New NxSeq HybCap Custom Kit - order design requirements and conditions

1. **Free custom HybCap™ Oligo Probe design conditions**
    a. LGC, Biosearch Technologies™ provides one free round of custom HybCap Oligo Probe design, including one free round of design revision if the initial design is not approved by the client. This includes:
        - Project consultation with an expert Biosearch Technologies Design Consultant
        - Separating target region(s) into probes of desired length/overlap
        - Screening probes against up to five (≤5) relevant genome(s) to check for probes that could hybridise to non-specific or over-represented (repeat) regions, thus biasing the hybrid capture results.
        - Summary report of relevant probe qualities (such as GC%, ΔG, repeat content)
        - One round of revision to the initial design, including altering the filtering or exclusion parameters if the client does not approve the first design
        - Full report of candidate probe sequences including recommendations on which probes to include in the final set/design, based on client's experimental goals
    b. Please note that the revision consists only of altering the filtering or exclusion parameters. Adding new sequences or altering probe length/overlap are not included during revision and will incur additional charges.
    c. If subsequent rounds of additional design and/or revision are required, these are available upon request for US$500 each, unless initiated by Biosearch Technologies.
    d. NxSeq™ HybCap Kits are not available through distributors for Biosearch Technologies and its affiliates.
    e. Client must provide written/digital approval of final design before production.

2. **Predesigned probe submissions**
    a. Clients can submit predesigned probes and Biosearch Technologies will synthesise those probes and provide them as part of a custom NxSeq HybCap Kit.
    b. The client is responsible for all probe design and screening as well as probe performance.
    c. Clients can submit their probe sequences as FASTA files (Appendix A).

3. **Ownership and confidentiality**
    a. Clients have full ownership of all probe sequences designed by Biosearch Technologies based on the client-provided target regions and sequences.
    b. All information provided to aid in custom probe design is held in strict confidentiality. This includes target sequences, sample info, genome(s), research goals, and more.

4. Compiling target sequences

a. Clients will generate one (1) FASTA or BED file, ordered by highest-to-lowest priority (if applicable), which includes their target regions of interest.

b. The sequence/interval file must be formatted according to the Submission Guidelines (see Appendix A).

c. Total target length should not exceed 200% of desired probe count or target space (e.g. 4 Mb total can be submitted for 2 Mb of target/21-40K set of probes)

d. Remove or flag repetitive sequences (recommended). Sequences should be hard- or soft-masked using a repetitive DNA screening tool (e.g. RepeatMasker or WindowMasker). We recommend screening against a database of taxonomically-relevant repetitive elements as well as for simple repeats and low-complexity DNA.

5. Submitting target sequences

a. Complete the NxSeq HybCap Oligo Probe project submission form and email to hybcap@lgcgroup.com.

b. Once submitted, your Biosearch Technologies Design Consultant will email you instructions to submit your target file and reference genome if necessary. Important submission format/file requirements are provided in Appendix A.

Appendix A: Sequence submission guidelines for a custom NxSeq HybCap probe design

To submit your sequences for a design, we accept sequences in either FASTA format or as coordinates from a reference genome (see below). Please note that we will design HybCap Oligo Probes from ALL sequences or coordinates that you provide. If you only want specific regions of those sequences in the probe set (e.g. exons only), please first curate your targets to only include those specific regions of interest.

1. Target region(s) provided as sequences

Acceptable: ***FASTA DNA sequence format, in plain text file***

Names

- All sequence names must be fully unique
- Allowed characters are letters, numbers, and dashes "-" ONLY (no other characters should be used)
- Spaces and underscores will be replaced with dashes
- Name length of 50 characters or less
- Recommended to incorporate species/locus names

Sequences

- Allowed characters are IUPAC bases
- Alignment gaps ("-") may be present, but will be ignored during probe design

Notes regarding non-ATCG bases in target region sequences

- A single N and/or short stretches of N's will be replaced with T's to facilitate probe design in these regions.
- Longer stretches of N's (e.g 10+ N's) will be skipped over during probe placement.
- Ambiguous base sequences (e.g. Y/M/R/S/W/K) are allowed but will be replaced by ONE random candidate base for manufacturing, since we only synthesise A/T/C/G bases.

2. Target region(s) provided as genome coordinates

Acceptable: ***BED ("Browser Extensible Data") format, in plain text file***

Detailed BED file format details

- Available at: https://genome.ucsc.edu/FAQ/FAQformat.html#format1

Quick summary of BED file content

- **Column 1**: chrom - The name of the chromosome (e.g. chr3, chrY, chr2_random) or scaffold (e.g. scaffold10671)
- **Column 2**: chromStart - The starting position of the target region in the chromosome or scaffold. The first base in a chromosome is numbered 0.
- **Column 3**: chromEnd - The ending position of the target region in the chromosome or scaffold. The chromEnd base is not included in the display of the feature. For example, the first 100 bases of a chromosome are defined as chromStart=0, chromEnd=100, and span the bases numbered 0-99.
- **Column 4**: Optional, e.g. target region descriptor

BED file example

Column 1	Column 2	Column 3	Column 4 (optional)
chr6	1000456	1000972	target_region_1
chr12	6541	7988	target_region_2
chr2	354001	355629	target_region-3

Specific instructions

- Provide link/copy of exact reference genome, otherwise coordinates will be incorrect
- Names of chromosome/contig/scaffolds must match genome entry names exactly
- Plain text file only (tab-delimited)
- Do not submit spreadsheet/Excel files

When you have prepared your sequences, please submit them according to the instructions provided by your Design Consultant.