Genotypes, Haplotypes, LD, and Selection

individual	site 1	site 2
1	{A,C}	{G,T}
2	{A,C}	{G,T}
3	{A,C}	{G,T}
4	{A,C}	{G,T}

(a) What is n for this dataset? What is π (average number of pairwise differences or average pairwise heterozygosity) for this dataset?

(b) The tables below show two different ways these genotypes could be sorted into haplotypes. Compute π again for the first table, do you get the same result?

haplotype	site 1	site 2
1	A	G
2	C	T
3	A	G
4	C	T
5	A	G
6	C	T
7	A	G
8	C	T

haplotype	site 1	site 2
1	A	G
2	C	T
3	A	T
4	C	G
5	A	G
6	C	T
7	A	T
8	C	G

(c) Compute LD (linkage disequilibrium) for both tables, is there a difference? Does this make sense?

(d) Compute LD for the following dataset.

haplotype	site 1	site 2
1	G	T
2	A	T
3	G	C
4	G	T
5	G	C
6	G	C
7	A	C
8	A	T