Opdracht 1:

1. General n-ary tree
2. Door als root node Buitenland, en met ja/nee edges naar subnodes van Europa en Nederland, van Nederland naar plaatsnamen, van Europa met ja/nee naar Australie en Europese landen, naar plaatsnamen etc tot locaties.
3. Nee, per node zijn er meer dan twee subnodes.

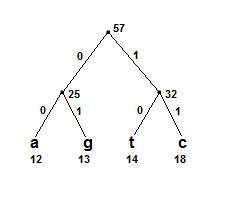
Opdracht 2:

1. a = 0,21

g = 0,22

t = 0,24

c = 0,31

1. 
2. a = 00

g = 01

t = 10

c = 11

1. 1000100111110000100111001010101010011111110111001001000001110010010001110111011100100111111000111110011100110111
2. Nee.
3. Door combinaties van 3 letters, dus effectief codons, als uitgangpunt te nemen ipv nucleotiden.
4. Ja, coderende DNA sequenties hebben meestal een andere codon-compositie dan niet coderend DNA.

Opdracht 3:

1. Door alleen de verschillen tussen een genoomsequentie en een referentiesequentie te coderen, waardoor maar een heel klein deel van de genoomsequentie nodig is.
2. Door in plaats van de Cambridge Reference Sequence een consensus sequentie voor de te coderen genoomsequenties te gebruiken kan er nog 23% extra compressie behaald worden.
3. Ze maken gebruik van een lossless compressie, er gaat geen data verloren.
4. Nee, je hebt te maken met een grote encoded string, en er is een decoderingstree op voorhand beschikbaar.