Supplemental sheet for genotyping data

Template form for the record and the handing-over of accompanying documentary information when passing genome wide data packages to third parties

Author:
Dr. Thomas Bettecken
Center for Applied Genotyping - CAGT
Max-Planck-Institut für Psychiatrie
München, Germany

Approved by the project team “Quality Management for High-Throughput-Genotyping” (<http://genotypisierung.tmf-ev.de>)



Version: 16 Sep 2010



Provided by TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.

TMF Product No. P999091



**How to use this document**

To **fill out** this form electronically, it should be in protected mode. This document is provided in protected mode without password by default. In protected mode, only the placeholders are editable. You can hop from one placeholder to the next with your tab-key and fill out this form. Protection can be set by chosing “Protect Document” in the “Tools” or “Extras” menu.

In order to **make changes** to this form (cover page, additional items, delete items etc.) you need to unprotect the document (chose “Unprotect Document” in the “Tools” or “Extras” menu, by default no password is required). Don’t forget to set back to protected mode after editing is completed. We recommend not to set a password, because here protection is only needed to activate the form placeholders, not to prevent from unauthorized editing.

In **printouts**, the placeholders and the shading of text is not visible.

# DNA samples

## Name and address of shipping laboratory/group

e.g. “Neurological Clinic, Technical University of Munich / TUM, Ismaninger Str. 21, 81675 München“

|  |
| --- |
|       |

## Name/description of project

e.g. “Control\_MS – Genomewide Case/Control Association Study for Multiple Sclerosis”

|  |
| --- |
|       |

## Date when samples were handed over from one group to another

|  |
| --- |
|       |

## Sample history (Who? What? When?)

e.g. “DNA samples from Multiple Sclerosis patients were collected as blood in the region of Munich, over the time period Jan. 2000 to December 2007. Controls were collected from a population survey in the Augsburg region (KORA) in the time period July 1004 to June 2006”

|  |
| --- |
|       |

## Name/descriptor of list with sample IDs

e.g. “MS\_Control\_SampleIDs\_WGG\_08Dec15”

|  |
| --- |
|       |

## Volumes and (nominal) concentrations of samples

e.g. “Samples have a volume of nominally 20 µl each in a concentration of 50 ng/µl”

|  |
| --- |
|       |

## Statement as to how sample identifiers were augmented/changed in the genotyping laboratory. Name of new sample list

e.g. ”Sample IDs handed over have been augmented by a preceding “MS\_TUM\_” and are deposited in a sample list called MS\_Control\_SampleIDs\_WGG\_CAGT\_09Jan02”

|  |
| --- |
|       |

## Sample handling prior to genotyping

e.g. quality control, concentration measurement, concentration adjustment, agarose gel test, …

|  |
| --- |
|       |

## Gender information

e.g. “Gender information was included in the list with sample IDs, as M for males and W for females. These notations have been translated to ‘Male’ and ‘Female’ in the sample sheets used in genotyping.”

|  |
| --- |
|       |

## Replicates (Intentional duplicates? Twins? Reference/control samples? …)

|  |
| --- |
|       |

## Is the use of the samples for the genotyping covered by informed consent?

e.g. “To all samples there is written informed consent provided by every patient and control individual. The informed consent documents are archived with the group at the Clinic of Neurology at TUM.”

|  |
| --- |
|       |

## Restrictions on storage and handling (e.g. max. storage period, …) or requests of what is to be done with the samples after completion of genotyping?

e.g. “No restrictions known” or “Remaining samples are to be returned to the sending laboratory after completion of genotyping.”

|  |
| --- |
|       |

# Genotyping process

## Volumes, concentrations and amounts of DND used for genotyping

e.g. “Per sample: Volume 15 µl, concentration 50 ng/µl, amount 750 ng”

|  |
| --- |
|       |

## Time frame for the genotyping lab work

e.g. “Jan 22 to Mar 28, 2009”

|  |
| --- |
|       |

## Genotyping process used

e.g. “Illumina Infinium II”

|  |
| --- |
|       |

## Genotyping product used (chip type, version)

e.g. “Illumina Sentrix BeadChip Array HumanHap300 Genotyping BeadChip 317k, TagSNP Phase I, v1.0”

|  |
| --- |
|       |

## Instrument (hardware) used for scanning the chips

type of instrument, serial number, e.g. “Illumina BeadStation 500G, S157”

|  |
| --- |
|       |

## Software/version used for chip scanning

e.g. “Illumina BeadScan Version 2.0”

|  |
| --- |
|      |

## Interruption of the genotyping process

e.g. “After the initial amplification step, samples were stored in -20°C until used for further processing.”

|  |
| --- |
|       |

## Any other comments concerning the genotyping process

|  |
| --- |
|       |

# Information concerning the genotype calling

## Software/version used for genotype calling

e.g. “Illumina BeadStudio Version 3.5 with Module Genotyping Version 3.2.12”

|  |
| --- |
|       |

## Auxiliary files (e.g. cluster files) used for genotype calling

e.g. “BDCHP-1x10-HUMANHAP300v1-1\_11219278\_C.egt”

|  |
| --- |
|       |

## How was the genotype calling done (one by one, all together, in groups/batches, which)

|  |
| --- |
|       |

## Was there manual assignment of genotypes for certain SNPs? If yes, exact description, what was done, name of changed cluster file.

|  |
| --- |
|       |

## Were there SNPs seen, presenting problems (e.g. low callrate, unreliable clusters)?

|  |
| --- |
|       |

## Did the calling software give a quality score for calling the genotypes? If yes, which range does this score cover and which cutoff for the quality score was used?

e.g. “To each SNP, the BeadStudio software is giving a quality score, named GC score. A cutoff of 0.25 was used for the genotype calling. Genotypes with a score of less than this value should not be used”.

|  |
| --- |
|       |

## Where there SNPs excluded from genotype calling? Reasons?

|  |
| --- |
|       |

## Any other comments concerning the genotype calling process

|  |
| --- |
|       |

# Information concerning quality of genotyping

## Minimum callrate for the genotype data delivered

e.g. “99.0%”

|  |
| --- |
|       |

## Maximum callrate for the genotype data delivered

e.g. “99.99%”

|  |
| --- |
|       |

## What is the average call rate, summed over all samples? Graphics of the distribution of call rates (samples, SNPs)

|  |
| --- |
|       |

## What was done with samples underperforming?

e.g. “All samples with call rates < 99.0% were genotyped a second time”

|  |
| --- |
|       |

## Where there samples underperforming more than once? Reason?

e.g. “Samples were genotyped not more than twice. Call rate was >99.5% in the second round”

|  |
| --- |
|       |

## Where there doubles? If yes, what is the genotype concordance rate from samples genotyped twice?

|  |
| --- |
|       |

## How does the phenotypic gender, as given in the list of sample IDs, match the genotypic gender ? How many discrepancies?

e.g. “Out of 500 samples, in 499 the phenotypic gender as given in the list of sample IDs matched the genotypic gender. The ID of the non matching sample is MS\_TUM\_08976”

|  |
| --- |
|       |

# Information concerning genotype delivery

## SNP manifest used (list with assignments of company’s SNP identifiers with public SNP identifiers, e.g. rs-number)

e.g. “BDCHP-1x10-HUMANHAP300v1-1\_11219278\_C.bpm”

|  |
| --- |
|       |

## Coding of genotypes

e.g. “Coding of genotypes: A, C, G, T”

|  |
| --- |
|       |

## Orientation of genotypes, respectively definition, how genotypes were extracted

e.g. “SNP genotypes are given in the Forward orientation, as defined in Illuminas SNP manifest and produced by the BeadStudio software”

|  |
| --- |
|       |

# Information concerning raw data and genotype archiving

## Where and under which project and sample identifiers the samples’ raw data can be retrieved at the genotyping center

|  |
| --- |
|       |

## System, coding and example of Digital Object Identifiers (DOI)

|  |
| --- |
|       |

# Laboratory and persons responsible for genotyping, date of genotype delivery

## Name of laboratory and person(s) responsible for the genotyping

|  |
| --- |
|       |

## Date of delivery of genotype and accompanying information

|  |
| --- |
|       |

# Any other comments and information

|  |
| --- |
|       |