



DESIGN REPORT

Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10

27 August 2024

Alessandro Davassi
Senior Bioinformatics Scientist - Lead Programmer

@TwistBioscience #SyntheticDNA

Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10

Data received by Twist Bioscience

- Coordinates of 11450 genomic regions

Genome assembly: mm10

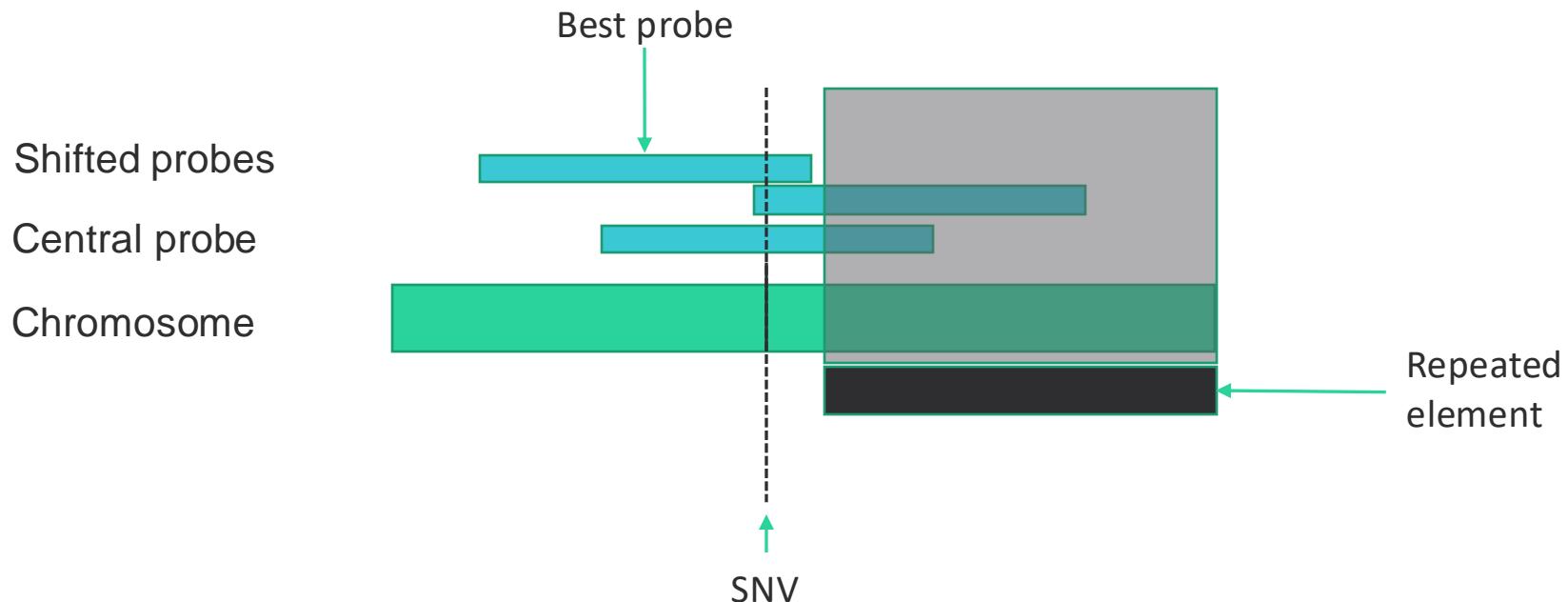


Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10

Design Instructions

Low stringency filters to remove probes on repetitive sequences have been applied.

Genome assembly: mm10



Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10

Design Summary

Number of probes: 9363

Coverage: 82.31%



Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10

Design Information	
Customer	Sanger
Design Name	Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm
DesignID	TE-99621817
Genome	mm10
Probe Length	120
Stringency	medium
Shadow Coverage	0
Tiling	1X
Coverage Summary	
Target Size (bp)	11,450
Total Merged Target Regions	11,405
Target Merged Regions with Probes	9,387
Number of probes	9,363
Number of unique probes after removing sequence duplicates	9,363
Number of probes removed due to repeats	2,013
Non Covered Whole Target Regions (#)	2,018
Non Covered Whole Target Regions (%)	17.69
Non Covered Whole Target Regions (bp)	2,026
Covered Size (bp)	9,424
Design Size (bp)	1,123,408
Overall Coverage (%)	82.31

Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10

Files shared

UCSC_Custom_Tracks_Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10_low_240827114323.bed

all_target_segments_not_covered_by_probes_Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10_low_240827114323.bed

all_target_segments_covered_by_probes_Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10_low_240827114323.bed

merged_probe_file_shareable_Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10_low_240827114323.bed

all_target_segments_with_zero_probes_Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10_low_240827114323.bed

report_Sanger_Mouse_phylo_REDUCED_v2_TE-99621817_mm10_low_240827114323.html

Definitions

- all_target_segments_covered_by_probes: all the target regions that are covered at 100% by probes
- all_target_segments_not_covered_by_probes: target regions that are not completely covered by probes
- report: Details number of 1) input targets 2) targets covered and 3) number of probes
- complete_target_regions_not_covered_by_probes: the target regions for which not a SINGLE probe could be placed for the WHOLE target region.
- UCSC: combined covered, not covered, whole not covered, and targets in a UCSC format with header
- merged probe file: all the genomic regions that are covered by probes



Appendix



Evaluating your custom panel

The bed files provided can be used to evaluate the following:

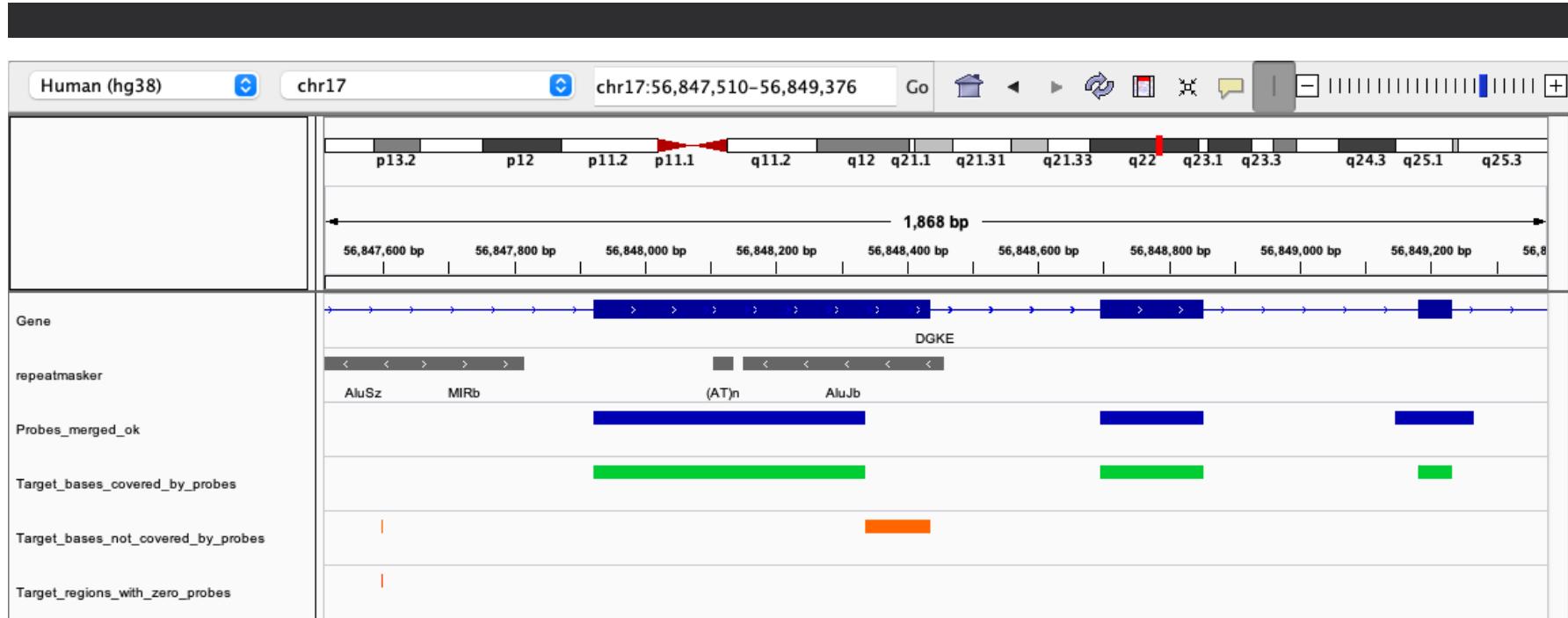
- Your target regions covered by probes: **Target_bases_covered_by_probes.bed**
- The merged probe coordinates: **Probes_merged_ok.bed**
- Your target regions which are NOT covered by probes: **Target_bases_not_covered_by_probes.bed** and **Target_regions_with_zero_probes.bed**

Consult the **Target_bases_not_covered_by_probes.bed** to assess if any of these uncovered regions are critical for your panel. If targets essential to your panel are filtered out because they occur in repeat regions, Twist can reintroduce probes. Doing so, however, will **increase off-target rates and lower the performance of your panel**. In addition, these regions may be challenging to uniquely map in downstream analysis and may result in poor coverage.

The bed files can be loaded into a text editor or into visualiser such as UCSC's genome browser or Integrative Genomics Viewer (IGV).



Visualizing your panel in IGV



Your bed files can be loaded into IGV for quick and easy visualisation of your covered and non covered target regions.

Track 1: Repetitive elements from repeat masker. Probes that overlap significantly with repeats are filtered out to limit off target rate.

Track 2: `Probes_merged_ok`, shows the merged probe footprint of your panel.

Track 3: `Target_bases_covered_by_probes`, are the target regions directly covered by probes.

Track 4: `Target_bases_not_covered_by_probes`, shows the bases within your target regions which are uncovered by probes. This can include partial as well as whole target regions which have been filtered due to repetitive elements.

Track 5: `Target_regions_with_zero_probes` shows the target regions which are completely uncovered by probes due to the presence of repetitive elements. No probes exist anywhere else on the target coordinate. These targets have been completely filtered out of your panel.

Visualizing your panel in UCSC Genome Browser

1. Navigate to:
<https://genome.ucsc.edu/>
2. My Data > Custom Tracks
3. Under Add Custom Tracks select your genome and assembly.
4. Upload the All_tracks.bed file and submit.
5. View in Genome Browser

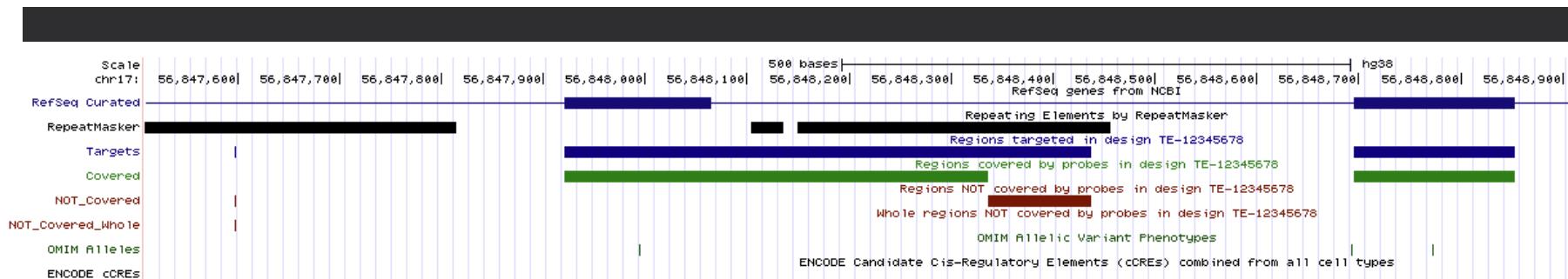
The screenshot shows the UCSC Genome Browser homepage with a navigation bar at the top. The 'My Data' menu is open, showing options like 'Custom Tracks', 'My Sessions', 'Track Hubs', 'Track Collection Builder', and 'Public Sessions'. Below this, the 'Add Custom Tracks' section is visible, where the user has selected 'Mammal' for clade, 'Human' for genome, and 'Feb. 2009 (GRCh37/hg19)' for assembly. A dropdown menu for assembly shows other options like 'Dec. 2013 (GRCh38/hg38)', which is currently selected. The 'Manage Custom Tracks' section shows a table of four tracks: 'Bases_Covered', 'Bases_Not_Covered', 'Target_Regions_With_Zero_Probes', and 'Targets_submitted', each with a description, type (bed), document count (2 or 1), position (chr1), and delete checkbox. Buttons for 'view in' (Genome Browser) and 'add custom tracks' are also present.

Name	Description	Type	Doc	Items	Pos	delete
Bases_Covered	Target bases covered by baits, TE-12345678	bed	2		chr1:	<input type="checkbox"/>
Bases_Not_Covered	Target bases not covered by baits, TE-12345678	bed	2		chr1:	<input type="checkbox"/>
Target_Regions_With_Zero_Probes	Target regions with zero probes, TE-12345678	bed	1		chr1:	<input type="checkbox"/>
Targets_submitted	Regions targeted in design TE-12345678	bed	4		chr1:	<input type="checkbox"/>

<https://genome.ucsc.edu/goldenPath/help/hgTracksHelp.html>



Visualizing your panel in UCSC Genome Browser



Once in the Genome Browser, your bed file will be represented as colored tracks:

Track 1: Repetitive elements from RepeatMasker. Probes that overlap significantly with repeats are filtered out to limit off target rate.

Track 2: Targets, shows all original submitted targets for your panel.

Track 3: Covered, are the target regions directly covered by probes.

Track 4: NOT_Covered, shows the bases within your target regions which are uncovered by probes. This can include partial as well as whole target regions which have been filtered due to repetitive elements.

Track 5: NOT_Covered_Whole shows the target regions which are completely uncovered by probes due to the presence of repetitive elements. No probes exist anywhere else on the target coordinate. These targets have been completely filtered out of your panel.

<https://genome.ucsc.edu/goldenPath/help/hgTracksHelp.html>



