

PCA of Pfam proportion in each sample



Principal component analysis was performed on the proportion of Pfam domains observed in assembled contigs in each sample, before any binning was performed. The horizontal axis shows the first axis of the decomposition, the vertical the second axis. Each axis is labelled with the percentage of variance which it explains. Polar stations are indicated with blue points, non-polar stations with red.

Genes were predicted from assembled contigs. The number of gene copies in the sample was estimated from the mean coverage of the contig when mapping reads back to the assembly. Each domain was assigned the sum of estimated gene copies in which that domain occurs.

This data is provided in comma separated format. Additionally, the same data mapped to GO terms is provided in the same format. We used the pfam2go mapping maintained by the Interpro team (<http://current.geneontology.org/ontology/external2go/pfam2go>)