

## Documentation for the Sequence of GRP94 (HSP90B1) Gene Dataset

### Dr. Yair Argon, Health ABC Ancillary Study #AS07-109

#### General description and dataset structure

There were 91 HABCIDs that were sequenced for the HSP90B1 gene for Dr. Yair Argon's Health ABC Ancillary Study #AS07-109. The sequence data consists of 46 files per participant in FASTA format. There were 23 primers used to sequence the entire HSP90B1 gene. Sequence data is separated into one file per primer used and for the forward and reverse strand. All 3,796 FASTA data files are in a self-extracting zip file that can be downloaded from the "Current Datasets" listing on the Health ABC website under the "Datasets/ Documentation" link. This zip file is called:

SeqHSP90B1.exe

A map file is provided with the file names for each HABCID with sequence data. For example file name 01.A01.1161.p01-F.seq, 01.A01 is the plate number, 1161 is the HABCID number, p01 is the primer #1, and F is for the forward strand. Please note that some filenames do not have the plate number before the HABCID.

Additional documents included in the SeqHSP90B1 zip file are:

1. Graphical location of the HSP90B1 gene sequenced (SeqHSP90B1\_GraphicLocation.pdf).
2. File with list of primers used (SeqHSP90B1\_Primers.pdf).
3. Map file for filenames and HABCIDs with sequence data (SeqHSP90B1\_Map.pdf).

Key variables:

HABCID                      HABC Enrollment ID without the 2-letter prefix

The FASTA files each contain data for one HABCID, which is a unique identifier for each participant.

To use the data, please contact the PI at your site.

#### Methods

The entire HSP90B1 gene was spanned with 23 primers, using 50ng DNA per sample. VariantSeqr™ is a newly developed system for PCR sequencing of exons and 5'-upstream regions of human genes, consisting of 420,000 pre-designed PCR primer pairs for nearly 16,000 human genes. Sequencing by this method also included the presumed splice sites, to inform on variants that lead to alternative splicing or exon skipping.