Q1: What are those 4 candidate SNPs?

[HINT, you will may want to check the first few links of search result]

In association studies of 4 candidate SNPs (rs12936231, rs8067378, rs9303277, and rs7216389) in 3 family-based childhood asthma cohorts, all 4 SNPs demonstrated significant evidence for association

Q2: What three genes do these variants overlap or effect? ZPBP2

Q3: What is the location of rs8067378 and what are the different alleles for rs8067378?

[HINT, alleles and location are listed at the top of the the Ensemble page. You may

search in a genome browser to find this information]  
Chromosome **17:39895095**

Q4: What are the downstream genes for rs8067378? Any genes named ZPBP2,

GSDMB, and ORMDL3?

Q5: What proportion of the Mexican Ancestry in Los Angeles sample population (MXL) are homozygous for the asthma associated SNP (G|G)? [HINT: You can download a CVS file for this population from ENSEMBLE and use the R functions read.csv(), and table() to answer this question] 14%

Q6. Back on the ENSEMBLE page, search for the particular sample HG00109. This is a male from the GBR population group. What is the genotype for this sample? G|G

Q7: How many sequences are there in the first file? What is the file size and format of the data? Make sure the format is fastqsanger here!

3863

Q8: Q8: Does the first sequence have good quality?

Good

Q9: What is the GC content and sequence length of the second fastq file? [HINT, you may check “Basic Statistics”]

54%GC, 3863 sequences

Q10: How about per base sequence quality? Does any base have a mean quality score below 20?

Good. No base has mean quality below 20

Q11: Where are most the accepted hits located? [HINT, you can view the SAM version of your accepted hits file in galaxy and also use the UCSC Genome Browser via following the galaxy provided link and focusing on particular regions as described above]

IKZF3,GSDMB,ORMDL3,GSDMA,LRRC3C,OSMD3

Q12: Following Q13, is there any interesting gene around that area? [HINT, you can find genes around accepted hits in either the UCSC Genome Browser or IGV - depending on which browser you prefer]

IKZF3,GSDMB,ORMDL3,GSDMA,LRRC3C,OSMD3

Q13: Cufflinks again produces multiple output files that you can inspect from your righthand-side galaxy history. From the “gene expression” output, what is the FPKM for the ORMDL3 gene? What are the other genes with above zero FPKM values?

136853