

## Experimental Design practical: Gene expression profile of wild and mutant HHEX in brain and liver development

**Hematopoietically Expressed Homeobox (HHEX) is a transcription factor which, plays an important role in the proper development of brain and liver in mouse. A mutant HHEX (where all Serine and Tyrosine residues are mutated to Alanine) is hyperactive and induces fetal death in mutant homozygous mice. We are interested in identifying gene expression changes in brain and liver and in determining key pathways involved, in response to hyperactivity of HHEX gene. Note that, samples are collected from 15 day-old fetuses that are homozygous wildtype (Wt/Wt), heterozygous mutant (Wt/Mt) and homozygous mutant (Mt/Mt), which manifest distinguishing morphological characteristics.**

Experimental design related question and answers:

- 1) What are the scientific questions of interest in this experiment?
  - a. Which genes show expression changes in response to mutant HHEX in brain and/or liver?
  - b. Do homozygous and heterozygous mutant HHEX mice show similar expression profiles?
  - c. Which genes show exclusive response in brain and liver respectively in response to hyperactivity of HHEX gene?
  - d. Which are the pathways that are involved in HHEX hyperactivity?
- 2) What are you measuring?
  - a. If you use RNA-seq you are measuring read counts.
  - b. If you use microarray you are measuring intensities.
- 3) What controls samples should be included in this experiment? Why is control needed in the experiment?
  - a. Wt/Wt mouse group is a control group that Wt/Mt and Mt/Mt groups will be compared against.
  - b. It is difficult, if not impossible, to measure the absolute gene expression changes. In differential expression experiments, one or more control group acts as a reference to compare the groups of interest to. Both Wt/Mt and Mt/Mt groups are compared against the control (Wt/Wt group).
- 4) How many replicates you need to include for each group? Discuss what factors might have influence in selecting the number of replicates?
  - a. Including replicates reduces variability in experimental results, increasing their significance.
  - b. There is no standard number of replicates recommended in an experiment. The number of replicates required, depends on various factors and therefore varies experiment to experiment. It is important to consider the number of replicates for each experiment that is carried out.
  - c. The absolute minimum number of replicates required is three per each experimental group. Statistical analysis cannot be carried if there are fewer than three replicates per group. The

Bioinformatics Core recommend a minimum of four replicates per group to allow for analysis to be carried out should a sample drop out of a group for whatever reason.

- d. In general a larger number of replicates should result in a more powerful experiment (i.e an experiment that is more likely to detect the effect size of interest should it exist in the population).
- e. Number of replicates required depends on the following factors
  - i. Amount of variation between experimental units. The more variability the harder it will be to detect the difference of interest therefore more replicates will be required.
  - ii. Effect size of interest (difference between control and treatment). To capture small effect size of interest, one needs to use more replicates. Therefore, it is important to know what size of effect you are interested in finding within the experiment.
  - iii. Statistical power required, in general the higher the number of replicates the higher the statistical power. 80% is usually taken as the amount of power to aim for in an experiment, that is there is an 80% chance of detecting a particular effect size given that it exists in the population. However, for the p-values to be repeatable 90% power is required.
  - iv. The Statistical significance level to be used in the statistical tests. This is usually taken to be 0.05 but can vary in different situations and should be pre-specified.
  - v. The type of alternative hypothesis (one sided or two sided alternate hypothesis)
  - vi. Other factors such as cost, ethics and ability to handle the number of replicates should also be considered.
  - vii. In our experiment one wants to see the effect of mutant HHEX on brain and liver development. Assuming that for each experimental group six biological replicates are required (note that both brain and liver samples can be collected from each fetus, this will give more power to any brain/liver comparisons). We need at least 18 mouse foetuses. Mouse average litter size is litter of 3-14; therefore samples may be collected from 2-6 different litters and litters should be from different male and female mouse pairs.

5) Which experimental groups will be included?

- a. There would be six experimental groups, 1) Brain from homozygous wild type mice, 2) Brain from homozygous mutant mice, 3) Brain from heterozygous mutant mice, 4) Liver from homozygous wild type mice 5) Liver from homozygous mutant mice and 6) Liver from heterozygous mutant mice.
- b. This is a 2 factorial experiment with three different levels.

	Wt/Wt	Wt/Mt	Mt/Mt
Brain	6	6	6

Liver	6	6	6
-------	---	---	---

Wt = Wild type

Mt = Mutant

- 6) How will any findings be validated?
  - a. Comparing your results with previous findings from literature
  - b. We could re-validate the findings using experimental techniques such as qPCR, in separate samples, such samples from completely different litters. Validating on samples from different litter gives generalizability to your findings.
- 7) What contrasts (sample group comparisons) you make with the data?
  - a. Brain Wt/Wt Vs Brain Wt/Mt for brain specific effect of heterozygous mutant
  - b. Brain Wt/Wt Vs Brain Mt/Mt for brain specific effect of homozygous mutant
  - c. Liver Wt/Wt Vs liver Wt/Mt for liver specific effect of heterozygous mutant
  - d. Liver Wt/Wt Vs liver Mt/Mt for liver specific effect of homozygous mutant
  - e. Brain Wt/Wt Vs Liver Wt/Wt, this comparison gives differentially expressed genes in brain and liver of wild type mouse.
  - f. Brain Wt/Mt Vs liver Wt/Mt, this comparison gives differentially expressed genes in brain and liver of mutant heterozygous mouse. Subtracting the, Brain Wt/Wt Vs Liver Wt/Wt, gene list gives brain and liver specific effect of heterozygous mutant genotype.
  - g. Like above comparison we can also find out the gives brain and liver specific effect of homozygous mutant genotype.
- 8) What are possible sources of bias and confounding variables in the experiment?
  - a. Gender of foetus
  - b. Age of the mice, animal breed, diet etc
  - c. People, if several people involved in sample collection and extraction.
  - d. Litter
  - e. Sequencing run
  - f. Index
  - g. Cage
  - h. Cage position
  - i. Technician handling the mice
  - j. Extraction batch
- 9) How can these sources of bias and confounding be controlled?
  - a. Randomisation
  - b. Blinding