**Workflow for the Distribution of cell lines and RNA samples**

The 500 European individuals that are sequenced to low coverage by the 1000 genomes project were sampled from five different populations (100 individuals / population): Utah residents with Northern and Western European ancestry (CEU), Toscani in Italy (TSI), British from England and Scotland (GBR), Finnish from Finland (FIN), Iberian populations in Spain (IBS). The CEU, TSI, GBR and FIN samples are now (April 7, 2011) available from the Coriell Institute for Medical Research. The IBS samples should be available in the next couple of weeks and will therefore be the samples that will be ordered/processed last.

Cell line Samples from the TSI, GBR, FIN and the IBS will be sent in batches of 25 from the Coriell Institute to ECACC London (25 samples / week). At ECACC the cells will be grown and RNA will be isolated. ECACC can process 50 cultures per two weeks for us. Then eight aliquots of the cell lines together with the isolated RNAs (1 aliquot per sample) will be sent to Geneva. We will further distribute the samples from here. The 100 CEU samples that we already have in our freezer will be sent to ECACC as well for culturing and RNA isolation – in order to have as comparable cell growth and RNA isolation conditions as possible across all analyzed samples. 32 of these samples were already sent to ECACC.

The senior management board of HPA/ECACC decided that a pilot with 32 samples had to be run and feasibility had to be demonstrated before they approve the processing of 500 samples. Thus, the 32 samples that were already sent will be used to test the protocols at ECACC for cell growth and subsequent RNA isolation. Two different protocols will be used for automated RNA isolation. The isolated RNA will be sent to Geneva. We will evaluate the RNA based on its integrity and quantity in Geneva. We will then decide by which protocol we will do the whole panel and whether the protocols used at ECACC will require adjustments.

The following table indicates the approximate time point of sample arrival at ECACC:

|  |  |  |  |
| --- | --- | --- | --- |
| **Cell line Batch** | **Expected Arrival at ECACC** | **Population** | **Destination lab** |
| 1 | End April | TSI | UNIGE (30 samples) MPIMG (20 samples) |
| 2 | Mid May | TSI | MPIMG (50 samples)  all 70 samples will be sent in one shipment |
| 3 | End May | FIN | UU (50 samples) |
| 4 | Mid June | FIN | LUNMC  (50 samples) |
| 5 | End June | GBR | HGMU (50 samples) |
| 6 | Mid July | GBR | ICMB (50 samples) |
| 7 | End July | CEU | UNIGE (50 samples) |
| 8 | Mid August | CEU | UNIGE (50 samples) |
| 9 | End August | ISB | CNAG/CRG/USC  (50 samples) |
| 10 | Mid September | ISB | CNAG/CRG/USC  (50 samples) |

The 500 cell cultures will be sent to all labs by mid September (when we have all here). I will send one copy (i.e. 500 cell lines) to each of the following labs: MPIMG, UU, LUNMC, HGMU, ICMB, CRG. Two cell line copies of the panel will stay in Geneva.

In Geneva we will also analyze a 10 percent overlap of the individuals that are sequenced in each of the other labs.

What you have to do:

1. Please provide me with contact information of a person in your institute that is responsible for receiving the samples (RNA / cell line samples) and put them in freezers accordingly (please provide name, e-mail address, and phone number). I won’t send of any dry ice parcel before I have confirmation that somebody will be physically there and take care of the parcel and its content.
2. Could you please send me a rough estimate of how long it takes for you to prepare the libraries and have them sequenced (+/- 1 month)? If there would be a long queue in one lab we may be able to give the samples analyzed in this lab some preference at ECACC. We should discuss this at our next phone call.

Protocols

All labs are expected to follow the protocols attached to this message. Each RNA sample will be prepped for the analysis of mRNA as well as small RNAs. Sequence from each sample at least 7 million clusters (try to aim for an average of 10 million clusters).

mRNAs: TruSeq protocol for High-Throughput analysis (starting page 71)

Perform 75 bp PE sequencing

Small RNAs: TruSeq protocol for small RNAs