# Issue with different reference alleles between gVCFs

The issue happened during joined genotyping of gVCFs from NFE (set1/set2.g.vcf) and ampliseq.g.vcf file. Somehow, there were 3 positions, with different **reference** alleles between the old (NFE) and new (ampliseq) gVCFs. It was completely unexpected because both sets of gVCF files were generated using the same reference and the same versions of GATK.

## Used reference genomes

In the headers of EF’s old NFE gVCFs it could be seen that she used:

grep reference --color=auto set1.g.vcf

reference\_sequence=

/scratch/medgen/resources/gatk\_bundle/b37/decompressed/human\_g1k\_v37.fasta

This should have been equivalent to the current reference that I have used to generate ampliseq gVCF:

/rds/project/erf33/rds-erf33-medgen/resources/gatk\_bundle/b37/decompressed/human\_g1k\_v37.fasta

Moreover, the current reference file time stamp was of 2015, which preceded the NFE analysis done by EF.

Manual check of the discrepant position in the current reference showed that it is in agreement with ampliseq.g.vcf and in disagreement with the old NFE gVCFs:

grep -P "4\t57273790" --color=auto set2.g.vcf

$samtools faidx $ref\_genome 4:57273790-57273790

## Used GATK versions

EF’s logs showed that I used the same version of GATK to generate ampliseq.g.vcf, as she used for generating NFE gVCFs: gatk-3.6-0

## Check-ups and outcome

I verified that the recent ampliseq g.VCF contains reference alleles consistent with the reference genome fastq:

grep -P "4\t57273790" --color=auto set2.g.vcf

$samtools faidx $ref\_genome 4:57273790-57273790

Then I extracted all positions-refs-alts from the sourse gVCF files (set1/2.txt and ampliseq.txt; primary location: cluster) and explored them in R (primary location: Macbook).

There was ~300,000 overlapped positions in the files. Only 3 of them were discrepant. Thus, I decide that the reference genome versions were the same and I may just ignore these 3 positions. However, the cause for such discrepancy is entirely unclear.