SNP:Sites in genome where two indicidual(of the same specie!) differ by a single base. The different values this position can take is referred to as 'Allele'. The set of alleles a person has defines its genotype :

TAGC
TTGC

Alleles: A,T Genotypes: AT,TT,AA

The reason for the 2 bases appearing in genotype is that each person has duplicate chromosomes except the sex chromosome. So one chromoso with A allele and the duplicate one with T makes it AT. Genotype thus refers to the particular set of SNP an individual has at a particular site. About 10 million SNPs exist in human population the rarest of them being at 1% frequency. Alleles of SNPs that are close together will also have a higher chance of being inherited together[The explanation revolves around how chromosome division during cell reproduction takes place, think about it.] The set of associated SNP alleles in a region of chromosome is acalled the haplotype .

Resolution:

Person's genotype may not define its haplotype . Consider this case a diploid individual with SNPs at two loci on the same chromosome. The first one can be A/T and the second one G/C. Thus the genotypes are:

AA,AT,TT ;; GG,GC,CC

Now there are 9 possible configurations possible:

	AA	AT	TT
GG	AG AG	AG TG	TG TG
GC	AG AC	AG TC or AC TG	TG TC
CC	AC AC	AC TC	TC TC

With homozygous config at one or more loci resolution is unambigous, it does not matter if you read it as AG or GA. but the problem is with both heterozygous loci.

Another good example:

ACATGT
ACCGCT
GTCGGA

These three sequences can be uniques identified using the combination of rend and green tagged SNPs(Look closely , they are not the only ones!)

Why HapMap?;

By describing the common patterns of genetic varion among humans and the specific chromosoms regions where the SNP muttual association is strong, it essentially curtails the need to study 10 million SNPs. Not all SNPs relays the same level of information and hence it is possible to zero down upon .2 to 1 million SNP tags that relay the greater chunk of info. So while you are ddoing a controlled study study with a disease and non-diseased group and then focus only on regions where the groups differ in their haplotype frequencies.

I am not sure what will constitute the population here(disease+non-disease or simply disease for disease and non-disease for nono-disease)..

TODO: Check what is D and r^2 in SNP studies.

Genotype Frequency: Ratio of number of individuals of a given phenotype to the total population.

There is inherent difference between allele frequency and genotype frequency:

Let allele frequency $= f$

Let Genotype frequency be represented by G

Consider diploaid specied (some pink flower) with alleles $'Rw'$

30 flowers with RR
20 flowers with Rw
50 flowers with ww

$f(w) = \frac{Rw+2*(ww)}{2*RR+2*ww+2*Rw}[\frac{Number of alleles}{Total alleles}]$

On the other hand, the genotype freuqcy is :

$$g(Rw) = \frac{Rw}{Rw + RR + ww}[\frac{Genotype}{Total polpulation}] \tag{1}$$

Thus genotpyic frequency is an indicative of richment of population in terms of a particular genotype.In our case the $g(Rw)$ is $\frac{30}{100}$