illumina®

Creating TruSeq[®] Custom Amplicon Designs with DesignStudio[™]

Improving the designability of your amplicon panel.

TruSeq Custom Amplicon (TSCA) is a complete amplicon sequencing system that can process up to ~650 Kb of cumulative genomic sequence. This document describes how to improve target designability with DesignStudio software for optimized sequencing results.

Each TSCA project is designed as an entire pool in order to minimize unfavorable probe-to-probe interactions. Gaps in designs can occur for several reasons: poor specificity, probe-probe interaction, GC content outside acceptable limits, or the presence of SNPs.

Probes and amplicons can be selected or deselected in the **Review Design** tab. Before adding new regions, the designs must be reset in the **Manage Targets** tab to avoid the potential interactions between existing and new target region probes. Since the limit on estimated amplicons is well above the orderable limit for a TSCA project, it may be beneficial to initially submit a larger pool of targets to understand which regions, if any, are difficult to design. Common solutions to improve the designability of TSCA projects are described below.

Why is it difficult to design an amplicon against certain targets?

There are several factors that can affect the designability of a target, including:

- Homologs: Including homologs in the same design can lead to low designability due to probes cross-binding in similar regions. Illumina recommends that customers split homologs into separate projects.
- GC content: Target regions with greater than 80% GC content can make design difficult, particularly if these regions are greater than 500 base pairs (bp) in length.
- Homopolymer sequences and repetitive elements: DesignStudio avoids these regions to ensure that probes have better specificity in the genome.
- Presence of SNPs: By default, DesignStudio avoids placing amplicon probes on regions with known variation (dbSNP), resulting in very conservative designs. Users can manually change the **Avoid SNPs** setting to **SNP OFF** for a less stringent design.

One of my targets is undesignable. What can I do to improve the coverage?

Increasing the size of the target region can rescue some previously undesignable regions. This gives DesignStudio more flexibility to fit a higher scoring amplicon over the desired target bases. There are several ways to increase the size of the target, including:

- Increase the exon padding up to 25 bp
- Select **FullRegion** and enter coordinates for each region as a separate target, then add bases upstream and downstream of each target region

Changing the **Avoid SNPs** setting to **SNP OFF** tells DesignStudio to ignore the location of SNPs in the design process. This can help rescue undesignable regions if DesignStudio avoided regions with low frequency or poorly annotated SNPs in the database. DesignStudio uses dbSNP 131 (human hg19), snp 128 (mouse mm9), snp 125 (rat rn4), and Ensembl UMD3.1 snps (cow) to position probes that avoid SNP locations, improving the performance of the design.

Changing the context of the regions can also improve coverage. For example, putting homologous target sequences into the same design can be problematic for specific amplification. Moving problematic regions into a separate design can frequently improve designability.

What does the amplicon score mean?

This score reflects the predicted success rate of amplicons by taking into account the target specificity, GC content, and other design parameters. DesignStudio only returns amplicon designs that are expected to provide a high-performing panel that meets product specifications (e.g., panel specificity, uniformity).

What assay performance can I expect from my TSCA design?

DesignStudio returns high-confidence amplicon designs that deliver unprecedented amplicon multiplexing performance, with specificity and uniformity > 70% and > 80%, respectively. In practice, Illumina has observed specificity and uniformity > 90% for hundreds of designs and thousands of amplicons.

Is there a limit to the amount of contiguous sequence permitted for a target?

A given target region is limited to 24 Kb of contiguous sequence to ensure reasonable design times with current computing resources. For enriching large contiguous stretches of genomic targets, Illumina recommends a hybridization-based assay such as Nextera[®] Custom Enrichment. Alternatively, customers can target only the smaller regions that are essential to the amplicon project.

What is the target proximity notification that results when I submit my targets for an amplicon design?

DesignStudio checks how close targets are in relation to one another. If two genomic targets are found to be within one amplicon's length apart, DesignStudio automatically merges the requested targets to design against one large target. This approach allows DesignStudio to tile amplicons across the region of interest and results in better coverage than if the two targets were independently designed.

How long will it take to complete the design?

Design time is dependent on the number of requested targets and estimated number of amplicons. Small designs can be completed in a matter of hours, while complex designs may take 2–3 days. This computing time ensures that the final design will provide robust multiplex amplicon performance.

How can I remove duplicate exons?

When uploading exon targets (with **Exon** specified in the **TargetType** column) using the **Coordinates** template, if more than one exon is contained within the overlapping sets of coordinates, the database will return a series of exons that may be duplicates. By selecting **Targets** rather than **Regions** at the upper right corner of the data grid, a **D** will be displayed in the **Info** column indicating which targets have been duplicated. Duplicates must be removed before the probe design process can begin.

Illumina • +1.800.809.4566 toll-free (U.S.) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

FOR RESEARCH USE ONLY

© 2013 Illumina, Inc. All rights reserved.

Illumina, IlluminaDx, BaseSpace, BeadArray, BeadXpress, cBot, CSPro, DASL, DesignStudio, Eco, GAlk, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, NuPCR, SeqMonitor, Solexa, TruSeq, TruSight, VeraCode, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners. Pub. No. 970-2013-001 Currrent as of 14 February 2013

