% Read desired accessions and positions from input xls file

desired\_accessions = xlsread('00\_Dataset\_Structure.xlsx','posADVRootNumber4DAC','B2:B121');

desired\_positions = xlsread('00\_Dataset\_Structure.xlsx','posADVRootNumber4DAC','C2:C205979');

% Read hdf5 file

snp = hdf5read('all\_chromosomes\_binary\_gzip.hdf5','snps');

all\_positions = hdf5read('all\_chromosomes\_binary\_gzip.hdf5','positions');

all\_accessions\_struct = hdf5read('all\_chromosomes\_binary\_gzip.hdf5','accessions');

all\_accessions = zeros(all\_accessions\_struct.length,1);

for i=1:all\_accessions\_struct.length

all\_accessions(i,1) = str2num(all\_accessions\_struct(i).Data);

end

number\_of\_desired\_accessions = length(desired\_accessions);

number\_of\_desired\_positions = length(desired\_positions);

result = zeros(number\_of\_desired\_accessions,number\_of\_desired\_positions);

tic

stored\_positions = zeros(number\_of\_desired\_positions,1);

for i=1:number\_of\_desired\_accessions

disp(100\*(i-1)/number\_of\_desired\_accessions);

current\_desired\_accession = desired\_accessions(i);

row\_of\_desired\_accession = find(all\_accessions == current\_desired\_accession);

k = 1;

for j=1:length(all\_positions)

current\_position = all\_positions(j);

is\_desired\_position = find(desired\_positions == current\_position);

if isempty(is\_desired\_position)

% Do nothing

else

result(i,k) = snp(row\_of\_desired\_accession,j);

stored\_positions(k,1) = current\_position;

k = k + 1;

end

end

end

elapsed\_time = toc