Directionality of Mouse Phenotype Gene Sets

You will notice that in the mouse gene set file described here that the names of the individual gene sets often, but not always, have a directionality or qualifier included. Such as: *increased* lymph node number; *increased* osteoblast cell number; *decreased* circulating testosterone level, *decreased* diameter of femur, *abnormal* liver physiology; *enlarged* embryo size, etc.

This directionality is related to the phenotypic effects of the genetic mutations where they were **originally defined**. These are catalogued in the MGI informatics database and can be identified individually by searching with the gene or the Mammalian Phenotype code (e.x. MP:0008534) in the MGI database.

When these gene sets are used in the analysis of orthologous data such as microarray or SNP data that directionality gets confused.

If a particular gene set is enriched, it simply means that that gene set is ALTERED. Even though the title of the gene set may say *enlarged embryo size* , it really means that the genes that were originally defined by gene knockouts (or whatever mutant)and resulted in enlarged embryo size, were enriched or altered in your experiment. It does not necessarily mean that your phenotype involves enlarged embryo size.

The directionality of the change of a gene set may or may not be in the same direction as the gene set title suggests. Go back and look at the physical phenotype. It may be that the embryo is decreased in size. This is true for all gene sets with a directionality or qualifier in the title. Be careful.

KGB

August 10, 2010