

Data required for designing a SeqSNP* assay

1. A reference genome file (FASTA format) or details of the publically available reference genome sequence to be used. Please provide us with a link for downloading, or provide the exact version of a genome reference sequence, e.g. Homo sapiens (assembly GRCh38.p11); hg38 will not suffice.
2. A file giving SNP/sequence variations locations (BED format); these locations refer to the given reference genome.

In more detail:

1. FASTA

A sequence file in FASTA format can contain several sequences. Each sequence in FASTA format begins with a single-line description, followed by lines of sequence data. The description line must begin with a greater-than (“>”) symbol in the first column. The word following the “>” symbol is the identifier of the sequence, and the rest of the line is the description (optional).

FASTA example:

>AB000263 |acc=AB000263|descr=Homo sapiens mRNA for prepro cortistatin like peptide, complete cds.|len=368

ACAAGATGCCATTGTCCCCGGCCTCCTGCTGCTGCTGCTCTCCGGGGCCACGGCCACCGCTGCCCTGCC

CCTGGxAGGGTGGCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGGCAGGAATAAGGAAAAGCAGC

CTCCTGACTTTCTCGCTTGGTGGTTTGAGTGGACCTCCCAGGCCAGTGCCGGGGCCCTCATAGGAGAGG

AAGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGCACCCCCCAGCAATCCGCGCGCCGGGACAGAATGCC

2. BED

1. **chrom** - The name of the chromosome (e.g. chr3, chrY, chr2_random) or scaffold (e.g. scaffold10671).
2. **chromStart** - The starting position of the **feature** in the chromosome or scaffold. The first base in a chromosome is numbered 0.
3. **chromEnd** - The ending position of the **feature** 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.

BED example:

field 1	field 2	field 3	optional field 4
chr6	20921570	20921571	example_SNP_1
chr12	65783448	65783449	example_SNP_2
chr2	40692483	40692484	example_SNP_3

* for further information about SeqSNP, please see SeqSNP guidance notes online: www.biosearchtech.com/services/sequencing/targeted-genotyping-by-sequencing-seqsnp

Integrated tools. Accelerated science.




 @LGCBiosearch | biosearch.com

All trademarks and registered trademarks mentioned herein are the property of their respective owners. All other trademarks and registered trademarks are the property of LGC and its subsidiaries. Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representative for details. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any retrieval system, without the written permission of the copyright holder. © LGC Limited, 2019. All rights reserved. GEN/0554/MW/0219

BIOSEARCH™
TECHNOLOGIES
GENOMIC ANALYSIS BY LGC