- Click on a feature in the Selected Chromosome or Detail View to highlight the feature in the Intersection Results table (the feature must be listed in the table to be highlighted).

You can perform the common table operations in the Intersection Results table (see *Common Table Operations* (on page 228)).

The table displays the following information for each overlapping set of items:

A Type — Type of item in A file with overlap.

A — Identifier used for item in A file.

Chromosome — Chromosome the items are located in.

A Min — Starting position of the A item.

A Max — Ending position of the A item.
For all segments, the segment start coordinates are always lower by one bp from the coordinate for the starting probe of the segment as reported in the graphs table while the end coordinate matches the coordinate for the ending probe as reported in the graphs table (see Appendix E, *Genomic Position Coordinates* on page 313).

A Size (kbp) — Size of the A item.

A CN State — Copy number of the segment in file A.

B Type — Type of Item in B file with overlap.

B — Identifier used for item in B file.

B Min — Starting position of the B item.

B Max	Ending position of the B item. For all segments, the segment start coordinates are always lower by one bp from the coordinate for the starting probe of the segment as reported in the graphs table while the end coordinate matches the coordinate for the ending probe as reported in the graphs table (see Appendix E, *Genomic Position Coordinates* on page 313)
B Size (kpb)	Size of the B item.
B CN State	Copy number of the segment in file B.
% A Touching B	How much of the A item is covered by the B item.
% B Touching A	How much of the B item is covered by the A item.
Shared Size (kbp)	Size of the overlap.

Reporting Results

IMPORTANT: Edit Mode must be OFF before attempting to make a PDF Report.

Chromosome Analysis Suite has tools for reporting results. It provides options for:

- *Exporting Graphic Views* (on page 269): Export the Karyoview, Selected Chromosome View, and Detail View as a PDF or PNG file.
- *Exporting Table Data* (on page 274): Export the table data as a PDF or TXT file, or copy selected data to the clipboard.
- *Combining PDFs into a Single Report* (on page 284)

IMPORTANT: The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.

NOTE: If you have trouble displaying non-English characters on screen or in exported PDF files, make sure that the font "Arial Unicode MS" is installed on the machine. Look in "C:\Windows\Fonts" for the file ARIALUNI.TTF.

Exporting Graphic Views

You can export the Karyoview, Selected Chromosome View, and Detail View as:

- PDF reports. *See Exporting as PDF Reports (on page 270)*
- PNG graphic file. *Export as PNG* (on page 273)

Exporting as PDF Reports

ChAS provides a variety of options for exporting graphic views as PDFs. The PDF Report displays the graphic with basic information about data files and settings (Figure 14.1, Figure 14.2).

IMPORTANT: The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.

Figure 14.1 Whole screen exported to PDF, p 1

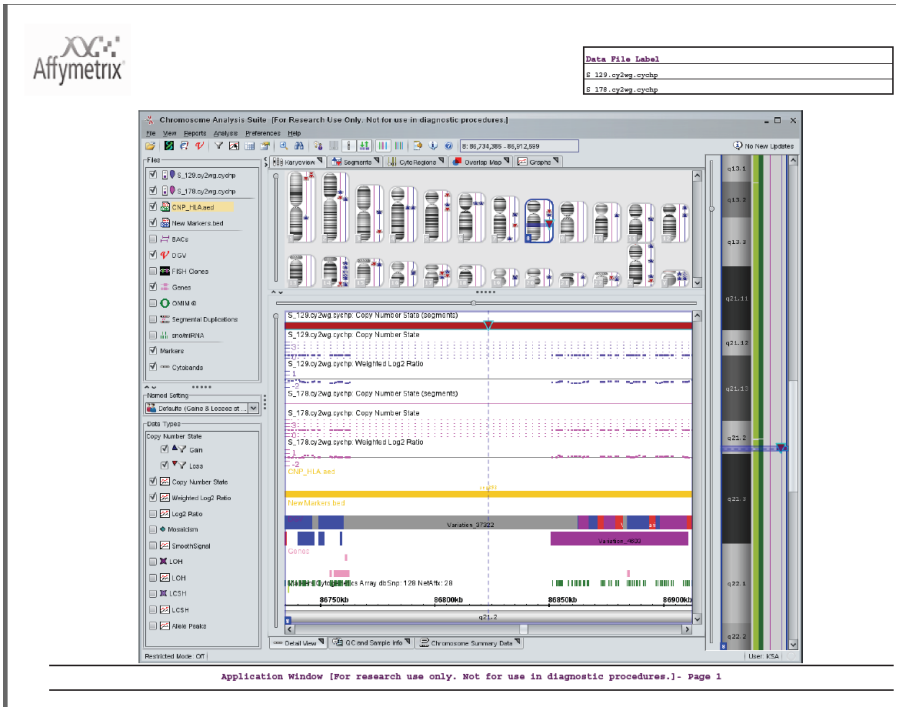


Figure 14.2 Whole screen exported to PDF, p 2

Data Files In View

File Label	File Name
S_129.cy2wg.cychp	S_129.cy2wg.cychp
S_178.cy2wg.cychp	S_178.cy2wg.cychp

Named Setting - Defaults (Gains & Losses at 100kbp, 50 markers)

Genome Filters

Segment Type	Filter Type	Minimum Filter Value
Gain	Segment Length	100000
Gain	Marker Count	50
Loss	Segment Length	100000
Loss	Marker Count	50

Settings for S_129.cy2wg.cychp

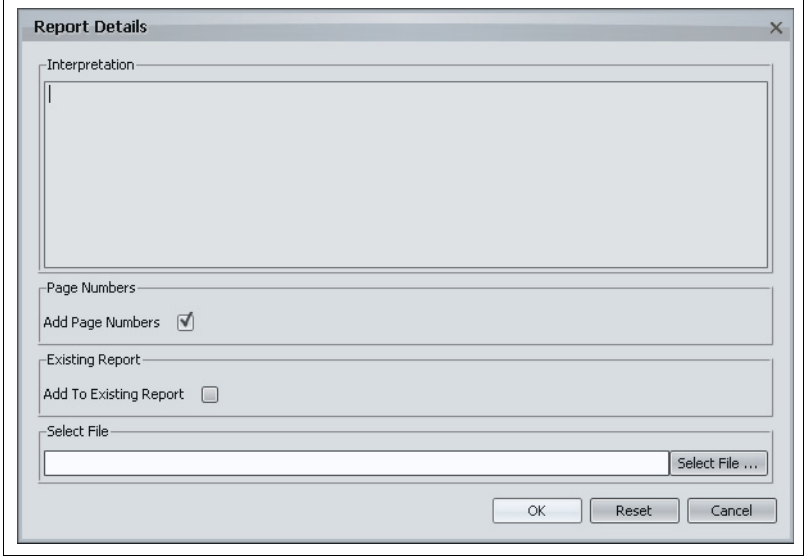
Setting Type	Setting
Smoothing	On
Smoothing Max Jump Limit	Off
Joining	50 markers
Joining Max Jump Limit	Off
Restricted Mode	Off

To export graphics information:

1. From the **Reports** menu, select the PDF option you want to use.
   - Export application window PDF - Creates PDF with entire software screen and information about data files
   - Export Karyoview PDF - Creates PDF with Karyoview and information about data files.
   - Export Selected Chromosome PDF - Creates PDF with Selected Chromosome View and information about data files.
   - Export Detail View PDF - Creates PDF with selected Detail View and information about data files.
   - Export Segments Table PDF - Creates PDF with Segment Table
   - Export QC and Sample Info PDF - Creates PDF with the QC and Sample Information
   - Export Chromosome Summary Data PDF - Creates PDF with the Chromosome Summary Data

   The Report Details dialog box appears (Figure 14.3).

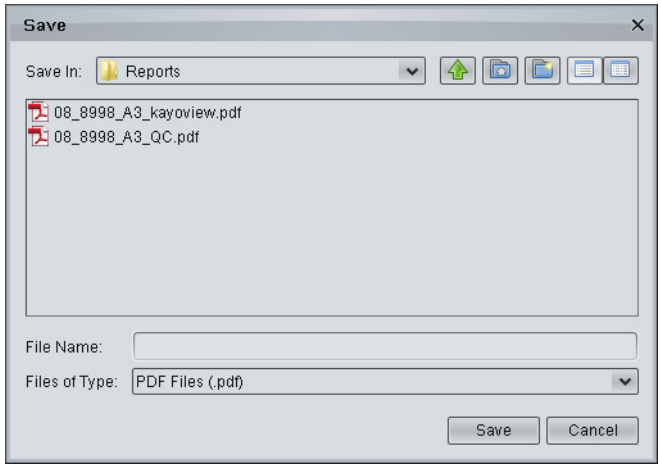
Figure 14.3 Report Details dialog box



2. Enter text into Interpretation box, if desired.
3. Select the option for adding page numbers, if desired.
4. Select the option for adding to an existing report, if desired. See *Combining PDFs into a Single Report* (on page 284).
5. Click **Select File**.

   The Save dialog box opens (Figure 14.4).

Figure 14.4 Save dialog box



6. Select a folder location for the file using the navigation tools.

   This folder location is automatically selected when exporting other reports during a session.

7. Enter a name for the PDF file.

   If you are adding the graphic to an existing PDF, select the PDF file.

8. Click **Save**.

9. Click **OK** in the Report Details dialog box

   A PDF report is created with the selected data type saved (Figure 14.5, Figure 14.6, Figure 14.7).

Figure 14.5 Detail View report, p 1

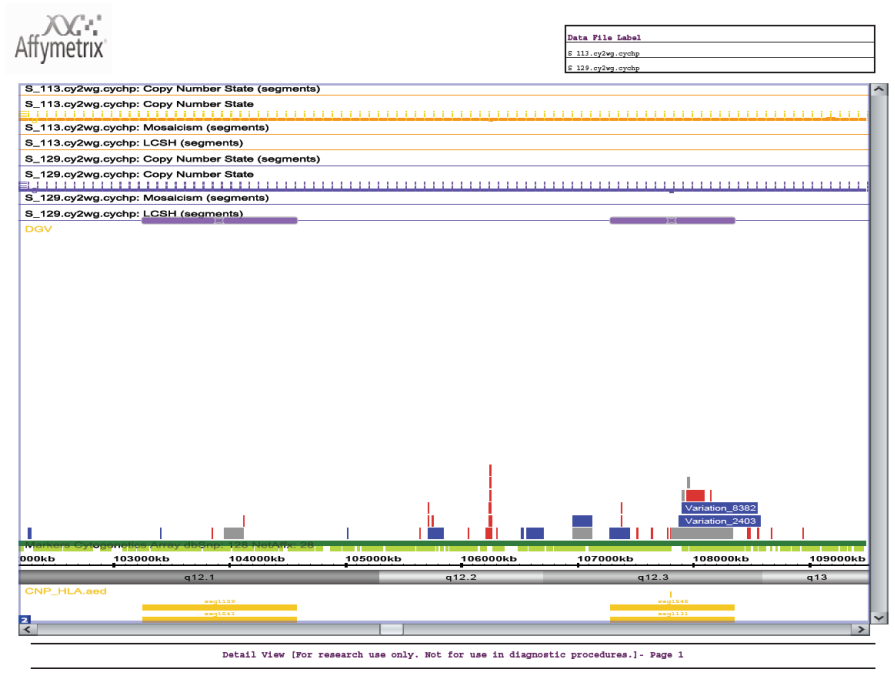


Figure 14.6 Detail View report, p 2

Data Files In View

File Label	File Name
S_113.cy2wg.cychp	S_113.cy2wg.cychp
S_129.cy2wg.cychp	S_129.cy2wg.cychp

Genome Filters

Segment Type	Filter Type	Minimum Filter Value
Gain	Segment Length	500000
Loss	Segment Length	200000

Settings for S_129.cy2wg.cychp

Setting Type	Setting
Smoothing	On
Smoothing Max Jump Limit	Off
Joining	50 markers
Joining Max Jump Limit	Off
Restricted Mode	Off

Settings for S_113.cy2wg.cychp

Setting Type	Setting
Smoothing	On
Smoothing Max Jump Limit	Off
Joining	50 markers

Detail View [For research use only. Not for use in diagnostic procedures.]- Page 2

Figure 14.7 Detail View report, p 3

Setting Type	Setting
Joining Max Jump Limit	Off
Restricted Mode	Off

Overlap Map - CNP_HLA.aed

Export as PNG

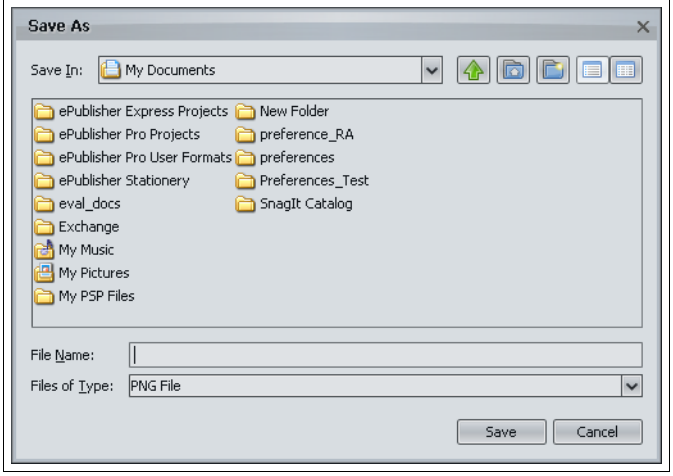
You can also create a PNG screenshot of the entire software screen.

IMPORTANT: The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.

To export a PNG screenshot:

1. From the Reports menu, select **Export application window PNG**.
   The Save As PNG dialog box opens (Figure 14.8).

Figure 14.8 Save As PNG dialog box



2. Select a folder location for the file using the navigation tools.

   This folder location is automatically selected when exporting other screenshots during a session.

3. Enter a name for the PNG file in the File Name box.

4. Click the **Save** button.

   The PNG file screen shot is saved in the selected location.

   The PNG file can be cropped in a graphics program like Paint and inserted into a word processing document if desired.

Exporting Table Data

ChAS provides several options for exporting table data:

- *Export Table Data in a PDF File* (on page 275)
- *Export Tables as TXT File* (on page 280)
- *Exporting a Segments Table with Modified Segments to a .TXT File* (on page 282)

Export Table Data in a PDF File

For information on how to choose and export preset column content (from previously saved table states).
See Saved Table States (on page 232)

IMPORTANT: The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.

NOTE: You can also export data from the Segment Table by selecting Export Segment Table PDF from the Reports Menu (see page).

NOTE: You can't export Graph table data as a PDF report.

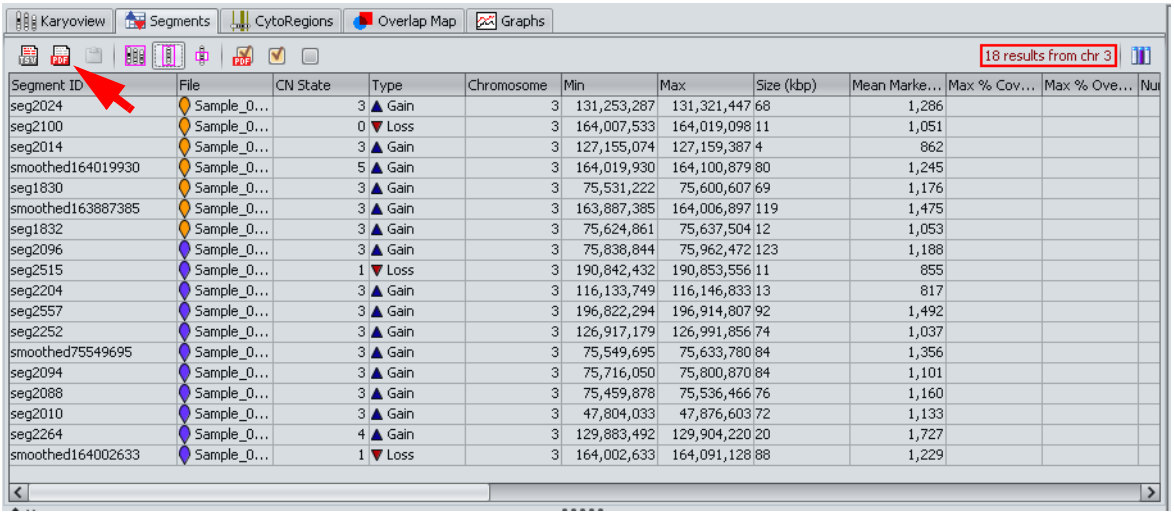
In order to track which of your segments have been modified (merged, created de novo, deleted, had their start or end coordinates edited, or had their type or state changed), you MUST perform the following BEFORE exporting the Segments Table to PDF. The only PDF Report capable of tracking which segments have been modified is the Segments Table PDF, all other PDFs show NO visual or textual distinction between modified and non-modified segments.

1. Click (top right of Segments table).
   The Select Columns window appears.
2. Scroll down, locate, then click to check the **Modified Segment** checkbox.
3. Click **OK**.
   The Modified Segment column is now added to your Segments table.

To export Table information in a PDF file:

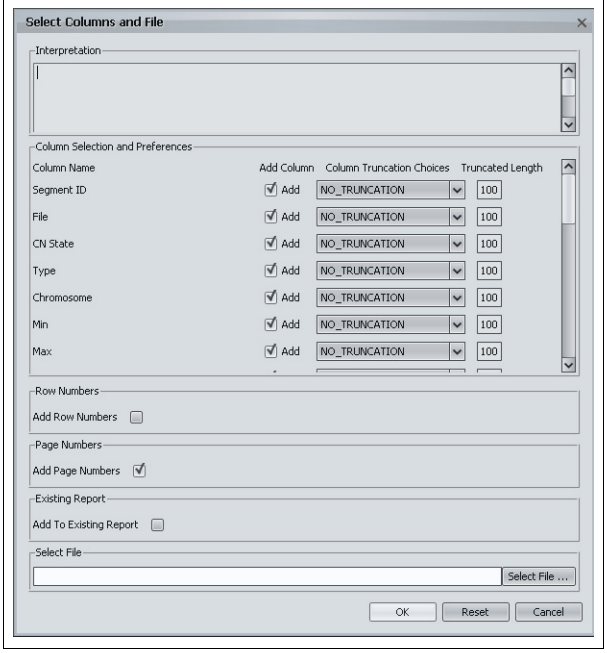
1. Click the PDF tool bar button (Figure 14.9).
   The Select Columns and Files window appears.
2. Enter the appropriate text in the Interpretation field. (Figure 14.10 on page 276)

Figure 14.9 Segments table with data filtered



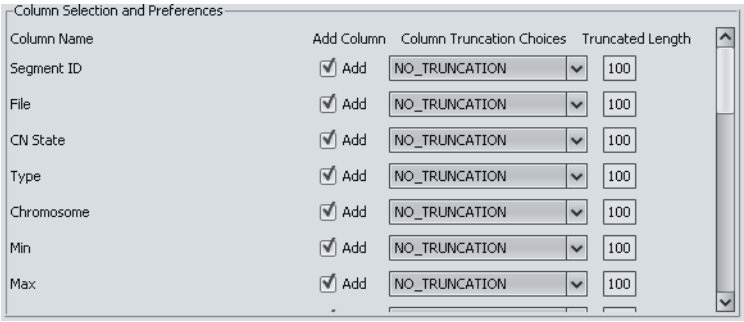
Segment ID	File	CN State	Type	Chromosome	Min	Max	Size (kbp)	Mean Marke...	Max % Cov...	Max % Ove...	Nu
seg2024	Sample_0...	3	Gain	3	131,253,287	131,321,447	68	1,286			
seg2100	Sample_0...	0	Loss	3	164,007,533	164,019,098	11	1,051			
seg2014	Sample_0...	3	Gain	3	127,155,074	127,159,387	4	862			
smoothed164019930	Sample_0...	5	Gain	3	164,019,930	164,100,879	80	1,245			
seg1830	Sample_0...	3	Gain	3	75,531,222	75,600,607	69	1,176			
smoothed163887385	Sample_0...	3	Gain	3	163,887,385	164,006,897	119	1,475			
seg1832	Sample_0...	3	Gain	3	75,624,861	75,637,504	12	1,053			
seg2096	Sample_0...	3	Gain	3	75,838,844	75,962,472	123	1,188			
seg2515	Sample_0...	1	Loss	3	190,842,432	190,853,556	11	855			
seg2204	Sample_0...	3	Gain	3	116,133,749	116,146,833	13	817			
seg2557	Sample_0...	3	Gain	3	196,822,294	196,914,807	92	1,492			
seg2252	Sample_0...	3	Gain	3	126,917,179	126,991,856	74	1,037			
smoothed75549695	Sample_0...	3	Gain	3	75,549,695	75,633,780	84	1,356			
seg2094	Sample_0...	3	Gain	3	75,716,050	75,800,870	84	1,101			
seg2088	Sample_0...	3	Gain	3	75,459,878	75,536,466	76	1,160			
seg2010	Sample_0...	3	Gain	3	47,804,033	47,876,603	72	1,133			
seg2264	Sample_0...	4	Gain	3	129,883,492	129,904,220	20	1,727			
smoothed164002633	Sample_0...	1	Loss	3	164,002,633	164,091,128	88	1,229			

Figure 14.10 Select Columns and File dialog box for Segment Table PDF



3. Select the columns to be displayed and the truncation rules for the content of the columns (Figure 14.11).

Figure 14.11 Column Selection and Preferences



Column Name Header of column in table.

Add Column Select checkbox to display column in PDF.

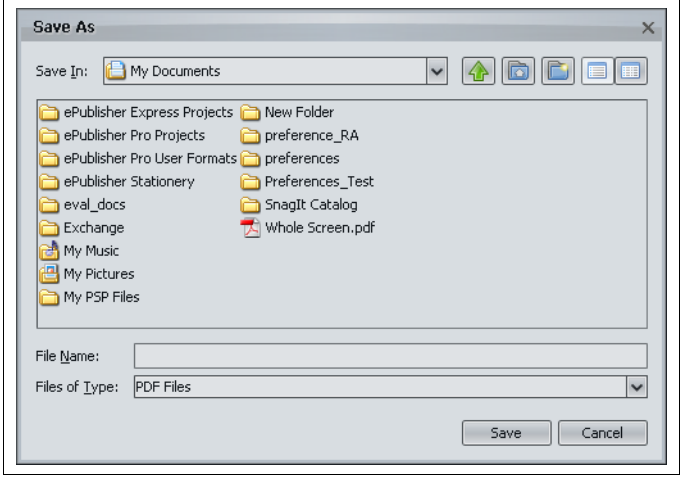
Column Truncation Choices Select Truncation Options:
- No_Truncation: field is exported as is, using wrap-around if necessary.
- Truncate_Beginning: Truncates content at the beginning of the field, leaving as many characters as specified in Truncated Length box.
- Truncate_Middle: Truncates content in the middle of the field, leaving characters at the beginning and end, with ellipses (…) to mark the truncated characters.
- Truncate_End: Truncates content at the end of the field, leaving as many characters as specified in Truncated Length box.

Truncated length Number of characters displayed after truncating the data.

4. Select the option for adding page and row numbers, if desired.
5. Select the option for adding to an existing report, if desired. See *Combining PDFs into a Single Report* (on page 284).
6. Click the **Select File** button.

   The Save As dialog box opens (Figure 14.12).

Figure 14.12 Save As for Segment Table PDF



7. Select a folder location for the file using the navigation tools.
8. Enter a name for the PDF file, or select a file for the information to be appended to.
9. Click **Save** in the Save As dialog box.
10. Click **OK** in the Select Columns and File dialog box.

    A PDF file (Figure 14.13 on page 278, Figure 14.14 on page 279, Figure 14.15 on page 280) is created with the selected data type saved.

The report displays:

- Table type
- Information on chromosome and genome region
- Interpretation
- Data files
- Genome or CytoRegion Segment Filters used
- Settings for Data Processing
- Microarray Nomenclature
- Details of the table data

Figure 14.13 PDF file, page 1

Segments Table

For research use only. Not for use in diagnostic procedures.

Data Files In View

File Label	File Name
11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_20110511.cyhd.cychp	11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_20110511.cyhd.cychp

Named Setting - Standard

Genome Filters

Segment Type	Filter Type	Minimum Filter Value
Gain	Segment Length	400000
Gain	Marker Count	50
Loss	Segment Length	400000
Loss	Marker Count	50

Settings for 11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_20110511.cyhd.cychp

Setting Type	Setting
Smoothing	On
Smoothing Max Jump Limit	Off
Joining	50 Markers, 200 kbp
Joining Max Jump Limit	Off
Restricted Mode	Off

Figure 14.14 PDF file, page 2

Setting Type	Setting
QC Metrics	mapd ≤ 0.25; snpQC ≥ 15; wavinessSd ≤ 0.12

Microarray Nomenclature

11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_20110511.cyhd.cychp:
arr[hg19] 5p15.33p15.31(408,191-9,147,251)x3, 5p15.31p15.1(9,153,447-17,457,890)x3, 5p15.1p14.3(17,478,490-19,427,622)x3, 5p14.3p13.2(19,573,864-34,799,067)x3, 5p13.2(35,095,856-37,121,657)x3, 5p13.2(37,373,774-37,867,140)x3, 8p23.3p23.2(2,095,440-3,238,755)x1, 8p23.2p23.1(5,515,516-8,657,463)x1, 8p23.1(9,799,952-10,428,420)x1, 8p22p21.3(18,632,189-19,150,880)x1, 8p21.3(20,871,989-21,381,652)x1, 8p21.3(21,453,224-22,037,634)x1, 8p21.3(22,489,199-23,142,234)x1, 8p21.1p12(28,608,428-29,153,037)x1, 8p11.23p11.22(36,596,983-38,556,532)x3, 8p11.22(38,712,599-39,246,760)x3, 8p11.21(39,843,294-40,701,486)x3, 8p11.21(40,708,420-41,128,710)x3, 8p11.21(41,546,091-42,359,484)x3, 8p11.21(42,615,579-43,026,713)x3, 8p11.1q12.1(43,308,594-57,006,476)x3, 8q12.1(57,079,698-61,538,423)x3, 8q12.2q12.3(61,633,287-63,981,227)x3, 8q12.3q13.1(64,174,027-66,162,323)x3, 8q13.1q21.11(66,832,680-74,564,645)x3, 8q21.11(74,678,627-75,100,051)x3, 8q21.11q21.12(75,200,634-79,831,539)x3, 8q21.12q21.13(80,073,939-80,655,778)x3, 8q21.13(80,745,550-81,253,112)x3, 8q21.13q21.3(81,413,406-88,202,737)x3, 8q21.3(88,349,344-90,639,110)x3, 8q21.3q22.1(91,017,130-95,394,662)x3, 8q22.1q22.2(95,927,685-100,366,870)x3, 8q22.2(100,556,943-101,505,174)x3, 8q22.3(101,653,351-103,058,024)x3, 8q22.3(103,318,496-104,712,568)x3, 8q22.3q23.1(105,565,117-110,426,607)x3, 8q23.2q23.3(110,551,033-114,048,383)x3, 8q23.3(114,065,943-115,632,043)x3, 8q23.3(115,650,683-116,541,758)x3, 8q23.3q24.12(116,594,211-120,882,163)x3, 8q24.12(120,885,308-121,317,090)x3, 8q24.12q24.13(121,490,686-124,103,255)x3, 8q24.13q24.21(124,437,925-129,442,671)x3, 8q24.21q24.3(129,458,472-140,761,415)x3, 8q24.3(140,809,247-141,802,431)x3, 8q24.3(142,224,672-146,295,771)x3, 11q22.1q22.2(101,950,388-102,575,754)x1, 11q23.1(110,945,179-111,874,060)x1, 14q32.33(104,410,114-107,053,817)x3, 16p11.2p11.1(32,524,764-35,283,214)x1, 17q21.31(44,212,823-44,784,639)x3, 21q21.3(27,447,808-28,032,671)x3, 21q21.3q22.11(31,424,401-31,852,180)x3, 22q11.21(20,184,050-20,761,384)x3, 22q11.22(22,673,638-23,258,438)x3, 22q11.23q12.1(25,569,645-26,249,277)x3, 22q12.3(34,397,452-35,048,300)x3

Segments Details

Figure 14.15 PDF file, page 3

File	CN State	Type	Chromosome	Cytoband Start	Size (kbp)	Gene Count	Genes	OMIM ® Genes Count	OMIM ® Genes	CytoRegions	Call	Interpretation
11-0816_LC_ONC134B_B10_PoP#2_CytoScan-PS_201105111.cyhd.cychp	3.0	Gain	5	p15.33	8739.06	45	AHRR, C5orf55, EXOC3, LOC25845, SLC9A3, CEP72, TPPP, ZDHHC11, BRD9, TRIP13, NKD2, SLC12A7, SLC6A19, SLC6A18, TERT, CLPTM1L, SLC6A3, LPCAT1, SDHAP3, LOC728613, MIR4277, MRPL36, NDUFS6, IRX4, IRX2, C5orf38, LOC285577, IRX1, LOC340094, ADAMTS16, KIAA0947, FLJ33360, MED10, UBE2QL1, LOC255167, NSUN2, SRD5A1, PAPD7, MIR4278, LOC442132, ADCY2, C5orf49, FASTKD3, MTRR, SEMA5A	0	N/A	Cytoregions Not Set		

Segments Table - Page 3

Tips for Exporting Tables

Follow the tips below to improve the export of table data in a PDF report:

- Use truncation
- Select only columns you need or select columns from a saved table state. (See *Saved Table States* (on page 232))
- Use filtering options (Segment filters, displaying only results for a chromosome or area in the detail view, etc.) to limit the number of values being exported

Export Tables as TXT File

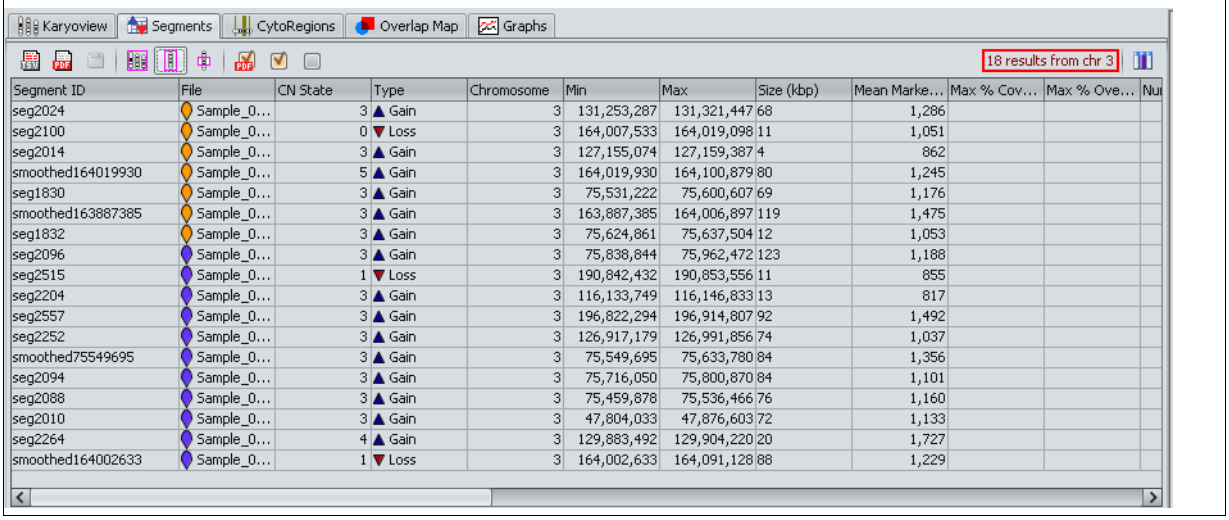
The TXT file format allows you to transfer data to other software for analysis.

IMPORTANT: The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.

To export table information as a text file:

1. Perform pre-filtering on the data in the table (Figure 14.16).

Figure 14.16 Segments table with data filtered

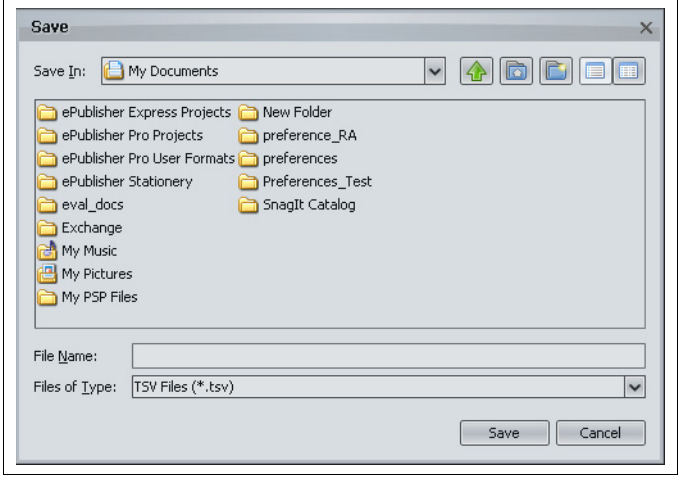


Segment ID	File	CN State	Type	Chromosome	Min	Max	Size (kbp)	Mean Marke...	Max % Cov...	Max % Ove...	Nu
seg2024	Sample_0...	3	Gain	3	131,253,287	131,321,447	68	1,286			
seg2100	Sample_0...	0	Loss	3	164,007,533	164,019,098	11	1,051			
seg2014	Sample_0...	3	Gain	3	127,155,074	127,159,387	4	862			
smoothed164019930	Sample_0...	5	Gain	3	164,019,930	164,100,879	80	1,245			
seg1830	Sample_0...	3	Gain	3	75,531,222	75,600,607	69	1,176			
smoothed163887385	Sample_0...	3	Gain	3	163,887,385	164,006,897	119	1,475			
seg1832	Sample_0...	3	Gain	3	75,624,861	75,637,504	12	1,053			
seg2096	Sample_0...	3	Gain	3	75,838,844	75,962,472	123	1,188			
seg2515	Sample_0...	1	Loss	3	190,842,432	190,853,556	11	855			
seg2204	Sample_0...	3	Gain	3	116,133,749	116,146,833	13	817			
seg2557	Sample_0...	3	Gain	3	196,822,294	196,914,807	92	1,492			
seg2252	Sample_0...	3	Gain	3	126,917,179	126,991,856	74	1,037			
smoothed75549695	Sample_0...	3	Gain	3	75,549,695	75,633,780	84	1,356			
seg2094	Sample_0...	3	Gain	3	75,716,050	75,800,870	84	1,101			
seg2088	Sample_0...	3	Gain	3	75,459,878	75,536,466	76	1,160			
seg2010	Sample_0...	3	Gain	3	47,804,033	47,876,603	72	1,133			
seg2264	Sample_0...	4	Gain	3	129,883,492	129,904,220	20	1,727			
smoothed164002633	Sample_0...	1	Loss	3	164,002,633	164,091,128	88	1,229			

2. In the Table tool bar, click the Export **TXT** button.

   The Save as TXT dialog box opens (Figure 14.17).

Figure 14.17 Save as TXT dialog box

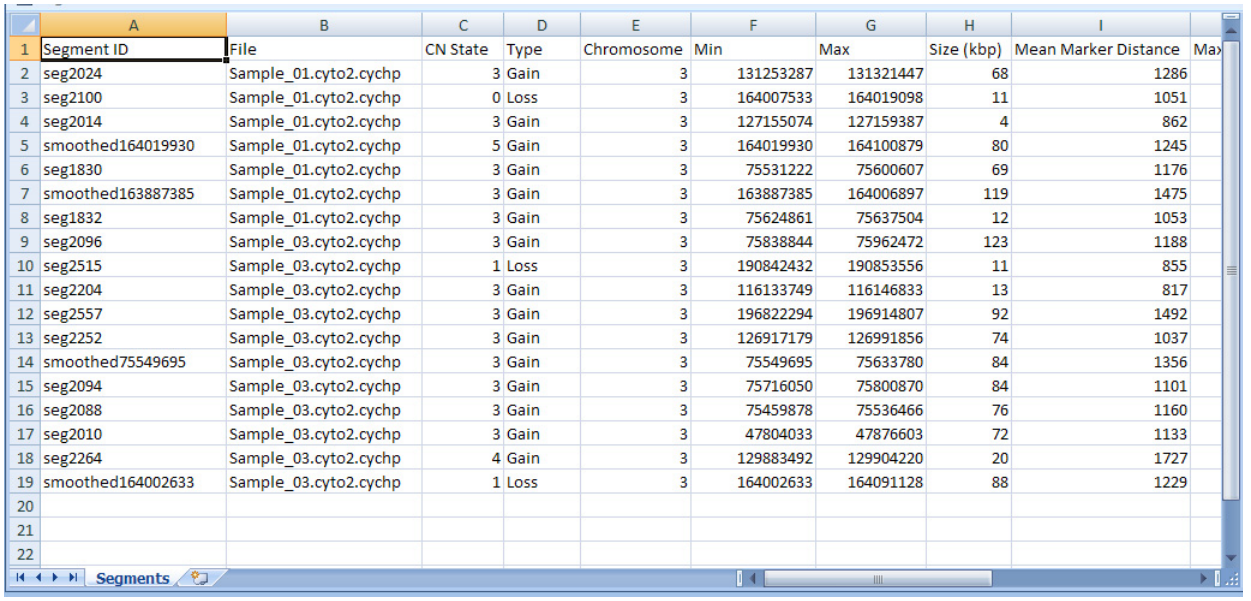


3. Select a folder location for the file using the navigation tools.
4. Enter a name for the TXT file.
5. Click **Save**.

   The TXT file is saved in the selected location (Figure 14.18).

   It can be opened using a text editing or spreadsheet program, or in other software designed to use Tab Separated Value TXT format.

Figure 14.18 TXT file opened in Excel



	A	B	C	D	E	F	G	H	I
1	Segment ID	File	CN State	Type	Chromosome	Min	Max	Size (kbp)	Mean Marker Distance
2	seg2024	Sample_01.cyto2.cychp	3	Gain	3	131253287	131321447	68	1286
3	seg2100	Sample_01.cyto2.cychp	0	Loss	3	164007533	164019098	11	1051
4	seg2014	Sample_01.cyto2.cychp	3	Gain	3	127155074	127159387	4	862
5	smoothed164019930	Sample_01.cyto2.cychp	5	Gain	3	164019930	164100879	80	1245
6	seg1830	Sample_01.cyto2.cychp	3	Gain	3	75531222	75600607	69	1176
7	smoothed163887385	Sample_01.cyto2.cychp	3	Gain	3	163887385	164006897	119	1475
8	seg1832	Sample_01.cyto2.cychp	3	Gain	3	75624861	75637504	12	1053
9	seg2096	Sample_03.cyto2.cychp	3	Gain	3	75838844	75962472	123	1188
10	seg2515	Sample_03.cyto2.cychp	1	Loss	3	190842432	190853556	11	855
11	seg2204	Sample_03.cyto2.cychp	3	Gain	3	116133749	116146833	13	817
12	seg2557	Sample_03.cyto2.cychp	3	Gain	3	196822294	196914807	92	1492
13	seg2252	Sample_03.cyto2.cychp	3	Gain	3	126917179	126991856	74	1037
14	smoothed75549695	Sample_03.cyto2.cychp	3	Gain	3	75549695	75633780	84	1356
15	seg2094	Sample_03.cyto2.cychp	3	Gain	3	75716050	75800870	84	1101
16	seg2088	Sample_03.cyto2.cychp	3	Gain	3	75459878	75536466	76	1160
17	seg2010	Sample_03.cyto2.cychp	3	Gain	3	47804033	47876603	72	1133
18	seg2264	Sample_03.cyto2.cychp	4	Gain	3	129883492	129904220	20	1727
19	smoothed164002633	Sample_03.cyto2.cychp	1	Loss	3	164002633	164091128	88	1229
20									
21									
22									

Segments

Exporting a Segments Table with Modified Segments to a .TXT File

In the Segments Table:

Materially Modified segments are shown in italic text when Edit Mode is ON, but not italicized when Edit Mode is OFF.

Calls and Interpretations don't cause a segment's row text to be italicized in Edit Mode.

Deleted segments are shown in strike-through text when Edit Mode is ON, and are not present in the table when Edit Mode is OFF.

In the TXT Exports of the Segments Table table, please note that Deleted segments will be part of the export when Edit Mode is ON, and will not be part of the export if Edit Mode is OFF

Exported .TXT File Examples

Figure 14.19 is an example of a Segment Table with Edit Mode **ON**. Note its italicized text representing materially modified segments and the strike-through text showing deletions.

Figure 14.19 Segment table - Edit Mode ON

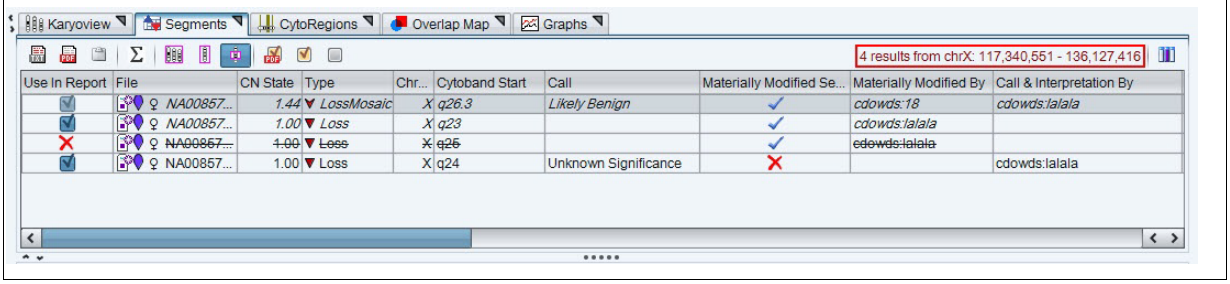
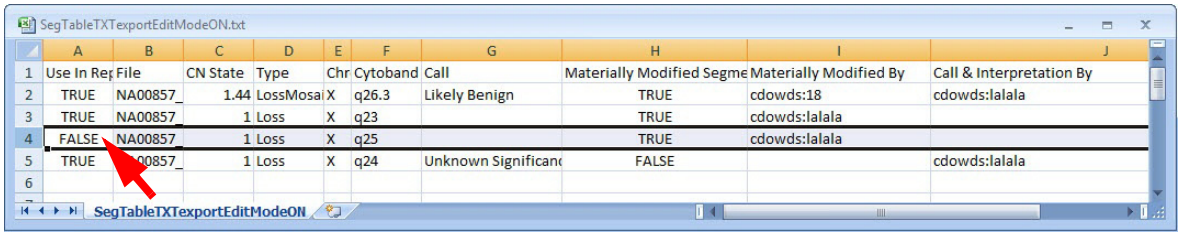


Figure 14.20 is an example of a Segment table that has been exported with the Edit Mode **ON**. Note the 4th row of the Use in Report column. In the case of this segment, it reads FALSE, because this segment was deleted.

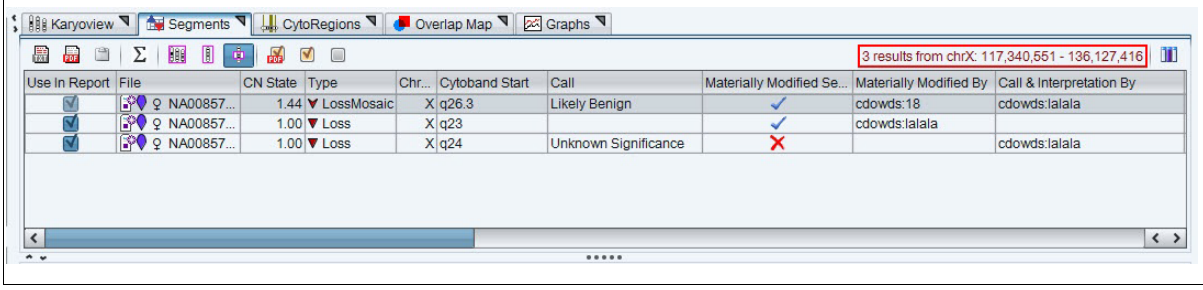
Figure 14.20 Segment table .TXT EXPORT - Edit Mode ON



	A	B	C	D	E	F	G	H	I	J
1	Use In Re	File	CN State	Type	Chr	Cytoband	Call	Materially Modified Segme	Materially Modified By	Call & Interpretation By
2	TRUE	NA00857_	1.44	LossMosai	X	q26.3	Likely Benign	TRUE	cdowds:18	cdowds:lalala
3	TRUE	NA00857_	1	Loss	X	q23		TRUE	cdowds:lalala	
4	FALSE	NA00857_	1	Loss	X	q25		TRUE	cdowds:lalala	
5	TRUE	00857_	1	Loss	X	q24	Unknown Significanc	FALSE		cdowds:lalala
6										

Figure 14.21 is an example of a Segment Table with Edit Mode **OFF**. The italicized text representing materially modified segments is no longer present. The deleted segment and strike-through text showing a deletion (shown in Figure 14.19 and Figure 14.20) also do not appear.

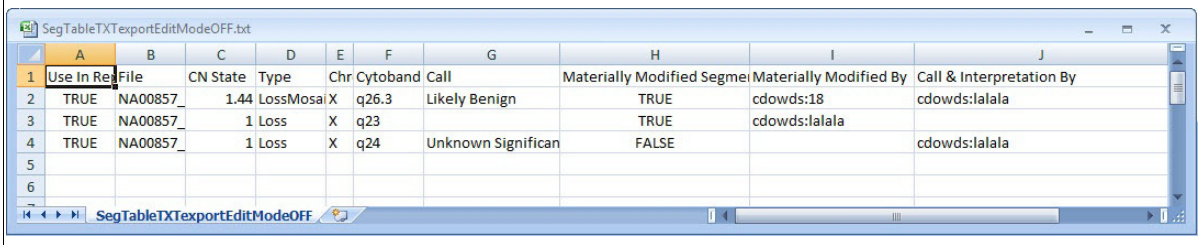
Figure 14.21 Segment table - Edit Mode OFF



Use In Report	File	CN State	Type	Chr...	Cytoband Start	Call	Materially Modified Se...	Materially Modified By	Call & Interpretation By
☑	♀ NA00857...	1.44	LossMosaic	X	q26.3	Likely Benign	✓	cdowds:18	cdowds:lalala
☑	♀ NA00857...	1.00	Loss	X	q23		✓	cdowds:lalala	
☑	♀ NA00857...	1.00	Loss	X	q24	Unknown Significance	✗		cdowds:lalala

Figure 14.22 is an example of how an exported .txt table appears with Edit Mode **OFF**. Note the Deleted segment shown in Figure 14.19 and Figure 14.20 is not present.

Figure 14.22 Segment table .TXT EXPORT - Edit Mode OFF



	A	B	C	D	E	F	G	H	I	J
1	Use In Re	File	CN State	Type	Chr	Cytoband	Call	Materially Modified Segme	Materially Modified By	Call & Interpretation By
2	TRUE	NA00857_	1.44	LossMosai	X	q26.3	Likely Benign	TRUE	cdowds:18	cdowds:lalala
3	TRUE	NA00857_	1	Loss	X	q23		TRUE	cdowds:lalala	
4	TRUE	NA00857_	1	Loss	X	q24	Unknown Significan	FALSE		cdowds:lalala
5										
6										

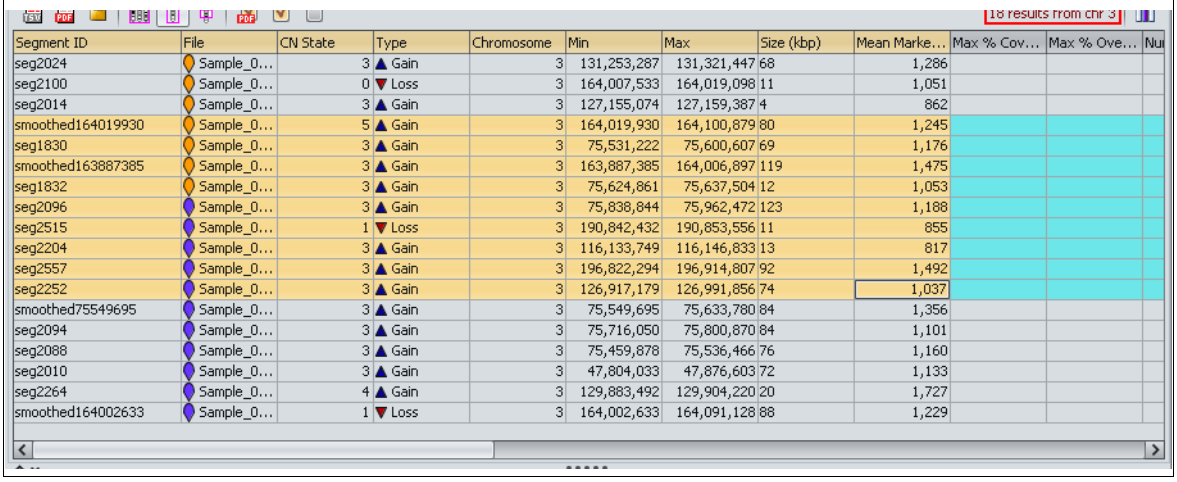
Transfer to Clipboard

You can copy data from selected cells to the clipboard for pasting into a text or spreadsheet file.

To copy table data to the Clipboard:

1. Select the cells you want to copy in the table (Figure 14.23).

Figure 14.23 Segment table with cells selected



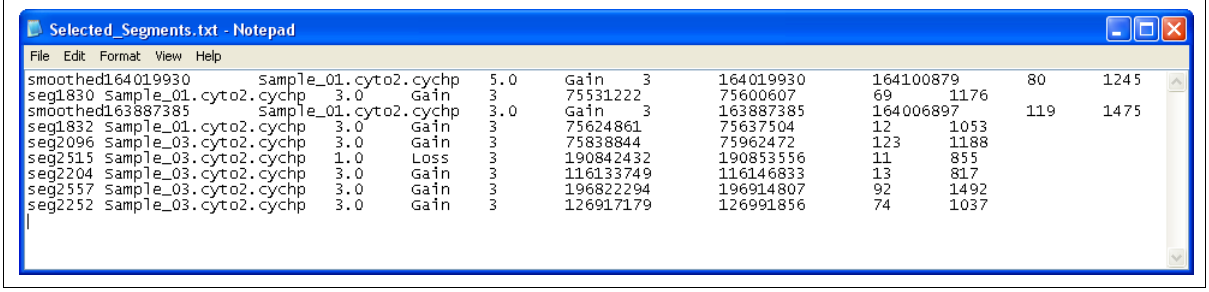
Segment ID	File	CN State	Type	Chromosome	Min	Max	Size (kbp)	Mean Marke...	Max % Cov...	Max % Ove...	Nu
seg2024	Sample_0...	3	Gain	3	131,253,287	131,321,447	68	1,286			
seg2100	Sample_0...	0	Loss	3	164,007,533	164,019,098	11	1,051			
seg2014	Sample_0...	3	Gain	3	127,155,074	127,159,387	4	862			
smoothed164019930	Sample_0...	5	Gain	3	164,019,930	164,100,879	80	1,245			
seg1830	Sample_0...	3	Gain	3	75,531,222	75,600,607	69	1,176			
smoothed163887385	Sample_0...	3	Gain	3	163,887,385	164,006,897	119	1,475			
seg1832	Sample_0...	3	Gain	3	75,624,861	75,637,504	12	1,053			
seg2096	Sample_0...	3	Gain	3	75,838,844	75,962,472	123	1,188			
seg2515	Sample_0...	1	Loss	3	190,842,432	190,853,556	11	855			
seg2204	Sample_0...	3	Gain	3	116,133,749	116,146,833	13	817			
seg2557	Sample_0...	3	Gain	3	196,822,294	196,914,807	92	1,492			
seg2252	Sample_0...	3	Gain	3	126,917,179	126,991,856	74	1,037			
smoothed75549695	Sample_0...	3	Gain	3	75,549,695	75,633,780	84	1,356			
seg2094	Sample_0...	3	Gain	3	75,716,050	75,800,870	84	1,101			
seg2088	Sample_0...	3	Gain	3	75,459,878	75,536,466	76	1,160			
seg2010	Sample_0...	3	Gain	3	47,804,033	47,876,603	72	1,133			
seg2264	Sample_0...	4	Gain	3	129,883,492	129,904,220	20	1,727			
smoothed164002633	Sample_0...	1	Loss	3	164,002,633	164,091,128	88	1,229			

2. Click the **Copy to Clipboard** button in the table tool bar.

   The selected data is copied to the clipboard.

   You can paste the data on the clipboard into a text or spreadsheet file (Figure 14.24).

Figure 14.24 Data pasted into text file



Combining PDFs into a Single Report

You can combine different PDFs into a single report with multiple pages and content.

You can do this by:

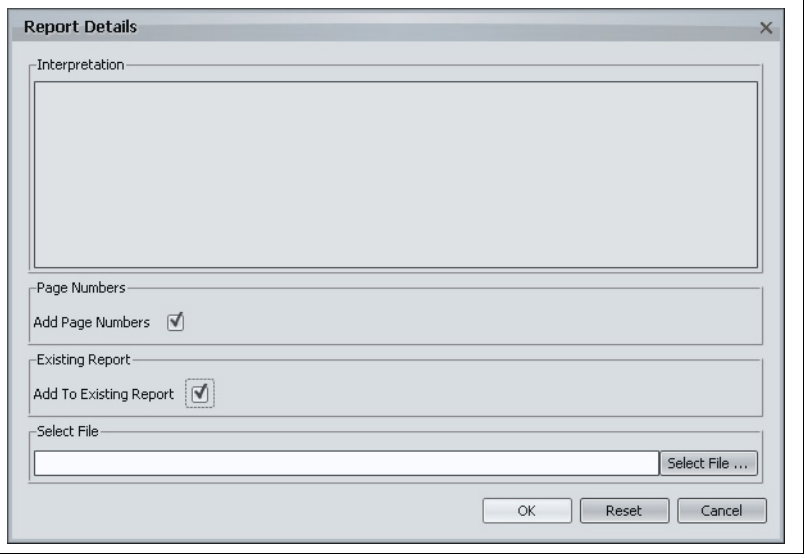
- Adding new data to an existing PDF Report
- Merging two or more existing PDF reports

IMPORTANT: **The results from ChAS are for Research Use Only. Not for use in diagnostic procedures.**

To add new PDF report data to an existing PDF Report:

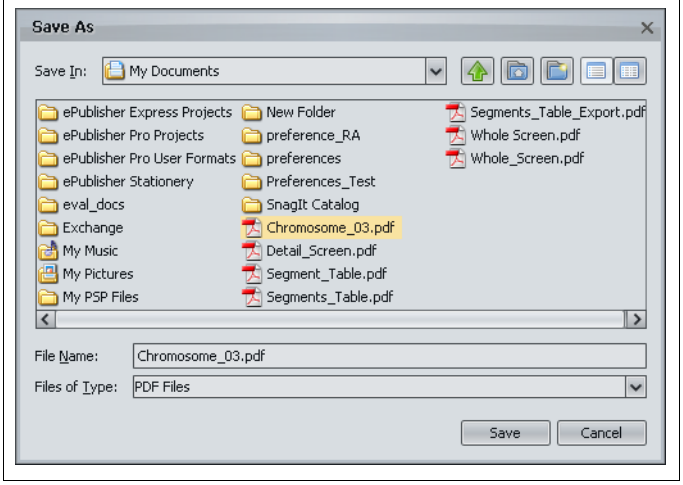
1. When creating a report, select the Add to Existing Report option in the Select Columns and File or Report Details dialog box (Figure 14.25).

Figure 14.25 Report Details dialog box with Add to Existing Report selected



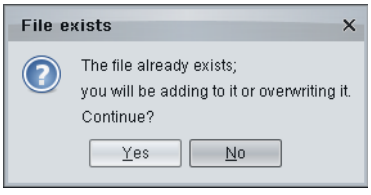
2. Click **Select File** and select an existing PDF Report for the data to be added to.
   The Save As dialog box opens (Figure 14.26).

Figure 14.26 Save As dialog box



3. Select a PDF file and click **Save** in the Save As dialog box.
   A confirmation message appears, asking if you want to overwrite or add to the data (Figure 14.27).

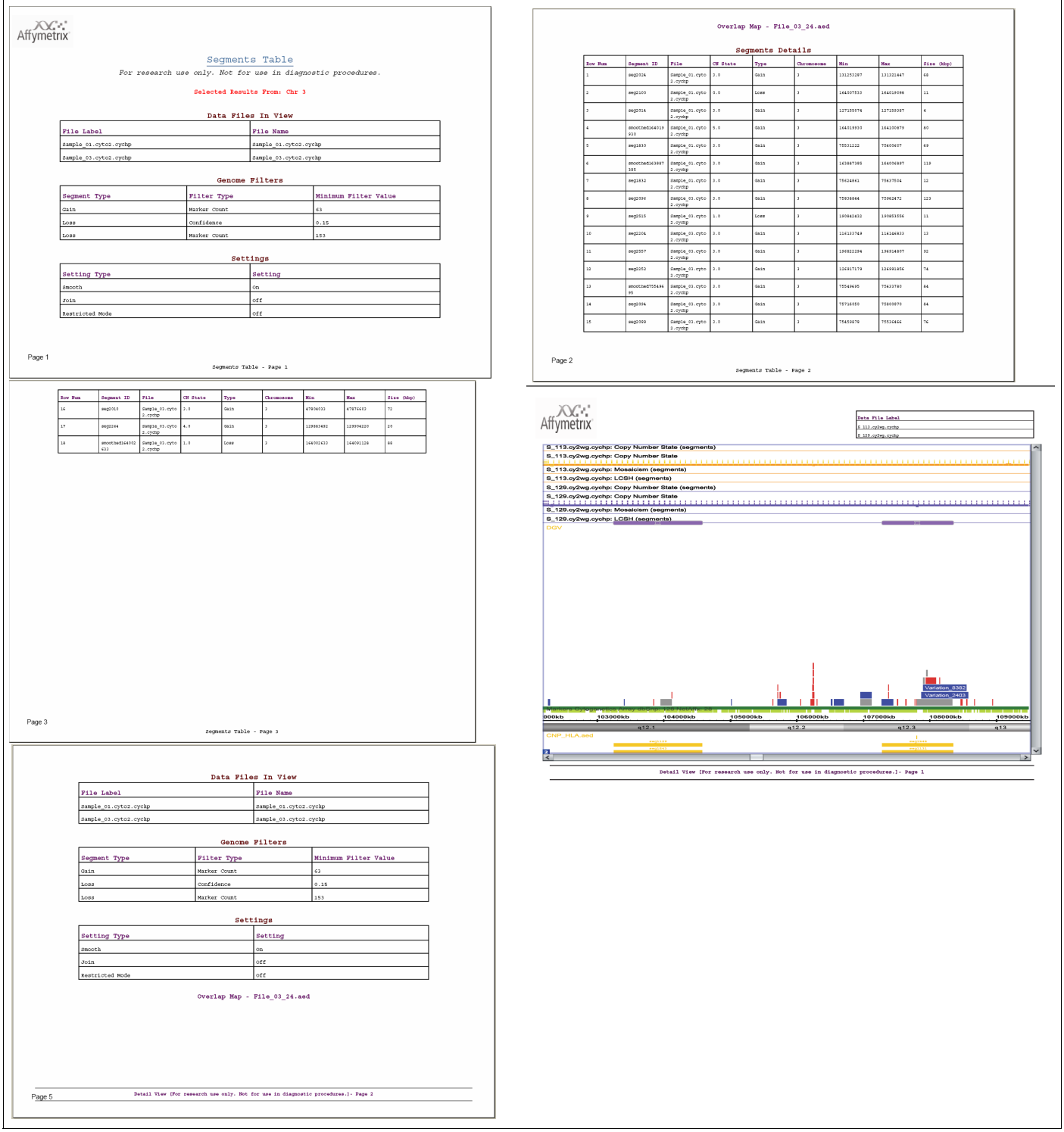
Figure 14.27 Confirmation message



4. Click Yes in the Confirm Rewrite notice to append the data in the selected file.

5. Click **OK** in the Select Columns and File or Report Details dialog box
   The new report data is combined with the existing report (Figure 14.28).

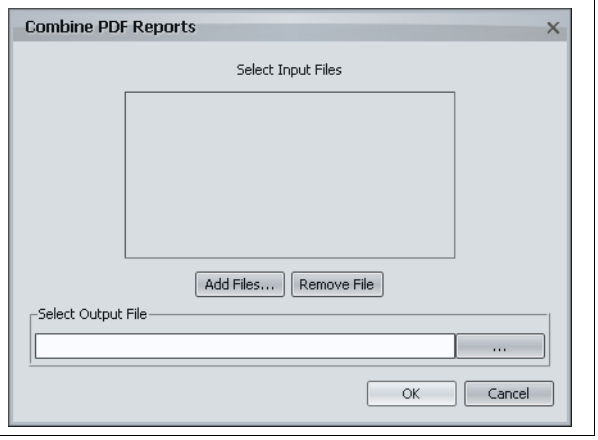
Figure 14.28 Combined PDF reports



To combine two existing PDF reports:

1. From the Report menu, select Combine PDFs into Report.
   The Combine PDF Reports dialog box opens (Figure 14.29).

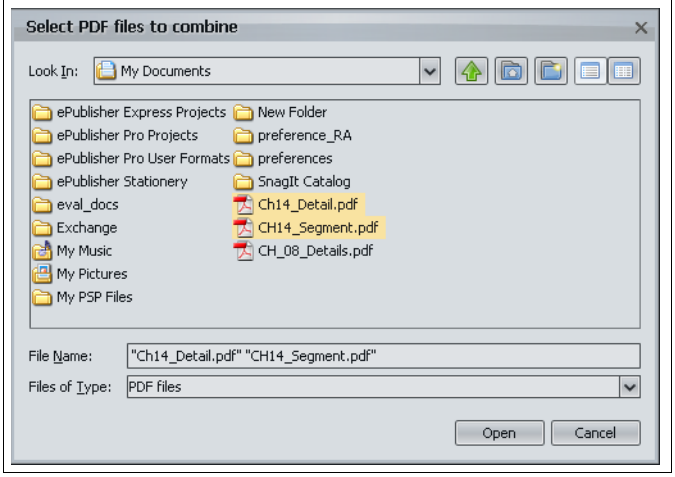
Figure 14.29 Combine PDF Reports dialog box



2. Click **Add Files…**.

   The Select PDF Files to Combine dialog box opens (Figure 14.30).

Figure 14.30 Select PDF Files to Combine dialog box



3. Select the PDF files to combine and click **Open** in the Select PDF Files dialog box.

   The selected files are displayed in the Select Input Files list.

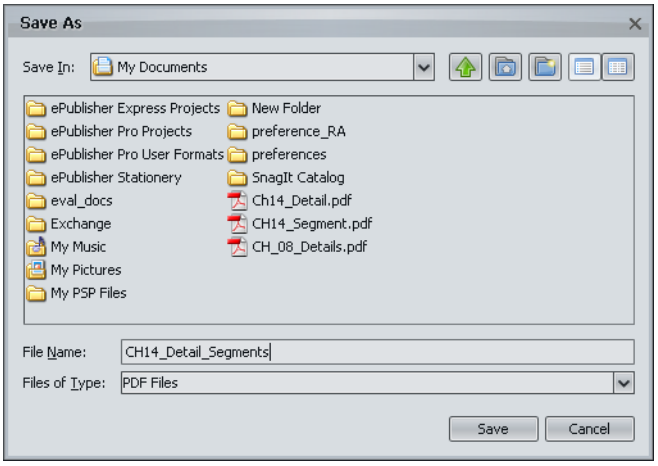
   You can use the Remove File button to remove a selected input file.

   Click and drag on a file in the list to change the order of data in the Combined PDF.

4. Click the **Select Output File** button.

   The Save As dialog box opens (Figure 14.31).

Figure 14.31 Save As dialog box



5. Enter a file name for the combined file and click **Save** in the Save As dialog box.
6. Click **OK** in the Combine PDF Reports dialog box.

   The Selected PDFs are combined.

User Profiles and Named Settings

ChAS provides many options for customizing the display of data and annotations. The User Profiles and Named Settings functions allow you to save your analysis and display settings. These functions are described in the following sections:

- *Types of Settings* (on page 289)
- *Creating and Using User Profiles* (on page 290)
- *Creating and Using Named Settings* (on page 292)
- *Exporting and Importing Preferences* (on page 294)

Types of Settings

ChAS provides two ways to store setup information:

- *User Profiles* (on page 289)
- *Named Settings* (on page 289)

They work differently and perform different functions.

User Profiles

A ChAS Browser User Profile stores your selections for various display settings as they were when the software was last shut down while using that user profile.

A new user profile can be created or selected only when starting the software.

The user profile saves the following display settings:

- Screen size, displayed tabs, and sizing of display areas
- The views displayed in ChAS, and the size of the display panes
- Available Named Settings: Different users can have different lists of named settings to choose from
- Name of the currently selected named setting
- Copies of the user's custom (not shared) named settings
- Data Display Configurations
- Region information files selected for CytoRegions and Overlap Map
- Which types of graph and segment data are turned on or off
- Display options for graph data (height, grid, values, etc.)
- Chromosome and area displayed.
- Selected Reference Annotation database (ChAS Browser NetAffx Genomic Annotations file)
- Loaded AED and BED files
- The files and Reference Annotations (Genes, DGV, etc.) that are checked or unchecked
- Custom color rules

Named Settings

A Named Setting stores the user's choices for:

- Which types of graph and segment data are turned on or off

- Segment Filter Settings
- Restricted Mode on/off

  The Named Setting doesn't save a particular CytoRegion file, but does inform you if no file is selected when you select a setting with restricted mode on.

  It is possible to apply a Named Setting with restricted mode using a different cytoregion file than was selected for the initial creation of the setting.

You can switch between different Named Settings in the same user profile to look at different types of data.

ChAS provides pre-configured (shared) Named Settings indicated by the icon as described in the table below (Table 15.1). These named settings cannot be deleted.

Table 15.1 Shared named settings

Named Setting	Genome Segment Filters	CytoRegion Segment Filters	Data Types
Standard	Gain: Marker Count = 50 Size (kbp) = 400; Loss: Marker Count = 50 Size (kbp) = 400	Gain: Marker Count = 50 Size (kbp) = 400 Loss: Marker Count = 50 Size (kbp) = 400	Gain Loss GainMosaic LossMosaic Copy Number State Weighted Log2 Ratio Allele Peaks Allele Difference
High Resolution	Gain: Marker Count = 50 Size (kbp) = 100; Loss: Marker Count = 50 Size (kbp) = 100	Gain: Marker Count = 25 Size (kbp) = 50 Loss: Marker Count = 25 Size (kbp) = 50	Gain Loss GainMosaic LossMosaic Copy Number State Weighted Log2 Ratio
LOH only (3Mb and 50 SNPs)	LOH: Marker Count = 50 Size (kbp) = 3000	LOH: Marker Count = 50 Size (kbp) = 3000	LOH Genotype Calls Allele Peaks Allele Difference
Differential Gains and Losses	Gain: Marker Count = 50 Size (kbp) = 400; Loss: Marker Count = 50 Size (kbp) = 100	Gain: Marker Count = 25 Size (kbp) = 50 Loss: Marker Count = 25 Size (kbp) = 50	Gain Loss GainMosaic LossMosaic Copy Number State Weighted Log2 Ratio
OncoScan Defaults	0	0	0

Creating and Using User Profiles

You can only select or create user profiles when starting ChAS.

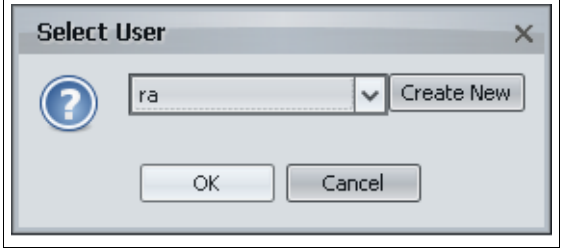
To create a user profile

1. Double-click on the ChAS icon on the desktop; or

   From the Windows Start Menu, select **Programs → Affymetrix →Chromosome Analysis Suite → Chromosome Analysis Suite**.

   The ChAS Splash Screen and the Select User dialog box open. (Figure 15.1)

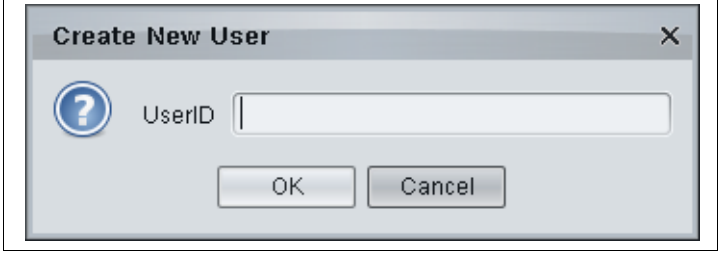
Figure 15.1 Select User dialog box



2. Click **Create New** in the Select User dialog box.

   The Create New User dialog box opens (Figure 15.2).

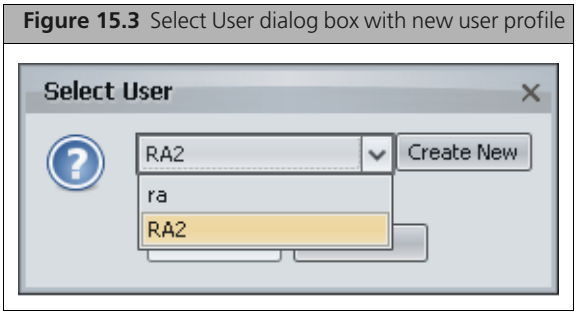
Figure 15.2 Create New User dialog box



3. Enter a name for the new profile in the User ID box.
4. Click **OK** in the Create New User dialog box.

   The new user appears in the drop-down User list in the Select User dialog box (Figure 15.3).

Figure 15.3 Select User dialog box with new user profile



5. Select the new user and click **OK**.

   Any changes you make to the setup of the software that is listed in *User Profiles* (on page 289) will be saved when you shut down the software and used the next time the software is opened with this user profile.

To delete a user profile:

- Go to the Windows folder where the user profiles are stored and delete the folder with the profile name you want to delete.

  You can see the location of the folder in the About window, as described in *Analysis File Locations* (on page 19).

Creating and Using Named Settings

You can save a snapshot of your favorite settings as a Named Setting. To apply a particular Named Setting to the active data (check marked in the Files List), make a selection from the Named Setting drop-down list. Some pre-configured Shared Named Settings may be available for use by all users. Only an administrator can add or remove Shared Named Settings, but any user can apply them to their data.

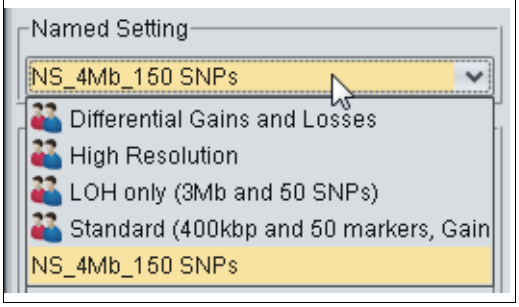
To save a named setting:

1. Set the display data settings as desired.
   These can include:
   - Which graphs and segment types are turned on or off
   - Segment Filter Settings
   - Restricted mode on/off
2. From the Preferences menu, select **Save Named Setting**.
   The Setting Name dialog box opens (Figure 15.4).

Figure 15.4 Named Setting List

3. Enter a name for the setting you want to create and click **OK**.
   The setting is saved and appears in the Named Setting dropdown list (Figure 15.5).

Figure 15.5 Named Setting dropdown list

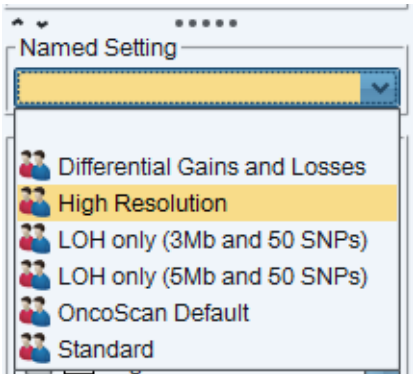


NOTE: The Named Setting saves the settings at the time it was created. Subsequent changes to the settings will not be saved in the Named Setting.

To select a Named Setting:

1. From the Named Setting drop-down list, select the setting (Figure 15.6).

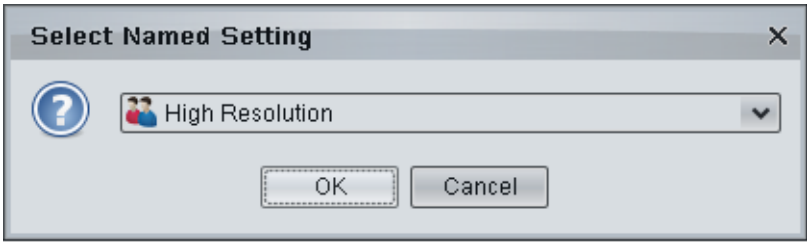
Figure 15.6 Named Setting dropdown list



Alternatively, from the Preference menu, select **Apply Named Setting…**

The Select Named Setting dialog box opens (Figure 15.7).

Figure 15.7 Select Named Setting dialog box



2. Select the Named Setting from the drop-down list and click **OK**.

   The selected setting is applied.

NOTE: A Named Setting is not modified by any changes that you make to the settings in ChAS. If you want to keep a copy of your new settings, you will need to save them as a new Named Setting.

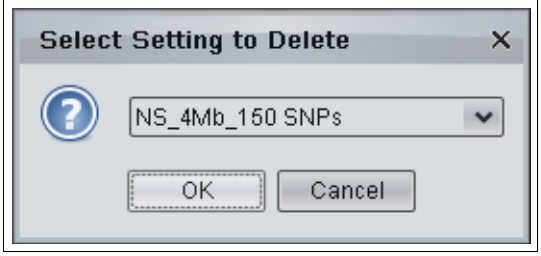
To delete a Named Setting:

1. From the Preferences menu, select **Delete Named Setting.**

   The Delete Setting dialog box opens (Figure 15.8).

NOTE: Shared Named Settings (denoted by the icon in the Named Setting list) do not appear in the Delete Setting drop-down list. Users cannot delete or modify a shared Named Setting.

Figure 15.8 Delete Setting dialog box



2. Select the setting you want to delete from the drop-down list and click **OK**.
   The setting is deleted.

Exporting and Importing Preferences

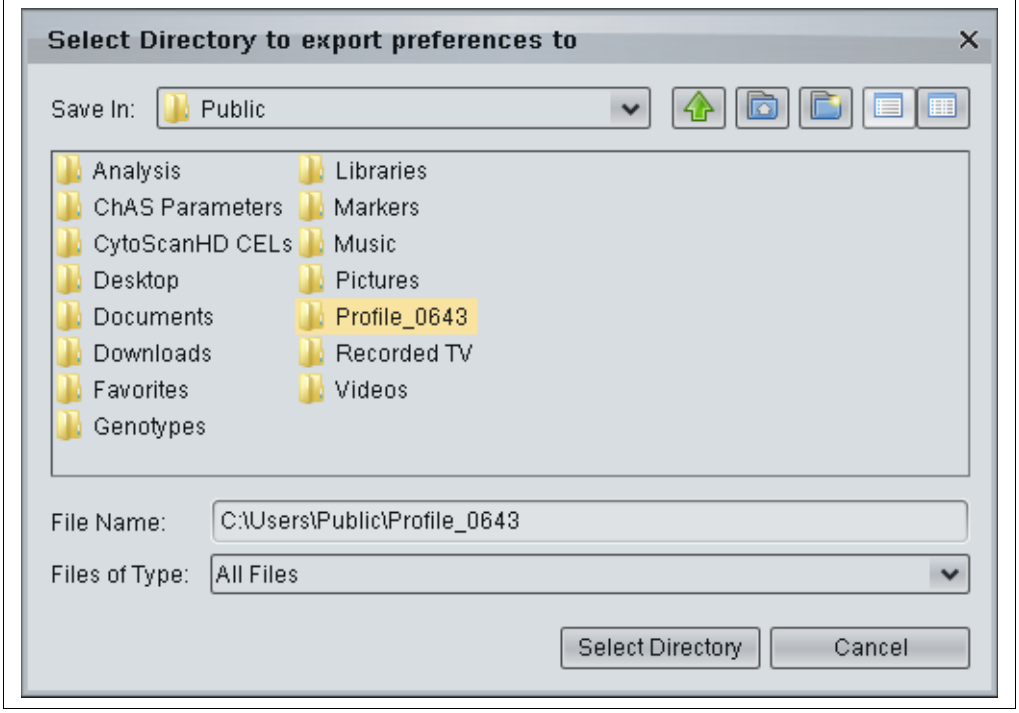
Preferences functions allow you to transfer most of the settings in a User Profile between one system and another.

NOTE: If you import "exported" preferences that reference a Shared Named Setting which no longer exists, such as a Shared Named Setting from ChAS 1.0.1 or ChAS 1.1, the profile will be changed to point to the default Shared Name Setting.

To export preferences:

1. From the Preferences menu, select **Export Preferences…**
   The Select Directory to export preferences to dialog box opens (Figure 15.9).

Figure 15.9 Select Directory to export preferences to



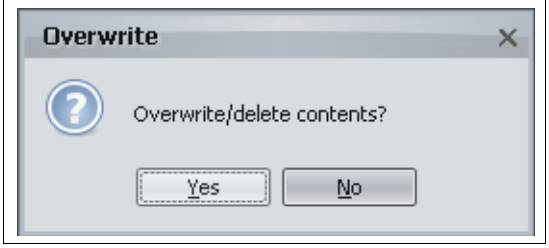
2. Use the navigation features of the dialog box to select or create a directory for the preferences.

NOTE: The software will create a folder called "preferences" in the directory you select or create. If you select a directory that already contains a "preferences" folder, it will be overwritten. When you want to import the preferences, select the directory that contains the "preferences" folder that is indicated by the icon.

3. Click **Select Directory.**

   If the directory already contains a "preferences" folder, the Overwrite notice appears (Figure 15.10).

Figure 15.10 Overwrite notice



4. Click **Yes** to export the preference files to the directory that you selected.

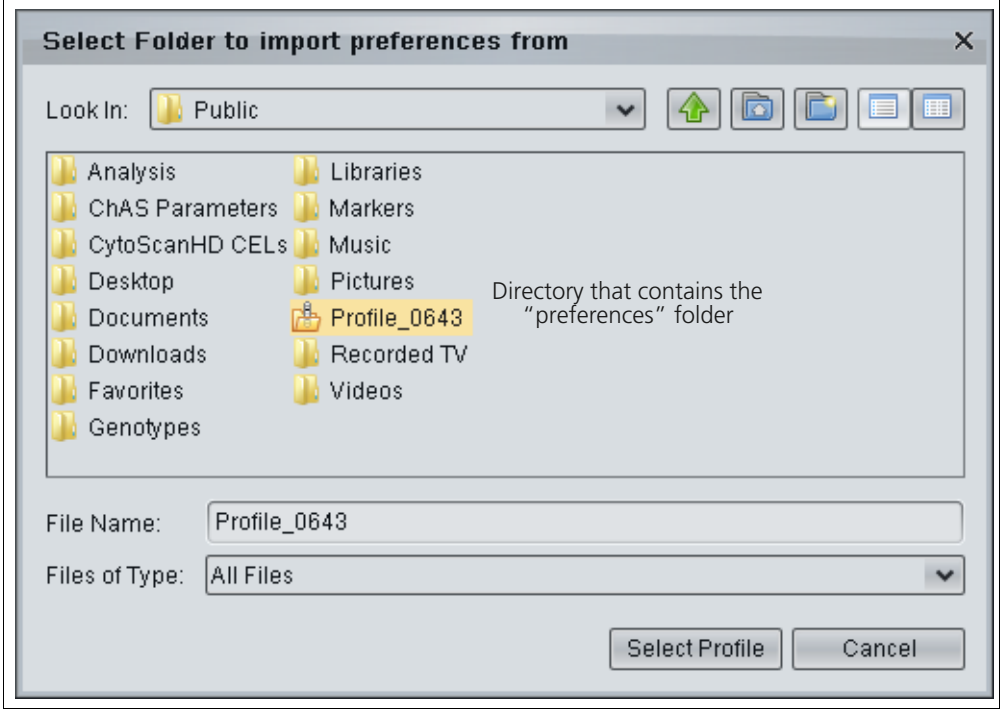
   You can then transfer the preferences to another user profile or system.

To import preferences:

1. From the Preferences menu, select **Import Preferences…**

   The Select Folder to import preferences from dialog box opens (Figure 15.11).

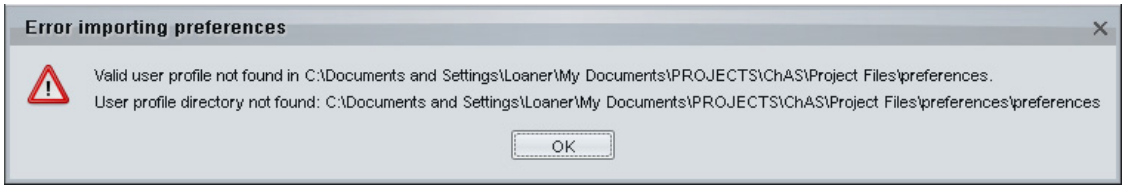
Figure 15.11 Select Folder to import preferences from



2. Use the navigation features of the dialog box to select the directory that the preferences were exported to (directories that contain a "preferences" folder are indicated by the icon.)
3. Click **Open to import the preference files.**

If you selected a directory that doesn't contain the "preferences" folder, the following notice appears (Figure 15.12).

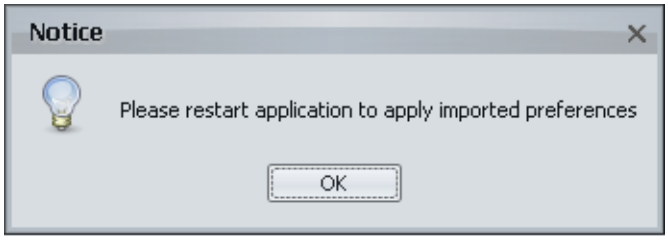
Figure 15.12 Error notice



Click OK and repeat steps 1 through 3, selecting the correct folder.

The Restart notice appears (Figure 15.13).

Figure 15.13 Restart notice



The imported preferences will not be applied until you restart ChAS.

Analysis Parameters

IMPORTANT: The CytoScan® HD Array has no ChAS user-modifiable parameters for Single Sample Analysis or Reference Model File Creation Analysis.

Analysis Parameters for Single Sample Analysis

- For information on CytoScan HD algorithm parameters and their values, see:

  http://media.affymetrix.com/support/developer/powertools/changelog/apt-copynumber-cyto.html

Reference Model File Creation

Reference Model File Creation is done with fixed parameters in the Reference Creation workflow.

It is essential that the input for Reference Model Creation include at least 44 total CEL files and at least 20 males and 20 females.

There are no user-adjustable parameters for the Reference Model File Creation

- For information on CytoScan HD algorithm parameters and their values, see:

  http://media.affymetrix.com/support/developer/powertools/changelog/apt-copynumber-cyto.html

AED File Format

Affymetrix Extensible Data files contain data that annotate positions on a genome. AED allows custom, typed fields, can be edited in the ChAS Browser's AED Editor feature (*Viewing and Batch Editing AED File Contents* (on page 216)), and supports internationalization.

AED files can be created by Chromosome Analysis Suite (ChAS) and loaded into ChAS as region information files to:

- Define CytoRegion and Overlap Map regions
- Record information of interest about features in the genome

They can be produced and edited in:

- ChAS (Recommended)
- Text-editing software (Not Recommended)
- Spreadsheet software such as Microsoft Excel (Not Recommended)

This appendix explains the formatting and use of Affymetrix Extensible Data (AED) files with ChAS.

- *AED File Description* (on page 298)
- *Property Name Elements* (on page 301)
- *Compatibility* (on page 304)
- *References* (on page 306)

AED File Description

An AED file contains a list of annotations, descriptions of features on a biological *sequence* such as a chromosome. This description is comprised of several *properties*—either properties defined by this specification, such as the annotation start and stop positions; or properties defined by users or third parties.

An AED file may also provide *metadata* which describe the particular group of annotations in the file as a whole, such as the author of the file or the genome assembly for which the annotations were produced.

Properties and metadata have certain *types* which define the semantics and constrain the range of values they may have. Properties should begin with a lowercase letter, while types should begin with an uppercase letter.

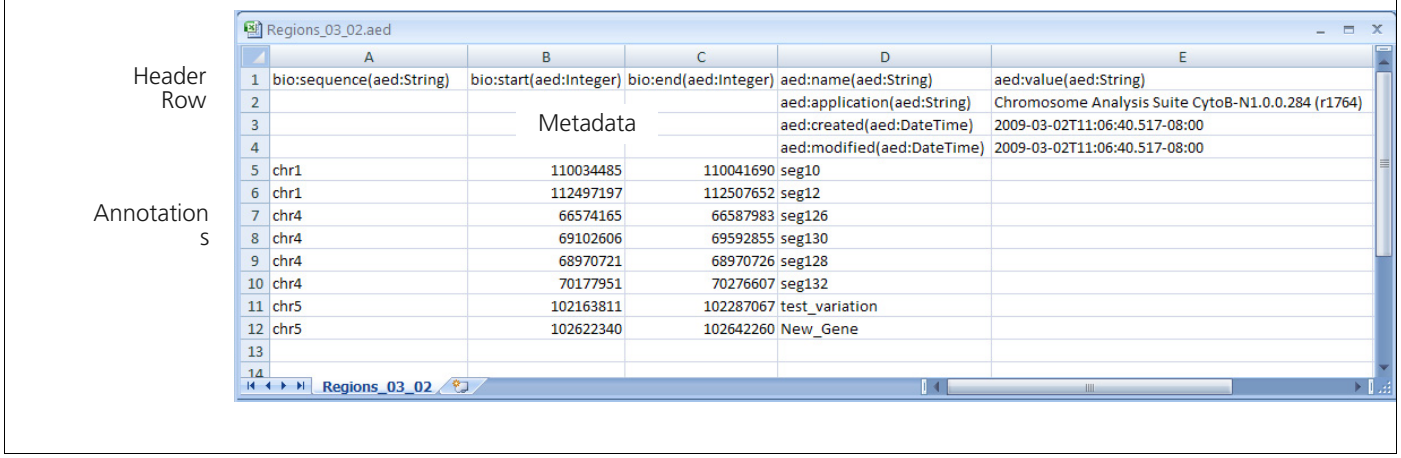
The AED file format uses a tab-delimited text format with the *.aed file extension. It uses Unicode character sets.

IMPORTANT: **AED supports only Unicode, which can be stored in one of various encodings (charsets such as UTF-8, UTF-16LE, and UTF-16BE). The AED file indicates the charset with an initial Byte Order mark (BOM). An AED file with no initial BOM is not recommended. An AED file that does not begin with a BOM will be interpreted as containing only the ASCII subset of Unicode, resulting in an error if any characters lie outside the range of ASCII. (With no indication of a charset, it is not possible to determine which non-ASCII characters were intended. File formats such as BED that make assumptions about non-ASCII characters have the potential to corrupt data when transported between systems.)**

It has the following components (Figure B.1):

- *Header row* (on page 299): names the properties that can be used in the annotations
- *Metadata Records* (on page 300) (optional): provides information about the AED file itself and the group of annotations it contains.
- *Annotations* (on page 300): for each feature annotated, an annotation row provides values for the properties listed in the header rows.

Figure B.1 AED file in Excel with required header fields for properties and metadata



	A	B	C	D	E
1	bio:sequence(aed:String)	bio:start(aed:Integer)	bio:end(aed:Integer)	aed:name(aed:String)	aed:value(aed:String)
2				aed:application(aed:String)	Chromosome Analysis Suite CytoB-N1.0.0.284 (r1764)
3				aed:created(aed:DateTime)	2009-03-02T11:06:40.517-08:00
4				aed:modified(aed:DateTime)	2009-03-02T11:06:40.517-08:00
5	chr1	110034485	110041690	seg10	
6	chr1	112497197	112507652	seg12	
7	chr4	66574165	66587983	seg126	
8	chr4	69102606	69592855	seg130	
9	chr4	68970721	68970726	seg128	
10	chr4	70177951	70276607	seg132	
11	chr5	102163811	102287067	test_variation	
12	chr5	102622340	102642260	New_Gene	
13					
14					

Header row

The header row of an AED file is a tab-delimited list of the properties that can be used to describe a region of the genome.

Each AED file header represents a property. Normal records in the file represent annotations, and the record fields represent annotation properties. Special metadata records represent metadata properties for the file as a whole, rather than for a particular annotation.

A property name has the following format:

namespacePrefix:propertyIdentifier(namespacePrefix:TypeIdentifier)

- `namespacePrefix` (optional) assigns the property or type to a vocabulary grouping called a namespace. The lack of a namespace prefix indicates that the property has been created by a user and is not part of the formal AED specification.

  The lack of a namespacePrefix indicates that the property is in the default/custom namespace; this namespace enables users to add properties to an annotation just by adding new columns, such as foo(aed:String) or bar(aed:Integer).

  See *Namespaces* (on page 301) for more information.
- `propertyIdentifier` names the property that the values in the column are for.

  Each annotation can be assigned an unlimited number of properties. Each property has a certain meaning, and this meaning is usually defined by the documentation for the property namespace. The purpose of the AED file is to indicate values for certain properties for each annotation. For example, by use of the aed:name(aed:String) column, the AED file indicates a string value to be used as the name each annotation).
- `TypeIdentifier` (always capitalized) specifies the data type of the value to be used for the property in the AED file.

Examples include:

- `bio:sequence(aed:String)`
- `aed:value(aed:String)`
- `medianMarkerDistance(aed:Integer)`

Required Fields

Fields may appear in any order, except that the following predefined fields must always appear in the following order at the beginning of the header:

- `bio:sequence(aed:String)`
- `bio:start(aed:Integer)`
- `bio:end(aed:Integer)`
- `aed:name(aed:String)`
- `aed:value(aed:String)`

  `[????](optional: you need to use this property if you are including metadata information in the file).`

IMPORTANT: ChAS validates the property types when importing an AED file. If a file header specifies a known property, but includes an incorrect data type for the property, the file will not be loaded. For example, "fish:score" is a known property with "whole number" data type. An AED file header that specifies "fish:score(aed:String)" would be treated as an error.

Metadata Records

Some records, instead of providing annotation about a location on a genome assembly, provide metadata information about the AED file itself (Figure B.2). These *metadata records* are identified by the presence of an empty string in the `bio:sequence` field. The bio:start and bio:end fields must also be left blank for metadata records. If there are metadata records present, the `aed:value` field is required.

In a metadata record, the value in the `aed:name` field is interpreted as the name of the metadata property, with type identification rules identical to those of the header fields. The value in the `aed:value` field is interpreted as the value of the metadata property, and the characters that make up its string value must follow the lexical and semantic rules specified by the type indicated in the `aed:name` field.

Figure B.2 Metadata entries

bio:sequence(aed:String)	bio:start(aed:Integer)	bio:end(aed:Integer)	aed:name(aed:String)	aed:value(aed:String)
			aed:application(aed:String)	Chromosome Analysis Suite CytoB-N1.0.0.284 (r1764)
			aed:created(aed:DateTime)	2009-03-02T11:06:40.517-08:00
			aed:modified(aed:DateTime)	2009-03-02T11:06:40.517-08:00

Blanks for bio:sequence and other properties

Metadata Property names

Metadata Property values

All other fields in a metadata record should be blank.

Annotations

The rows below the Metadata properties are the annotations. Each row is a tab-delimited list of values. Each value must have the correct data type, as described in the property name for that value.

Property Name Elements

The property name elements are described in more detail in the following sections:

- *Namespaces* (on page 301)
- *Properties* (on page 301)
- *Types* (on page 303)

Namespaces

The name of each type and property in AED is considered part of a vocabulary grouping called a *namespace*. Namespaces prevent clashes between names defined by disparate parties, as well as unambiguously identify commonly used types and properties so that identical semantics may be assured. A namespace is identified by a Uniform Resource Identifier (URI) as defined in **RFC 3986**.

A type or property identifies its namespace by a namespace prefix followed by a colon character. If no namespace prefix is present, the type or property is considered part of the AED default namespace. The part of the type or property after the namespace prefix is considered its *simple name*.

AED has several build-in namespaces, with predefined namespace URIs and prefixes:

Table B.1 Namespaces

Namespace Prefix	Description	Examples

	Custom and experimental properties not yet established as standard.	`foo`
		`bar`
		`myProperty`

`aed`	AED-specific properties and types.	`aed:name` `aed:Integer`
`bio`	Descriptions of biological entities.	`bio:sequence` `bio:Strand`
`style`	Information related to information representation, visually or otherwise.	`style:color`

If any other namespace is used in an AED file, it must be declared in the metadata section of the file using the special namespace prefix. The simple name of the metadata header indicates the prefix being declared, and the value (of type aed:URI) indicates the namespace to be associated with the prefix. For example, to associate the prefix "*example*" with the URI http://example.com/namespace/, and the "*affx*" with the URL *http://affymetrix.com/ontology/*, use the following metadata record:

```
namespace:example(aed:URI)    http://example.com/namespace/

namespace:affx (aed:URL)      http://affymetrix.com/ontology
```

Properties

This section describes the properties defined by the AED specification. By convention property names begin with lowercase letters. A predefined property is only required if indicated. Some properties are only useful as metadata, and these are so indicated.

Affymetrix Properties

This section describes the properties defined by the AED specification. By convention, property names begin with lowercase letters. A predefined property is only required if indicated. Some properties are only useful as metadata, and these are so indicated.

Table B.2 Affymetrix Properties

Name	Type	Description
`Affx:ucscGenomeVersion (aed:String)`	`aed:String`	**(Metadata)** The genome assembly version using UCSC names, for example "hg19".

AED Properties

These properties are parts of the AED namespace.

Table B.3 AEDProperties

Name	Type	Description
`aed:application`	`aed:String`	**(Metadata)** The name of the application that produced the AED file, if metadata, or the annotation.
`aed:category`	`aed:String`	Identifies the group and optionally subgroups into which the resource is classified. Subcategories, if any, should be delimited using the forward slash character `'/'` (U+002F) with no whitespace. (Example: `copynumber/gain`)
`aed:created`	`aed:DateTime`	**(Metadata)** The point in time the data was created; this is not necessarily the time the file was created.
`aed:counter`	`aed:Integer`	A general purpose field to be incremented when user-defined circumstances occur. A common use for this field is to indicate, the number of samples in which the condition has been observed.
`aed:modified`	`aed:DateTime`	**(Metadata)** The point in time the data was modified; this is not necessarily the time the file was modified.
`aed:name`	`aed:String`	**(Required.)** Indicates the name of the record. In a metadata record, this value is interpreted as the name and type of the metadata property.
`aed:note`	`aed:String`	A user-defined explanation or comment regarding the annotation.
`aed:value`	`aed:String`	**(Required only if metadata records are present.)** In a metadata record, this value is interpreted as the value of the metadata property. An AED processor must ignore this field for all non-metadata records.
`aed:uuid`	`aed:UUID`	**(Metadata)** The universally unique identifier of the resource. Although allowed, it is not always advised to identify user-editable resources such as AED documents with UUIDs, as copying and manually editing such resources can result in multiple such resources with identical UUIDs, negating the purpose of UUIDs.

Biology Properties

Table B.4 Biology Properties

Name	Type	Description
`bio:assembly`	`aed:URI`	**(Metadata)** A URI indicating the genome assembly used. Currently the DAS GlobalSeqIDs are recommended.
`bio:state`	`aed:Rational`	The algorithm-determined state (e.g. copy number) of an annotation.
`bio:confidence`	`aed:Rational`	A value between 0.0 and 1.0, inclusive, indicating the confidence that an annotation call is accurate.
`bio:end`	`aed:Integer`	**(Required.)** The zero-based ending position, exclusive, of the record along the sequence.
`bio:markerCount`	`aed:Integer`	The number of markers such as probes that intersect an annotation.
`bio:sequence`	`aed:String`	**(Required.)** The name of the chromosome (e.g. `chr3`, `chrY`), contig (e.g. `ctg5`), or scaffold (e.g. `scaffold90210`). The special value of an empty string (“”) indicates that the record is a metadata record, giving special meaning to values in other fields in the record.
`bio:start`	`aed:Integer`	**(Required.)** The zero-based starting position, inclusive, of the record along the sequence.
`bio:strand`	`bio:Strand`	The sequence strand on which a feature lies.

Style Properties

Style properties are used to control the display of the annotation.

Table B.5 Style Properties

Name	Type	Description
`style:color`	`aed:Color`	The color to be used when visually depicting the annotation.

Types

AED defines the following types. By convention type names begin with an uppercase letter. The lexical form is applied to the resulting character sequence derived after following the quoting rules of **RFC 4180**.

Table B.6

Type	Description	Lexical Form	Examples
`aed:String`	A sequence of Unicode characters.	*character**	▪ `Abc` ▪ `a "fun" test!`
`aed:Integer`	The positive whole numbers, the negative whole numbers, and zero.	[-] *digit+*	▪ `123` ▪ `0` ▪ `-5000`

Table B.6

Type	Description	Lexical Form	Examples
`aed:Rational`	A rational number. Currently only rational numbers with finite decimal expansion are allowed. AED processors typically implement this type using floating point values. This type does not allow floating point not-a-number (NaN) values.	[-]*digit*+.*digit*+ [e[+|-]*digit*+] The lexical form "NaN" is explicitly prohibited.	▪ `-123.0` ▪ `0.0` ▪ `123.45` ▪ `1.2e+5`
`aed:Boolean`	A binary true/false value.	"false" | "FALSE" | "true" | "TRUE"	▪ `false` ▪ `FALSE` ▪ `TRUE`
`aed:URI`	A Uniform Resource Identifier (URI).	See RFC 3986.	▪ `http://example.com/` ▪ `urn:uuid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6` ▪ `mailto:jdoe@example.com`
`aed:DateTime`	A timestamp with the absolute date and time and an identified time zone. The lexical form is a subset of **ISO 8601** with required milliseconds and time zone offset.	*YYYY-MM-DDThh:mm:ss.s*+ (+|-)*hh:mm*	▪ `2008-09-12T18:45:43.779-07:00`
`aed:Color`	A color. Currently only supports colors in the RGB color space. The lexical form represents red, green, and blue components, respectively, each supporting a decimal integer value 0-255)	rgb(*digit*+, *digit*+, *digit*+)	▪ `rgb(0, 0, 0)` ▪ `rgb(200, 50, 128)` ▪ `rgb(255, 255, 255)`
`bio:Strand`	Represents a sequence strand relative to the landmark or assembly. The lack of an indicated strand encompasses the semantics of "unknown", and "non-stranded."	One of the following characters: + (U+002B) Forward strand. - (U+002D) Reverse strand.	▪ + ▪ -
`aed:UUID`	A Universally Unique Identifier (UUID) as specified by **RFC 4122** (in canonical form, not as a URN).	*xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx*	▪ `f81d4fae-7dec-11d0-a765-00a0c91e6bf6`

Compatibility

UCSC Browser Extensible Data (BED)

The BED file format, developed at UCSC, is widely used for transfer of simple region coordinates. However, the format has been interpreted and implemented in multiple ways by various software within and outside of UCSC. Some implementations require a TAB delimited format, others require a space-delimited format, and still others accept both. Characters outside of the ASCII character set are not well supported. We created the AED format with very strict and explicit definitions so as to avoid some of these compatibility issues.

Although the AED format is preferred, ChAS supports both the import and export of data in BED format. When exporting data in BED format, ChAS exports only the basic 4-column tab-delimited BED format containing the position and name of each item. If the names of any of your items contain spaces or non-ASCII characters, there is no guarantee that all programs will be able to interpret those names correctly.

When importing data in BED format, ChAS supports the reading of BED files with anywhere from 4 to 12 columns.

- The file must be TAB delimited
- Only ASCII characters should be used
- The values for thickStart and thickEnd will be ignored for display purposes
- The value for itemRgb will be honored for display purposes
- The values for blockCount, blockSizes and blockStarts can be used to import and display data with intron/exon structure, such as genes.
- Formatting rules in the BED header are ignored
- BED files containing multiple tracks are not supported; use a separate BED file for each track.

The UCSC Browser, as well as ChAS, uses the strict definition of BED where chromStart is not allowed to be greater than chromEnd. ChAS will accept import of BED files even if this convention is violated, but will auto-correct and export BED files properly with chromStart $\leq$ chromEnd.

AED has been structured to facilitate as much as possible migration of data rows to and from BED. Starting with existing AED and BED files, data records from AED may be transferred to BED by using:

- The “Export” function from inside ChAS (recommended)
- A text editor (not recommended) if the AED files are first prepared in the following manner:
  - Remove all fields except for bio:sequence, bio:start, bio:end, and aed:name.
  - Ensure that no non-ASCII characters are included. (The treatment of non-ASCII characters by a BED processor is undefined.)
  - Ensure that no name contains whitespace characters.
  - Data rows from the first four columns of a BED file can be transferred to AED with no constraints as long as the columns are delimited by TAB.

Microsoft Excel and Other Spreadsheet Applications

Though not recommended, an AED file may be edited using most spreadsheet programs that support tab-separated value (TSV) files and that recognize a byte order mark (BOM). An AED file can be edited in Microsoft Excel, for example, using the following rules:

When loading an AED file into Microsoft Excel as a TSV file, make sure that the Unicode code page for the correct encoding is selected (preferred), or accept the default “Windows (ANSI)” code page (which should still recognize Unicode characters if the correct BOM is present in the file).

When saving an AED file from Microsoft Excel, make sure the “Unicode Text” type is selected. This will result in a file encoded in UTF-16LE, which is still a valid AED file as it begins with the appropriate BOM.

Microsoft Notepad and Other Text Editors

Though not recommended, an AED file may be edited by any text editor that supports Unicode and that uses a byte order mark (BOM) to indicate the charset. The version of Microsoft Notepad in Windows XP, for example, will both correctly read text files marked with a BOM and save text files using the appropriate BOM if the following rules are followed:

When saving an AED file from Microsoft Notepad, make sure the encoding is set to “UTF-8” or “Unicode”.

For other text editors, make sure the correct preferences are set both to recognize and write BOMs for files.

Text Editors

EmEditor <*http://www.emeditor.com/*> is a commercial text editor that has extremely good Unicode and BOM support, and is able to open up gigantic text files.

PSPad <*http://www.pspad.com/*> is a free text editor that has particularly extensive Unicode and BOM support and is available in many localizations.

UniPad <*http://www.unipad.org/*> is a shareware text editor that correctly handles Unicode and BOM, and provides a wide range of built-in glyphs for representing Unicode code points that cannot be viewed on most other text editors.

References

ISO 8601: ISO 8601:2004(E): *Data elements and interchange formats — Information interchange — Representation of dates and times.* International Organization for Standardization, 2004-12-01.

Microsoft Byte Order Mark: *http://msdn.microsoft.com/en-us/library/ms776429(VS.85).aspx*

RFC 3986: *RFC 3986: Uniform Resource Identifier (URI): Generic Syntax*. T. Berners-Lee, R. Fielding, and L. Masinter. Internet Engineering Task Force, 2005. http://tools.ietf.org/html/rfc3986

RFC 4122: *RFC 4122: A Universally Unique IDentifier (UUID) URN Namespace*. P. Leach, M. Mealling, and R. Salz. Internet Engineering Task Force, 2005. *http://tools.ietf.org/html/rfc4122*

RFC 4180: *RFC 4180: Common Format and MIME Type for Comma-Separated Values (CSV) Files*. Y. Shafranovich. Internet Engineering Task Force, 2005. *http://tools.ietf.org/html/rfc4180*

Unicode Byte Order Mark FAQ: *http://unicode.org/faq/utf_bom.html*

ChAS Properties and Types

Starting with version 1.1.0, ChAS has adopted the framework underlying AED files as its native framework for identifying and storing properties and value types. Every property of annotations, files, and other objects within the software is now identified by a URI behind the scenes. Furthermore, the types of values given to these properties are the same types available within AED files. The AED framework therefore provides a consistent and pervasive approach to describing entities throughout the application and seamlessly across AED files and other types of files such as CxCHP files.

Identifying Properties within ChAS

Standard AED Property Style

Because every user-accessible property within ChAS complies with the AED framework, any property may be entered in standard `prefix:simpleName` style, just as it would appear in an AED file. For example, the creation date of an entity may be entered using the `aed:created` property name.

The predefined AED prefixes defined for AED files may always be used. Unlike AED files, which allow declaration of arbitrary prefixes with additional namespaces, ChAS has an additional list of predefined namespace prefix associations valid only within the context of the ChAS user interface. These prefixes may be used to refer to the corresponding namespaces with no explicit namespace declaration. For example, the `fish` prefix may be used to refer to FISH namespace properties (e.g. `fish:labs`) with no need to explicitly associate the FISH namespace URI with the namespace. Table C.1 lists the namespace prefixes recognized within the ChAS user interface.

Table C.1 ChAS namespace prefixes

Namespace URI	Prefix	Label
`http://affymetrix.com/ontology/aed/`	`aed`	General
`http://affymetrix.com/ontology/aed/biology/`	`bio`	Biology
`http://affymetrix.com/ontology/aed/style/`	`style`	Style
`http://affymetrix.com/ontology/aed/default/`	(none)	Custom
`http://affymetrix.com/ontology/`	`affx`	Affx
`http://affymetrix.com/ontology/algorithm/`	`alg`	Algorithm
`http://affymetrix.com/ontology/algorithm/option/`	`algopt`	Algorithm Option
`http://affymetrix.com/ontology/algorithm/state/`	`algstate`	Algorithm State
`http://affymetrix.com/ontology/arr/`	`arr`	ARR
`http://affymetrix.com/ontology/chp/`	`chp`	CHP
`http://affymetrix.com/ontology/netaffx/`	`netaffx`	NetAffx
`http://affymetrix.com/ontology/genome.ucsc.edu/bed/`	`bed`	BED

Table C.1 ChAS namespace prefixes

Namespace URI	Prefix	Label
`http://affymetrix.com/ontology/projects.tcag.ca/variation/`	`dgv`	DGV
`http://affymetrix.com/ontology/www.genome.ucsc.edu/fishClones/`	`fish`	FISH
`http://affymetrix.com/ontology/www.ncbi.nlm.nih.gov/omim/`	`omim`	OMIM
`http://affymetrix.com/ontology/www.pubmed.gov/`	`pubmed`	PubMed
`http://affymetrix.com/ontology/www.ncbi.nlm.nih.gov/RefSeq/`	`refseq`	RefSeq
`http://affymetrix.com/ontology/www.genome.ucsc.edu/genomicSuperDups/`	`superdup`	Segmental Duplications

ChAS Property Style

Properties may also be entered by the user in a more attractive format specific to ChAS, consisting of the property simple name followed by a space and a special label in parentheses to identify the namespace. The namespace label is presented in the table above. For example, the `fish:labs` property may also be entered as labs (`FISH`).

The ChAS style is the default format for display within ChAS. Regardless of what format you use to enter a property, it will be displayed in its ChAS style.

Full URI Property Style

Any property, including those with predefined namespace prefixes, may be entered using its full canonical URI. This allows entry of unknown third-properties in other namespaces. If the property is in a known namespace, it will be displayed using the ChAS property style described above. For example, the `fish:labs` property may be entered using the full URI `http://affymetrix.com/ontology/www.genome.ucsc.edu/fishClones/labs`; the property will be displayed as labs (`FISH`).

Properties may be entered in ChAS three different ways:

- **ChAS style (default)**: `labs (FISH)`
- **AED style:** `fish:labs`
- **AED URI:** `http://affymetrix.com/ontology/www.genome.ucsc.edu/fishClones/labs`

Identifying Value Types within ChAS

With the incorporation of the AED framework, ChAS now uses the same AED types used within AED files. Because there is a small, fixed set of types, within the application user interface these types are always selected using a prepopulated list such as a drop-down list control. The types are displayed using special user-friendly names, even though they refer to the same corresponding types within AED files. These types and their ChAS style labels appear below.

- aed:Boolean - Boolean
- aed:Color - Color

- aed:DateTime - DateTime
- aed:Integer - Whole Number
- aed:Rational - Decimal Number
- aed:Strand - Strand
- aed:String - Text
- aed:URI - URI
- aed:UUID - UUID

Some types may not be available in some contexts.

Automatic Conversion of CxCHP Headers to Properties

The file format used by CxCHP files uses a separate approach for describing properties and value types. However, there exist equivalent AED types for the types used by CxCHP files. Moreover, the header names used in CxCHP files are logically grouped into categories by certain prefixes. ChAS uses these groupings to place CxCHP headers into AED namespaces when loading CxCHP files. ChAS also modifies the names of the headers to make them more readable and to provide consistency with other AED properties used throughout the application.

Header Name Conversion

ChAS uses the following namespaces when converting CHP header names based on the header name prefix.

Table C.2 CHP header name conversion to AED namespace

CHP Header Name Prefix	AED Namespace
`affymetrix-algorithm-`	`http://affymetrix.com/ontology/algorithm/`
`affymetrix-algorithm-param-`	`http://affymetrix.com/ontology/algorithm/`
`affymetrix-algorithm-param-option-`	`http://affymetrix.com/ontology/algorithm/option/`
`affymetrix-algorithm-param-state-`	`http://affymetrix.com/ontology/algorithm/state/`
`affymetrix-chipsummary-`	`http://affymetrix.com/ontology/chp/summary/`
(all others)	`http://affymetrix.com/ontology/chp/`

After determining the appropriate namespace to use, ChAS removes the header name prefix and modifies the remaining characters according to the following rules (simplified):

1. All beginning uppercase letters are converted to lowercase.
2. All separator characters (such as '-' and '_') are removed.
3. The characters immediately following separators are converted to uppercase.

For example, the CHP header `affymetrix-algorithm-param-option-gender-override-file` will be converted to `optionGenderOverrideFile` and placed in the `http://affymetrix.com/ontology/algorithm/` namespace.

Converted Properties

ChAS performs further special conversions on the following header parameters for historical and consistency reasons.

Table C.3 Converted Properties

CHP Header Name	Property
affymetrix-array-type	http://affymetrix.com/ontology/arrayType
affymetrix-chipsummary-snp-qc	http://affymetrix.com/ontology/chp/summary/snpQC
affymetrix-chipsummary-MAPD	http://affymetrix.com/ontology/chp/summary/mapd

Derived Properties

The following properties are each assigned to a file property derived from one or more header parameters, CxCHP file attributes, or other information, in the given order:

http://affymetrix.com/ontology/aed/created (aed:created)

1. The file creation time in the Calvin generic data header.
2. The create_date header parameter.
3. The create-date header parameter.

http://affymetrix.com/ontology/aed/modified (aed:modified)

1. File system last modified date.

http://affymetrix.com/ontology/algorithm/annotationFile (alg:annotationFile)

1. The affymetrix-algorithm-param-state-annotation-file header parameter.
2. The affymetrix-algorithm-param-cn-annotation-file header parameter.
3. The affymetrix-algorithm-param-mapfile header parameter.
4. The affymetrix-algorithm-param-option-annotation-file header parameter.

http://affymetrix.com/ontology/algorithm/parameterFile (alg:parameterFile)

1. The affymetrix-algorithm-param-state-config-file header parameter.
2. The affymetrix-algorithm-param-config-file header parameter.
3. The affymetrix-algorithm-param-paramfile header parameter.
4. The affymetrix-algorithm-param-option-config-file header parameter.

http://affymetrix.com/ontology/algorithm/referenceFile (alg:referenceFile)

1. The affymetrix-algorithm-param-state-reference-file header parameter.
2. The affymetrix-algorithm-param-reference-file header parameter.
3. The affymetrix-algorithm-param-reference-mdlfile header parameter.

ChAS Browser NetAffxGenomic Annotations

Homo Sapiens Database Files

NetAffxGenomic Annotation Files

Source of Content

NetAffxGenomic Annotation files are used by the ChAS Browser to display recent snapshots of genomic annotations downloaded from public databases.

The UCSC Genome Browser is the source of the data that populates the following Browser tracks: BACs, FISH Clones, Genes, Segmental Duplications, sno/miRNA, and Cytobands. UCSC was also the source of OMIM data for ChAS Browser file versions NA31, 32, and 32.1.

The Database of Genomic Variants is the source of the data displayed in the DGV track. The last update made to the DGV site as of February 2013 in the original format was in November of 2010 (v10),.

The OMIM database (with curation and processing assistance from UCSC) is the source of the data displayed in the OMIM Genes and OMIM Phenotype Loci tracks present in the ChAS Browser file v version of NA32.3.

NA33: The date stamps of the download of data from teh UCSC, DGV, and OMIM data for the hg19/GRCh7 based NA33 ChAS Browser NetAffxGenomic Annotation file <NetAffxGenomic AnnotatonsHomo-sapiens.hg19.na32.3db> are:

- TBD
- TBD
- TBD

NA32.3: The date stamps of the download of data from the UCSC, DGV and OMIM data for the hg19/GRCh37 based NA32.3 ChAS Browser NetAffxGenomic Annotations file <NetAffxGenomicAnnotations.Homo_sapiens.hg19.na32.3.db> are:

- UCSC: 15 June, 2012
- DGV: November, 2010 (v10)
- OMIM: 23 November, 2012

NA32 and NA32.1: The date stamps of the UCSC (including OMIM) and DGV data for the hg19/GRCh37 based NA32 and NA32.1 ChAS Browser NetAffxGenomic Annotations files <NetAffxGenomicAnnotations.Homo_sapiens.hg19.na32.db> and <NetAffxGenomicAnnotations.Homo_sapiens.hg19.na32.1.db> are:

- UCSC/OMIM: 11 March, 2011
- DGV: November, 2010 (v10)

NA31: The date stamps of the UCSC (including OMIM) and DGV data for the hg18/NCBI 36 based NA31 ChAS Browser NetAffxGenomic Annotations files <NetAffxGenomicAnnotations.Homo_sapiens.hg18.na31.db> are:

- UCSC/OMIM: February, 2010

- DGV: November, 2010 (v10)

IMPORTANT: It is VERY highly recommended to confirm findings obtained using the ChAS Browser's NetAffxGenomic Annotations file contents by linking out to the current UCSC, DGV, and ENSEMBL data bases using the ChAS software coordinates. See *Going to External Websites* on page 140.

Genomic Position Coordinates

There are multiple conventions and file formats to describe locations in chromosomal DNA sequences. Here we describe a few of the issues as they relate to ChAS.

Genome Assemblies

First, it is important to know which set of DNA sequences is being used as the reference. For the human genome, the reference assembly is available for download from public sources such as UCSC and Ensembl. Those two sites currently use identical genome assemblies, but refer to them by different names. UCSC uses names such as “hg17”, “hg18” and “hg19”. The identical genome assemblies are known as “NCBI35”, “NCBI36” and “GRCh37” at Ensembl. Assemblies at NCBI can have a decimal point as well, for example, “36.3” or “37.1”. For positions on the chromosomes 1-22, X and Y, there is no difference between assemblies “36.1”, “36.2” and “36.3” and we expect the same will be true for future “point” releases.

SNP and Marker Positions

When referring to individual positions on a chromosome, such as the positions of SNPs, it is sufficient to give a single coordinate. There are different conventions about whether to consider the first DNA base pair on the chromosome as position 0 or position 1.

For SNP marker positions, all of the following consistently use a 1-based index position coordinate: CYCHP files, CNCHP files, NetAffx detail pages for SNP markers, NCBI pages for SNP positions of dbSNP entries, and the Graphs Table in ChAS 1.0.

Consider the (randomly-chosen) SNP marker “S-2MOG4” from the Cytogenetics Array. This marker is designed to correspond to the SNP with ID “rs7641618” in the dbSNP database. The NCBI website reports the position as chr3:186255049 on NCBI37 (hg19) or chr3:187737743 on NCBI36 (hg18). On the NetAffx website, currently still based on NCBI36 (hg18), the identical coordinate chr3:187737743 is also given for S-2MOG4. The same coordinate value is given in CYCHP files and in the ChAS graphs table. Refer to *http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?type=rs&rs=7641618* for this particular example.

The Cytogenetics Array also contains copy-number markers which are not based on SNP positions. For these markers, we continue to use a 1-based index position. Unlike the case for SNPs, there is no particular single base pair that the marker corresponds to. The convention in CYCHP files is to use the position of the first DNA pair corresponding to the position where the marker hybridizes with the DNA. When two or more markers on the Cytogenetics Array have the same start position on a chromosome, the coordinate of one of them will be shifted by one or occasionally a few more bases such that each marker is reported at a unique position.

Segment Positions

Affymetrix uses the BED or AED file format for storing and sharing region files between software. The BED format was created by UCSC for use with their genome browser, and is also used in other software. The AED format was created by Affymetrix for use with ChAS and possible future software, but used the BED format as a starting point.

The BED file format is explicitly defined to use a 0-based coordinate system where the second column (chromStart) in the file is the position of the first base-pair and the third column (chromEnd) is the position of the last base-pair plus one. Another way of saying this is that the start index is inclusive and the end index is exclusive. As an example, to refer to the first 100 based on the chromosome, you would use chromStart=0 and chromEnd=100. The length of any region is always given simply by chromEnd minus chromStart.

The UCSC browser strictly requires that chromStart not be larger than chromEnd. In order to support file outputs from non-conforming programs, ChAS will accept BED files where chromStart > chromEnd. It will simply switch those two coordinates and act as if the coordinates were given in the correct order.

Since a SNP has, by definition, a length of one base-pair, the proper way to represent a SNP position is with chromEnd = chromStart + 1. The UCSC browser does allow chromStart to be equal to chromEnd. But this is used for representing insertion points, and is not used to represent SNP positions. Because the AED format was intended to be compatible with BED format, we use the same coordinate system.

In order to make the ChAS Segments Table (and other segments-based tables) resemble the BED and AED output files, the Segments Table uses these same BED-like coordinates. Each segment is defined to include the DNA position of its starting array marker and its ending array marker.

For example, suppose there are three markers with the following positions on a chromosome given in the CYCHP file: Marker A at 1000, Marker B at 2000, Marker C at 3000. Marker positions in the CYCHP file are 1-based index positions. To represent these in a BED file, we would need a file like this:

Chr3 999 1000 markerA

Chr3 1999 2000 markerB

Chr3 2999 3000 markerC

If there were a segment starting at markerA and ending at markerC, we would need to represent it in a BED or AED file as:

Chr3 999 3000 segment_1

It would be represented in exactly the same way in the ChAS Segments Table. The length of this segment is 2001, because it includes the DNA bases at markerA and the marker at markerC.

Thus there will appear to be a discrepancy between the Segments Table and the Graphs Table for the start of the segment because the Graph Table will say markerA is at position 1000 (1-based index) while in the Segment Table, the segment starts at position 999. The end position of 3000 for the segment agrees with the position of markerC. This is not a mistake; it reflects the fact the Segments Table uses BED/AED-style coordinates.

Difference Between ChAS 1.0 and ChAS 1.1

ChAS 1.0 unintentionally failed to correct the CxCHP coordinates to match BED conventions. Thus the Segments Table and AED files exported from ChAS 1.0 were shifted one position higher on the chromosome. Since ChAS 1.1 allows the direct export of BED files which can be loaded directly into the UCSC browser, this correction has been made so that the segment borders line-up precisely with SNP positions.

Comparing ChAS 1.0 with ChAS 1.1, you will see no difference in the Graphs Table, since the 1-based coordinate convention is retained for compatibility with NCBI and NetAffx websites, but you will see a difference of 1 base position for the Segments Table and AED outputs since we have corrected our coordinates to match the BED-style 0-based coordinates.

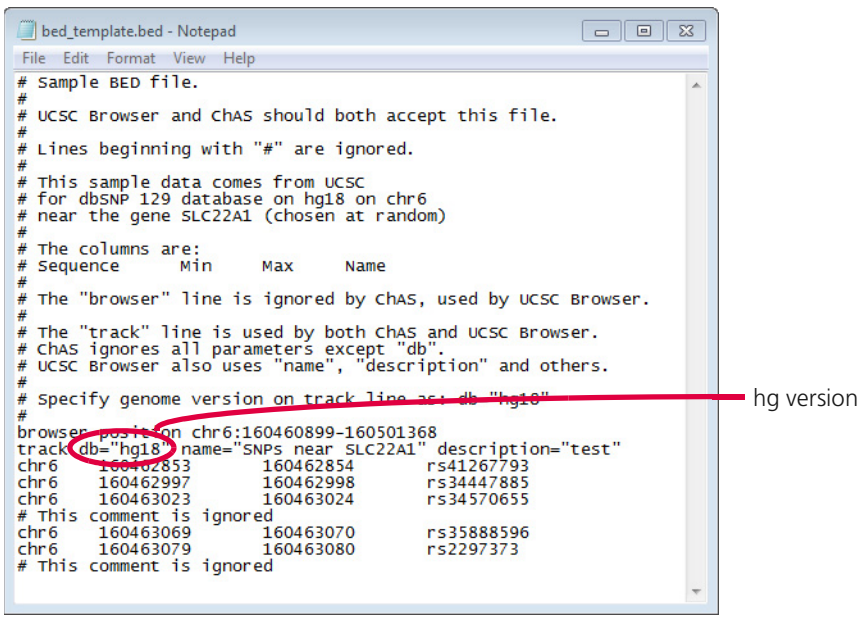
There are no additional changes in ChAS 1.2.

BED Files

Editing BED Files

A BED file is essentially a tab-delimited file (Figure F.1).

Figure F.1 Example BED file



Using a text editor (not a spreadsheet application like Microsoft Excel) to edit BED files is recommended. Text editors such as WordPad and NotePad work well.

Editing a BED file using a spreadsheet application such as Excel is not recommended because these programs may not preserve the correct BED file format. For example, when exporting data from Excel into tab-delimited text,

Excel may add quotation marks around some text, which would cause the file to be invalid and unusable with ChAS or other applications.

There is no easy way to prevent Excel from adding extra quotation marks which corrupt the output. Advanced Excel users can use macro programming to create special output formats. Other options include:

- Do not use Excel to edit BED files. You may use a text editor, but be certain to separate the columns with TAB characters and do not use non-ASCII characters. The BED format was not designed with such characters in mind; therefore, problems may occur when you try to share these files with others. ChAS will reject BED files containing non-ASCII characters, and will never export non-ASCII characters into a BED file.
- After exporting a BED file from Excel, edit the file in another application to remove extra quotation marks.
- Use AED format for the data, then use ChAS to export to BED format if needed.
- Be careful to create a BED file that does not cause Excel to add quotation marks. Do not include the **itemRgb** column, quotation marks, or special characters in a track line. For example, the following is an acceptable track line: **Track db=hg19 name=My_Track description=This_is_my_data**

CytoScan® Algorithms and QC Metrics

Algorithm Overview

This section provides a high level overview of how copy number calls are generated within the software. The copy number workflow starts with the intensities on the array, include normalization and scaling, reference set ratios, Log2 transformation, CN state segmentation, and how CN segment calls are made.

Feature Identification and Signal Extraction

Affymetrix® GeneChip® Cartridge Microarrays are scanned on the GeneChip® Scanner and processed by the AGCC scanner software package. AGCC aligns a grid on the DAT file (the original scanned image) to identify each microarray feature and calculates the signal from each feature. This process uses the .DAT file, containing the raw signal, and creates a .CEL file, which contains a single signal intensity for each feature. The .CEL file is used for all downstream analyses.

Single Sample CytoScan® Workflow

Beginning with the raw signal data in the CEL file, the Single Sample CytoScan® Workflow implements a series of steps that perform probe set summarization, normalization, removal of variation caused by known properties and residual variation, and completing with calling genotypes, copy number segments and LOH segments.

A brief overview of each step performed by the CytoScan® Workflow is shown in Figure G.2 on page 318. In addition, a rough sketch of Analysis Pipeline (for Single Sample Analysis) is demonstrated in Figure G.2 on page 318.

Figure G.1 CytoScan® Workflow Overview. Steps related to probes used for copy number determination run down the left and those steps used for SNP probes run down the right side of the diagram

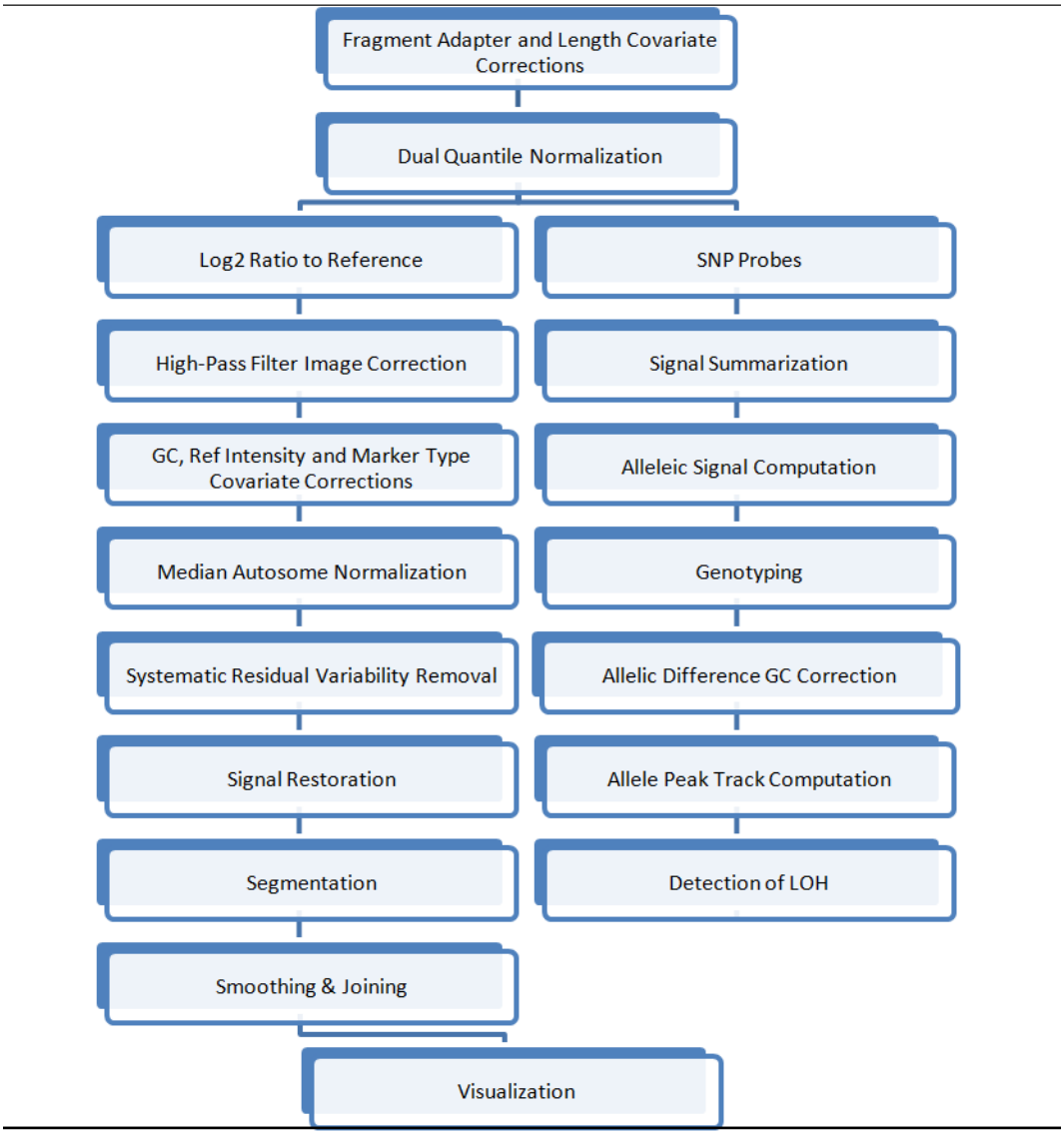
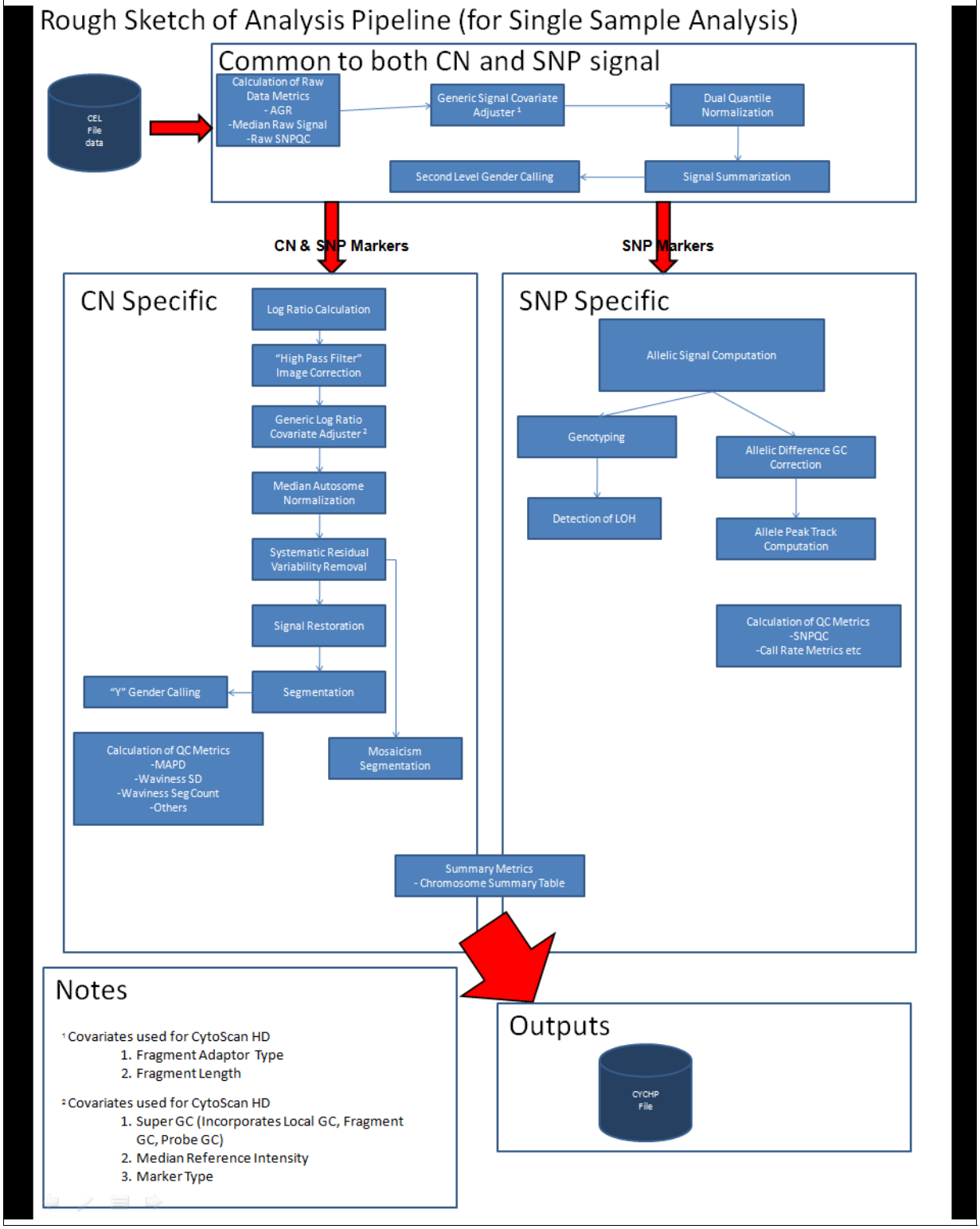


Figure G.2 Rough sketch of Analysis Pipeline (for Single Sample Analysis)



Signal-Level Covariate Adjustors

The first level of covariate adjustors operate on the raw signal.

Fragment Adapter Covariate Adjustor

After the Nsp I restriction digest, an Nsp I-specific adaptor is ligated onto the cohesive end termini. Since Nsp I is a 6-nucleotide cutter with degenerate sites, meaning that they contain one or more base pairs that are not specifically defined, these ends are of various sequences and the ligated adaptors are a variety of sequences. The exact sequences of the cut site and ligation adaptor have an effect on the overall efficiency of ligation and subsequent PCR amplification. The Adaptor Covariate Adjustor corrects for these differences by normalizing the signals for each adaptor/cut site sequence class to an overall median.

Fragment Length Covariate Adjustor

The length of each Nsp I fragment impacts the efficiency of PCR amplification and therefore the signal. Fragments of 300-500 bp are amplified with the highest efficiency and the degree of amplification tapers off as the fragments get longer. The Length Covariate Adjustor corrects for these differences by normalizing the signals for a series of fragment size bins to an overall median.

Dual Quantile Normalization

Dual quantile normalization is simply a two-phase process where probes used for copy number detection and probes used for SNP genotype detection are normalized separately. In both cases, a normalization sketch is built using the autosomal probes in the reference set. The normalization sketch is the prototype distribution of probe intensities that defines what this distribution looks like for all arrays. The single sample autosomal probes are fit to the sketch and the X and Y probes are interpolated into the distribution.

Quantile normalization makes the assumption that the distribution of probes on the array is fairly consistent from array to array. Since the X-chromosome is one of the largest chromosomes (155Mbp, ~5% of the genome), differences between males and females would stretch this assumption. That is why the quantile normalization focuses on creating an autosomal sketch and normalizing the autosome to it. The X and Y chromosome probes are then handled in a special way. Each of them is matched to the closest pre-normalization signal value. Based on that match, their normalized signal should be close to the signal for the very same autosomal probe. So the normalized values for X and Y probes are simply "looked-up" in the pre-normalization autosomal sketch, and transformed to the post-normalization value.

Copy Number Workflow

Log2 Ratio Calculation

Log2 Ratios for each marker are calculated relative to the reference signal profile. The Log2 Ratio is simply $\text{Log2}(\text{sample}_m) - \text{Log2}(\text{reference}_m)$, for each marker, "m".

High Pass Filter Image Correction

Since most probes map to genomic markers associated with a normal copy number, most Log2 Ratios should be centered at a value of zero. Also, since markers from any genomic region are scattered across the surface of the microarray, regions of altered copy number will not appear as regional changes on the microarray image.

Some samples do reveal spatial trends away from zero that are gradual and this spatial bias when scattered back across the genome exhibits itself as added noise in the Log2 Ratios. The High Pass Filter Image Correction identifies these gradual spatial trends and adjusts Log2 Ratios to remove the spatial bias and lower the level of noise.Log2 Ratio-Level Covariate Adjustors

Log2 Ratio-Level Covariate Adjustors

Super GC Covariate Adjustor

The GC content of genomic DNA sequence impacts probe signal dose-response and therefore probe Log2 Ratios. The sequence GC content of the microarray probe impacts hybridization kinetics. In addition, the genomic GC content of the Nsp I fragment and the 500 kbp surrounding the probe (local GC) all impact the efficiency of target preparation in the genomic region of each probe. The super GC covariate adjustor combines the probe GC content, the fragment GC content and the local GC content into one covariate that corrects for Log2 Ratio differences based on the combination of GC contents associated with each probe.

Reference Intensity Covariate Adjustor

Probes in different intensity categories have different dose responses in Log2 Ratio space. Using the Reference Set probes to define bins based on probe intensity, the single sample probes are binned and the median of the distribution of Log2 Ratios within each bin is adjusted to the median Log2 Ratio of the corresponding bin from the reference set.

Marker Type Covariate Adjustor

Polymorphic probes designed for SNP detection and non-polymorphic probes designed for copy number detection have different properties and different dose responses. The Marker Type Covariate Adjustor normalizes the median Log2 Ratios of SNP and CN markers to account for differences in Log2 Ratios between the two groups.

Median Autosome Normalization

This final level of normalization simply shifts the median Log2 Ratio of the autosomes to a copy-number state equal to 2, i.e. a Log2 Ratio of 0.

Systematic Residual Variability Removal

Even after all of the Covariate Adjustors, there is some residual variation with unknown origins. During product development we have introduced variation into the protocol in an attempt to capture other forms of unanticipated variation. The Systematic Residual Variability Removal step matches sample variability to the residual variability of the reference set, and when matched, corrects the data to remove the residual.

Signal Restoration

Signal restoration is an application of Bayes wavelet shrinkage to the Log2 Ratios associated with markers. The result of this transformation is an overall reduction in variation around the local mean of Log2 Ratios. In this context, “local” means a region consisting of a small set of markers upstream and downstream from a given marker. The resulting data can be viewed as the Weighted Log2 Ratio in ChAS and serves as the input to the segmentation algorithm. The wavelet shrinkage method is augmented by a reduction of influence of outliers when local means are computed.

Segmentation

Copy Number Calls for each Marker based on Log2 Ratios

For CytoScan® HD and CytoScan® 750K, markers are individually assigned a copy number call by a Hidden Markov model (HMM). The sample specific inputs to the HMM are the Weighted Log2 Ratios generated by the Signal Restoration module.

The weighted Log2 Ratios are centered on copy number (CN) = 2. In theory, when Log2 Ratio = 0 then CN = 2, when Log2 Ratio = -1 the CN = 1, etc. In truth, microarrays, or any hybridization-based technology, exhibit Log2 Ratio compression due to many factors, so the Log2 Ratios never exhibit the amplitude expected by the math. The following table shows theoretical and actual Log2 Ratios for different Copy Number States.

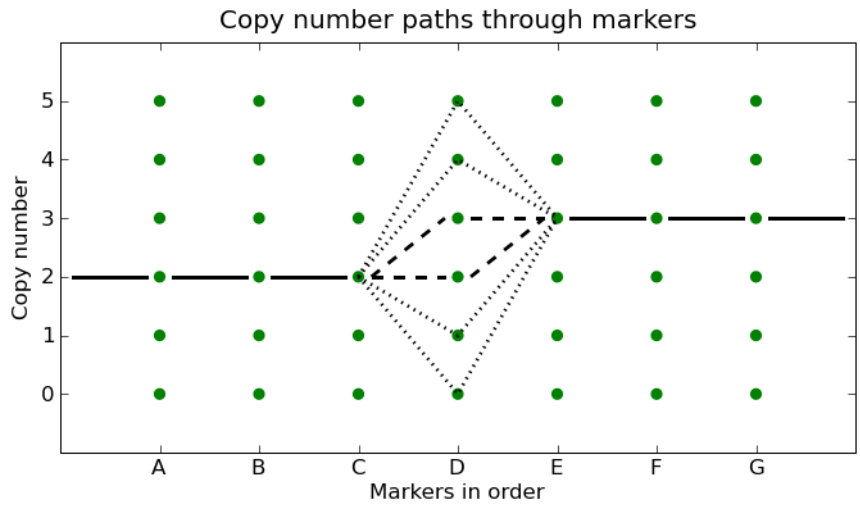
Table G.1 Copy number states and Log2 Ratios

Copy Number Truth	Theoretical Log2 Ratio	Actual Log2 Ratio
1	-1	-0.45
2	0	0
3	0.58	0.3

The actual Log2 Ratios observed are best derived from a very large data set with well-characterized copy number changes. To this end, we have analyzed over 1400 samples that have copy number changes across 75% of the genome and have established stable empirical values for these expected Log2 Ratios. These values, as well as the dispersion characteristics of the Log2 Ratio data, are used as inputs to the HMM along with the weighted Log2 Ratios of the sample data.

The HMM uses these inputs to convert observed Log2 Ratios into a CN state for each marker. It uses a table of transition probabilities that express the probability of changing from any CN state to another. As can be seen in the following example (Figure G.3), there are many potential paths through the possible CN states of a set of markers.

Figure G.3 Potential CN paths through a set of markers



The HMM uses the Viterbi algorithm to calculate the most probable path through the set of markers using the transition probabilities between each pair of CN states. Essentially, the graph of potential CN states is the “hidden” layer of the HMM, and the measure Log2 Ratios are the observed layer. The HMM algorithm finds the most probably CN states given the observed Log2 Ratios.

Segment Formation

Once markers are assigned Copy Number States by the HMM, contiguous stretches of adjacent markers ordered by chromosome position having the same state are aggregated into segments. These segments are described in a segment table within the resulting CYCHP file that provides for each segment, the common Copy Number State, the number of markers in the segment, the genomic marker position that initiates the segment and the genomic marker position that terminates the segment.

Enforce Minimal Segment Length

Default parameters enforce a minimum segment length of 5 markers. This is a subjective choice of parameter that implicitly states that the user is not interested in segments with fewer markers than the minimum. The algorithm that enforces minimum segment size distributes markers from any segment with fewer than the minimum to its larger neighboring segments by changing the copy number call on the modified markers to conform with those of the neighbors.

Smoothing & Joining

To stabilize the calling of copy number gains or losses, the ChAS software implements a smoothing step. Smoothing will combine adjacent segments that are both gains, even if they are not the same Copy Number State. For example, smoothing will combine a set of adjacent segments of Copy Number State 3 and 4 into one segment and assign it the most prevalent Copy Number State of the markers in the original segments, (rounding up for gains and rounding done for losses in case of a tie). Smoothing will also combine Copy Number States 0 and 1. But smoothing will not combine gains with losses or either with normal segments.

Joining combines segments of gains or losses if they are separated by small spans of normal copy number segments. The default value defining “small spans” in ChAS is ?50 markers and ?200 kbp. Small segments of less than 50 markers/200kbp of normal copy number are removed, and the adjacent gain segments are joined. Likewise for flanking loss segments. This is a dynamic process in ChAS, in that smoothing and joining can be turned on and off, and parameters altered, resulting in modifications of the displayed segments, but not altering the underlying CNState graph.

Segment Table Output

The final result of the copy number pipeline is a table of segments identified in the sample. The table in the CYCHP file includes segments of normal and non-normal copy number. Segments called on the X- and Y-chromosomes are characterized as normal or non-normal using gender information and adjusting for the Pseudo Autosomal Regions (PAR) that are present on the X and Y. In ChAS, the segment table display only shows segments of non-normal copy number.

Mosaicism Segment Algorithm

The algorithm for detection of copy number aberrations in the presence of mosaicism considers single copy deletions and gains. The algorithm is tuned to be most accurate when the normal/expected Copy Number State is two. The algorithm targets detection of changes of approximately 5MB or more in size. Copy number change events less than this size may be detected; however estimates of the mosaicism level will be less accurate. The algorithm considers only a discrete number of mosaicism levels which are set at 30%, 50% and 70%.

To detect mosaicism, it is assumed that the range of log ratios has been broken into a series of bands according to the detection level (30% or greater, 50% or greater, 70%-100% bands) and that log ratios within each band denote a specific copy number change (mosaic) event. The band from 0% change to

30% change serves as the copy number neutral region. Percentage bands can be positive or negative, depending on whether the change is a gain or loss, respectively. Given this information, the Mosaicism algorithm:

1. Computes the running means of Log2 Ratios over N markers for each chromosome. This window of N markers moves one marker at a time for each new mean. The algorithm considers marker number N = 6000 (corresponding roughly to 5MB).
2. Detects places where copy number change regions are entered as mosaic events (i.e. left the copy number neutral region and into at least the 30% band's boundary). Band ranges were determined empirically via titration experiments.
3. Identifies the band in which the most extreme value of the mosaic event falls, and uses this band as the level of mosaicism to report.
4. Determines the segment boundaries of the mosaic events. Estimates of the boundaries are computed based on the level of mosaicism, and the exit and entry points from the copy number neutral (normal) region into the mosaicism band regions.

Limitations of the Mosaicism Algorithm for CytoScan Arrays

- The algorithm is designed to detect only mosaicism between approximately 30%-70% mosaicism for copy numbers between 1 and 3 for regions on the order of 5000 markers in size or larger.
- The endpoint location of mosaic segments is less precise than the copy number segmentation, with endpoint variation within 500 markers being typical for segments of 5000 markers or larger in size.
- Some regions of full integer copy number 1 or 3 below 5000 markers in length may be incorrectly called as mosaic segments.
- Some regions of copy number below 1 or above 3, mosaic or otherwise, less than 5000 markers in length may be incorrectly called as mosaic segments.
- Mosaic segments are not generated for the Y chromosome.

SNP Marker Workflow

Signal Summarization

CytoScan® HD and CytoScan® 750K contain 6 probes for each SNP probe set, 3 targeting each allele. The first step of the SNP-specific workflow is to summarize the previously-normalized probe intensities for the A and B alleles, yielding allelic signal values.

Allelic Signal Computation

For each marker, the Allelic Difference is calculated as the difference between the summarized signal of the A allele minus B allele, standardized such that an A-allele genotype is scaled to a positive value, and the B allele is scaled to a negative value. The standardization is determined based on median values for this difference under different genotype configurations determined by the reference set. In this way a homozygous AA maps to approximately +1, and a homozygous BB allele maps to approximately -1, with the heterozygote mapping to approximately 0. Additionally, single A and B allele signals will map to 0.5 and -0.5, respectively. This scaling provides a useful way of discerning two copies of an A allele from a single copy, enabling detection of regions of copy-neutral LOH (e.g. IBD) from hemizygous LOH.

Genotyping

Genotyping for CytoScan® HD and CytoScan® 750K is accomplished using the BRLMM-P algorithm described in the Affymetrix White Paper: BRLMM-P: A Genotype Calling Method for the SNP Array 5.0 (2007).

Allelic Difference GC Correction

Systematic changes in Allelic Differences can be related to differences in GC content. For instance, on a given sample Allelic Differences representing AA and BB genotype markers might get progressively closer or further from each other as the GC content changes. It is assumed that such changes represent unwanted variability. The Allelic Difference GC correction determines differences in the structure of the allelic differences associated with GC and then removes these differences. For CytoScan® HD the super GC covariate is used. For CytoScan® 750K the Local GC covariate is used.

Allele Peak Track Computation

The allele peak tracks for CytoScan® HD and CytoScan® 750K are generated by signal processing the allele differences for each maker to improve the resolution of the tracks. The first step is the identification of local structure for each region by using a sliding window of markers. Within each window, a set of peak estimates are determined. Allelic difference signals for each marker are then shrunk toward the nearest peak, reducing some individual marker variability.

Detection of LOH

The LOH algorithm frames the problem in terms of a statistical hypothesis test. Given a specific region containing *N* SNP markers with heterozygous and homozygous genotype calls, decide between the following two hypotheses:

Null Hypothesis: Region is LOH
Alternative Hypothesis: Region is non-LOH

To decide between the two hypotheses the number of heterozygous calls is compared with a critical value that is computed for each sample. When the number of heterozygous calls is above the critical value, then the alternative hypothesis is favored, i.e. region is not LOH. If there are not a sufficient number of heterozygous calls then the decision is made in favor of LOH. The algorithm moves the region of *N* markers along the genome to determine LOH events. Further details are provided in the Affymetrix While Paper: The Loss of Heterozygosity (LOH) Algorithm in Genotyping Console 2.0.

Array Data QC Metrics

This section provides a high level overview of the key QC metrics used with the CytoScan® Arrays.

Median of the Absolute values of all Pairwise Differences (MAPD)

MAPD is a global measure of the variation of all microarray probes across the genome. It represents the median of the distribution of changes in Log2 Ratio between adjacent probes. Since it measures differences between adjacent probes, it is a measure of short-range noise in the microarray data. Based on an empirical testing dataset, we have determined that array data with MAPD > 0.25 has too much noise to provide reliable copy number calls.

Waviness SD

Waviness-SD is a global measure of variation of microarray probes that is insensitive to short-range variation and focuses on long-range variation. Based on an empirical testing dataset, we have determined that array data with Waviness-SD > 0.12 has either sample or processing batch effects that will reduce the quality of the copy number calls. Elevated Waviness-SD is not always an indication of too much noise. Elevated Waviness with good MAPD and SNPQC metrics can occur in samples with many copy number changes or very large regions of change. It is therefore advised to check the data when observing elevated Waviness with good MAPD and SNPQC.

SNPQC

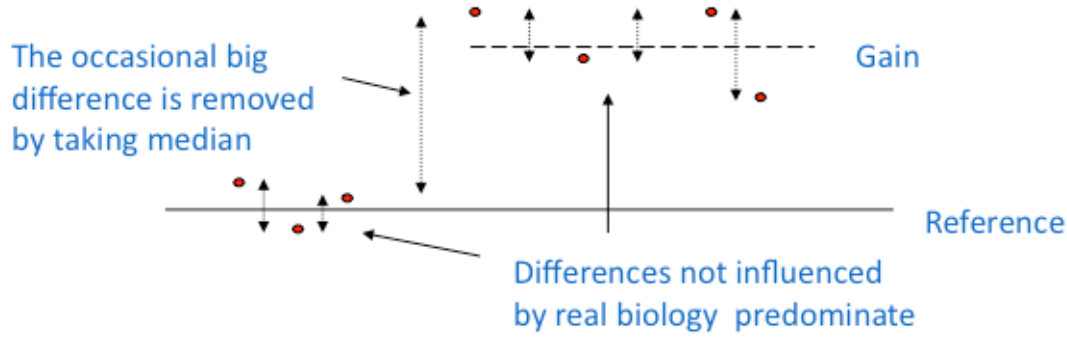
SNPQC is a measure of how well genotype alleles are resolved in the microarray data. Based on an empirical testing dataset, we have determined that array data with SNPQC < 15 is of poorer quality than is required to meet genotyping QC standards.

MAPD – Detailed Description

For quality assessment purposes, we define metrics that assess whether the microarray data is useful for copy number (CN) analysis. One of these metrics is Median of the Absolute values of all Pairwise Differences (MAPD).

MAPD is defined as the Median of the Absolute values of all Pairwise Differences between Log2 Ratios for a given chip. Each pair is defined as adjacent in terms of genomic distance, with SNP markers and CN markers being treated equally. Hence, any two markers that are adjacent on the genome are a pair. Except at the beginning and the end of a chromosome, every marker belongs to two pairs (Figure G.4).

Figure G.4



Formally, if xi: is the Log2 Ratio for marker i:

MAPD = median($|x_{i-1} - x_i|$, with *i* ordered by genomic position)

MAPD is a per-microarray estimate of variability, like standard deviation (SD) or interquartile range (IQR). If the Log2 Ratios are distributed normally with a constant SD, then MAPD/0.96 is equal to SD and MAPD*1.41 is equal to IQR. However, unlike SD or IQR, using MAPD is robust against high biological variability in Log2 Ratios induced by conditions such as cancer.

Variability in Log2 Ratios in a microarray arises from two distinct sources:

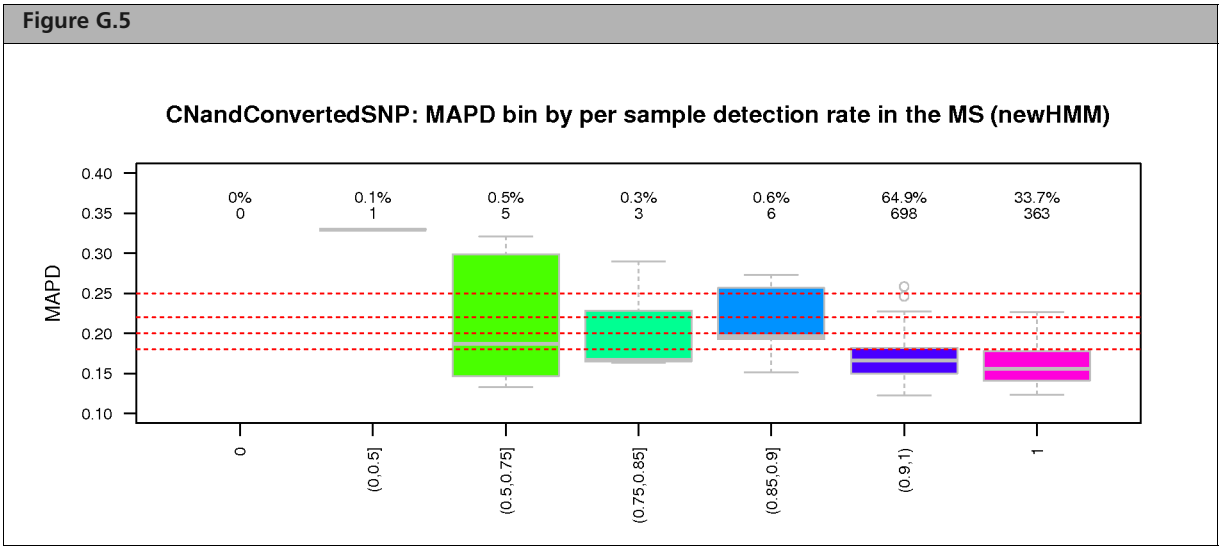
- Intrinsic variability in the starting material, hybridization cocktail preparation, microarray or scanner
- Apparent variability induced by the fact that the reference may have systematic differences from this microarray

Regardless of the source of the variability, increased variability decreases the quality of CN calls.

Effect of MAPD on Functional Performance

As a measure of performance, we measured copy number gain and loss using samples with large chromosome aberrations that spanned approximately 70% of the genome. With this dataset of nearly 1500 microarrays we measured the sensitivity for detecting regions of copy number change across all of these regions. The sensitivity of detecting an aberration on each array was binned into groups of varying sensitivities, and plotted versus MAPD for each array in the following graph (Figure G.5).

Figure G.5



The bins of detection sensitivity are displayed as coordinates along the x-axis, with 0% detection at the left and 100% at the right. The number of arrays is listed above each box plot. The majority of the arrays had sensitivities above 90%. Based on this analysis, we established a QC cutoff for MAPD of 0.25. Arrays with MAPD above 0.25 cannot be reliably used to determine copy number.

Waviness-SD – Detailed Description

For quality assessment purposes, we define metrics that assess whether the microarray data is useful for copy number (CN) analysis. In addition to MAPD (above) we define an alternate form of measurement of variance in the array data that is called Waviness-SD, where SD stands for Standard Deviation.

Waviness refers to an effect seen in all genomic microarrays (see Maroni et al. (2007) Genome Biology **8**:R228) where long-range variation is observed, often associated with regional genomic differences like local GC-content changes.

Waviness-SD is a QC metric that focuses on measuring these long-range effects. As described separately, MAPD is a metric that measures short-range variation, the variation of adjacent probes. The long-range variation measurement is accomplished by calculating the variation in Log2 Ratios across the whole genome and subtracting out the short-range variation, specifically, for autosomal probes:

Define:

X_i as the Log2 Ratios of autosomal probes

And Z_i as the variance between adjacent probes:

$Z_i = X_{2i+1} - X_{2i}$

Waviness-SD is the total variance (X_i) minus the local variance (Z_i):

Waviness-SD = sqrt(Var(X_i)-Var(Z_i)/2)

While this metric is useful in most cases, it does make the assumption that most of the genome is of normal copy number. This assumption may not be reasonable for some types of cancer samples with large amounts of genomic copy number variations, or for multiple-chromosome constitutive trisomies, where a considerable fraction of the genome is duplicated.

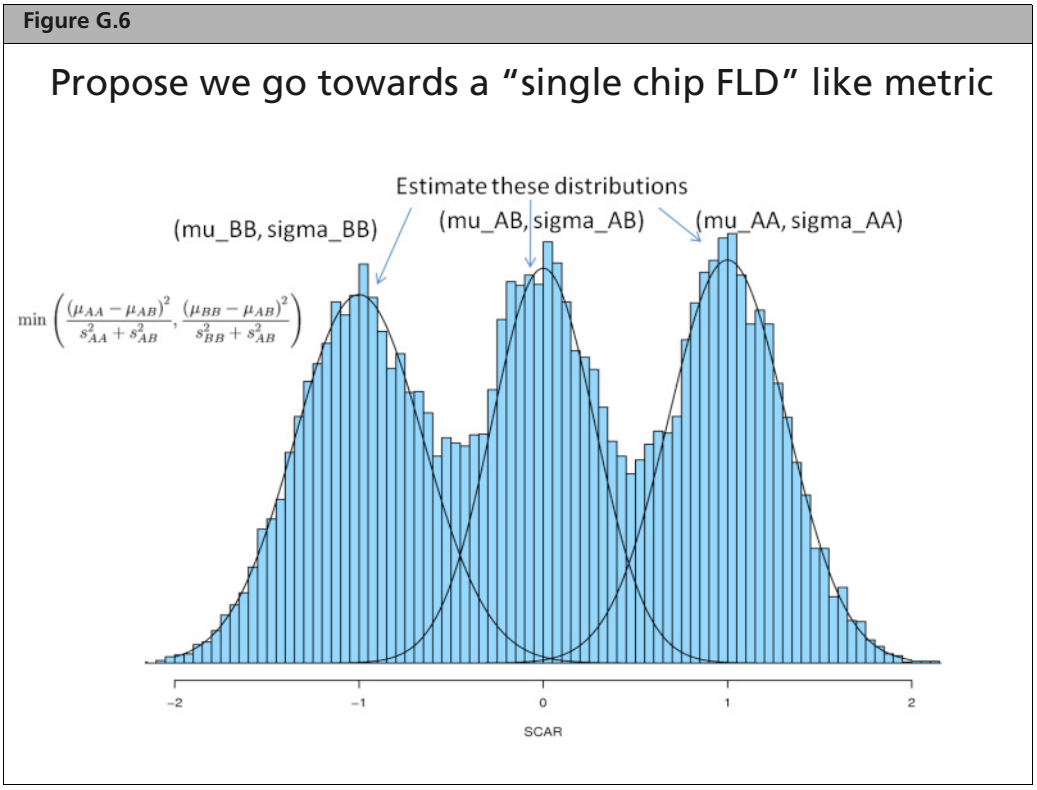
For most samples, a Waviness-SD value below 0.12 for CytoScan® HD or CytoScan® 750K arrays indicates that the long-range variation is within levels that can be accommodated by the CytoScan® algorithms. But a high Waviness-SD measure on a sample with good MAPD and SNPQC metric values should be checked for the presence of large regions of copy number change to assess whether it is a sample effect or a QC failure.

Waviness-SD can be a good indicator of process drift since it measures long-range variation relative to the CytoScan® HD or CytoScan® 750K reference profile. A general rise of Waviness-SD for all samples coming from your laboratory may be an indication of a change of protocol, technique or reagents.

SNPQC – Detailed Description

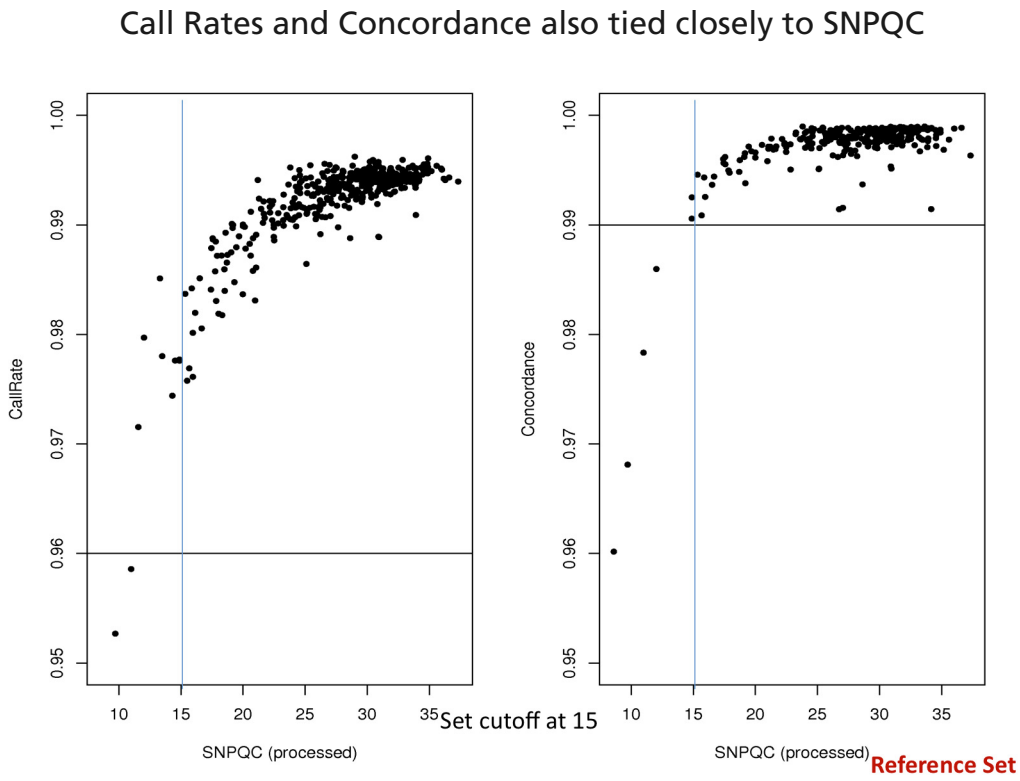
SNPQC is a metric that estimates the distributions of homozygous AA, heterozygous AB and homozygous BB alleles and calculates the distance between them. The better the separation of these distributions, the better the ability to identify a genotype based on its cluster position (Figure G.6).

Figure G.6



SNPQC correlates well with genotype performance, as measured by Call Rate and Concordance to published HapMap genotypes. To establish this relationship, we scored 380 microarrays from the Reference Set by calculating SNPQC, Call Rate and Concordance. The following graphs show the relationships between SNPQC and the other two metrics (Figure G.7).

Figure G.7

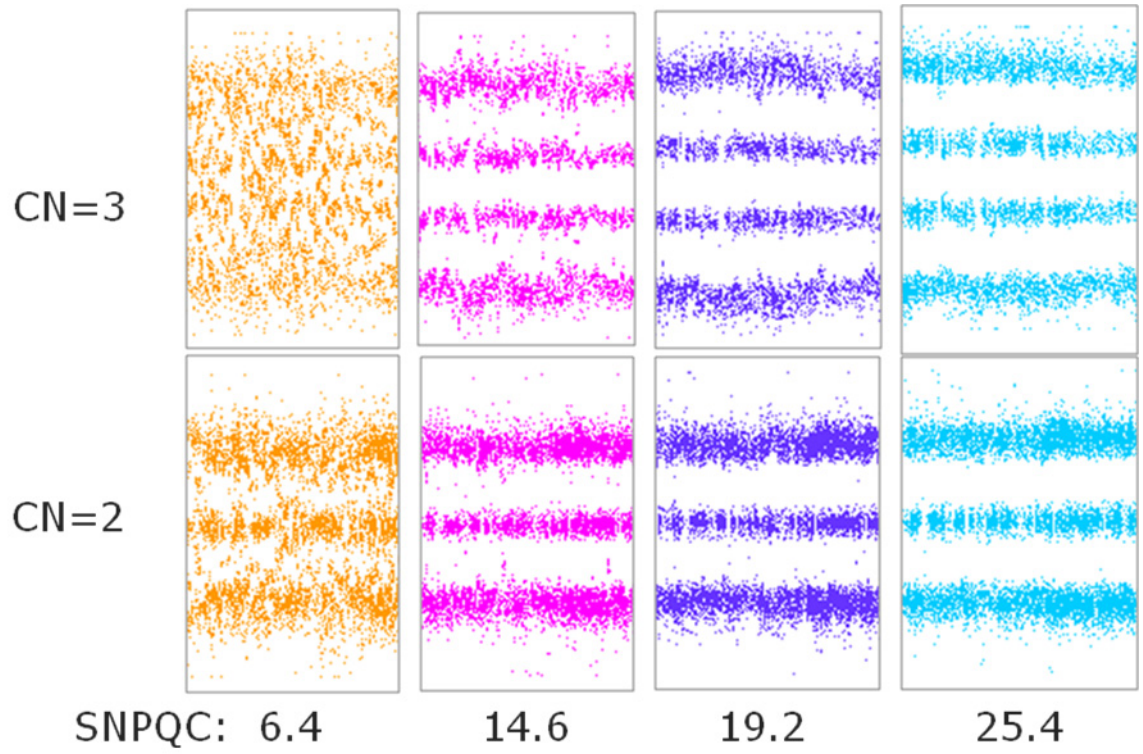


The left panel shows that when SNPQC > 15, Call Rate is above 98%. The right panel shows that when SNPQC > 15, Concordance is above 99%. This functional mapping of SNPQC has allowed us to set a functional threshold for this QC metric at 15. Microarrays with SNPQC ? 15 are considered of high quality and interpretation of the data is possible.

Effect of SNPQC on Functional Performance

SNPQC provides insight into the overall level of data quality from a SNP perspective. The key consideration when evaluating the SNPQC value is to ensure the threshold is exceeded. The quality of the SNP allele data is compromised, and is noisier and more difficult to interpret when the SNPQC values are below the recommended acceptance threshold as illustrated in the figure below. When the SNPQC value is below 15, the noise within the array is higher than normal which compromises the overall data quality and clarity of results. However, when the SNPQC value is above 15, the data is of excellent quality and can be relied upon as robust with regard to performance (Figure G.8).

Figure G.8 Examples of Allele Track data quality at various levels of SNPQC. The lower row of figures show data for a CN=2 and the upper for CN=3 regions. The panels from left-to-right represent increasing SNPQC quality. The functional threshold for SNPQC is 15, so all values above 15 show excellent data quality.



The key consideration is whether the SNPQC value is above or below the threshold value and not the absolute magnitude. As long as the SNPQC value exceeds the threshold there is a retention in the data quality as illustrated by the graphs to the right which demonstrate clear allelic data across a broad range of SNPQC values which exceed the recommended threshold. SNPQC is one of the metrics used to assess array quality and should be helpful towards determining which experimental data sets are of satisfactory quality to continue with subsequent interpretation.

NOTE: For detailed information on algorithms and QC metrics for the OncoScan FFPE Assay array, please refer to the OncoScan Console User Manual (P/N 703195).