Genotyping Part II

Dear all,

The genotyping has finally finished. I apologize for the delay, but after running the missiing portions, I noticed that the last 5 chromosomes did not have as many SNPs per bp as the first five. I wanted to make sure this wasn't an artifact of the initial failed analysis, so I reran the genotyping in a single run across the whole genome. The pattern is still there if looking at a restrictive call rate SNP set, but I think this is a characteristic of the genome/assembly. I'll write more on that later, but I wanted to get you the data as soon as I had it.

I'm still in the process of putting the INDEL calls up on the FTP, but all the SNP and raw data are there.

The login information is below:

Criteria	Credential
Server address	KITT.uri.edu
Port	2292
User	user1
Password:	...OyGen.2.0.1.8

Everything will be within the directory `/home/user1/files/`, and I **DO NOT** recommend trying to download the whole thing. It's 2.9 TB. You will likely be interested in the contents of `./VCF_files/` which contain the relevant variant files:

File Name	Variants	Number of Loci	Size
Combined.SNP.TRSdp5g75FnDNAmaf01.vcf.gz	Filtered SNPs with a minor allele frequency > 0.01	35,413,548	93 GB
Combined.SNP.TRSdp5g75FnDNAmaf05.vcf.gz	Filtered SNPs with a minor allele frequency > 0.05	16,752,436	52 GB
Combined.SNP.TRSdp5g75FnDNA.vcf.gz	Filtered SNPs	46,767,159	110 GB
	Filtered SNPs with a		

Combined.SNP.TRSdp5g95FnDNAmaf05.vcf.gz	minor allele frequency > 0.05 and missing data < 5%	9,559,587	31 GB
Combined.SNP.TRSdp5g1FnDNAmaf052alleles.vcf.gz	Filtered SNPs with a minor allele frequency > 0.05 and no missing data, 2 alleles	4,116,298	12 GB
Combined.TotalRawSNPs.vcf.gz	The original raw variant VCF	104,843,589	138 GB
Combined.SNP.TRSdp5g75mtDNA.recode.vcf.gz	Filtered SNPs from mitochondrial genome	1369	2.8 MB

Please let me know if you have any questions,

Jon