



Alignment ved hjælp af dotplot-metoden

Denne vejledning indeholder en række små øvelser og opgaver der illustrerer, hvordan dotplot-metoden kan benyttes til at undersøge parvise sekvens-alignments. På nettet findes en række mere eller mindre brugervenlige programmer til at lave dotplots af lange sekvens-sammenligninger, her vil vi primært fokusere på den generelle metode og arbejde med små eksempler der kan løses både i hånden og ved hjælp af det tilhørende Excelark.

Excelarket DotplotMetoden.xls, kan downloades på:

www.bioteknologibogen.dk/bioteknologi-6/data/DotPlotMetoden.xlsx

Dotplottets primære fordel er at det kan hjælpe med at identificere små lokale alignments i en sammenligning af to sekvenser, hvor de to sekvenser har en høj ensartethed. Ud fra identificeringen af disse lokale alignments er det muligt at identificere en lang række fænomener der er svære at håndtere med andre globale alignment-metoder. Af muligheder kan blandt andet nævnes:

- Indelmutationer
- Translokationer
- Inversioner
- Exon-/intron-struktur i genomsekvens for et gen ved sammenligning med mRNA
- Repetitive elementer

Derudover kan man selvfølgelig også identificere simple punktmutationer. De forskellige typer mutationer er alle beskrevet i Bioteknologi 6, tema 12, siderne 13-18. Herunder kan du finde små øvelser der illustrerer dotplot-metodens evner til at identificere nogle af ovenstående fænomener.

Introduktion til dotplot-metoden

Denne øvelse er en introduktion til brug af dotplot-metoden. De enkelte afsnit indeholder både en vejledning og små opgaver der skal løses undervejs. For en grundigere gennemgang af teorien se også forklaring i Bioteknologi 6 på siderne 15-16.

I dette eksempel skal du lave en dotplot-alignment af de første 10 aminosyrer af glucose-6-fosfat-dehydrogenase-genet (G6PD) fra mennesket (*Homo sapiens*) og ulven (*Canis lupus*).

1. Start med at opskrive de to sekvenser i de grå felter og udfyld dernæst skemaet ved at sætte en prik (dot) i alle felter hvor de to sekvenser er ens.

Sekvens 1 (Menneske): MGRRGSAPGN Sekvens 2 (Ulv): MSRRGAAPGN

		Menneske									
Ulv											

2. Download Excelarket DotplotMetoden.xlsx, beskrevet i introduktionen.
3. Indtast de to sekvenser i Excelarket som i eksemplet vist herunder, og tjek at du får samme resultat som i din håndlavede udgave.

		Testsekvens 1									
		M	A	M	T	C	G	N	K	L	W
Testsekvens 2	M	.		.							
	A		.								
	G				.						
	T				.						
	C					.					
	G						.				
	N							.			
	K								.		
	L									.	
	W										.

Eksempel på hvordan man kan bruge Excelarket til hurtigt at få lavet et dotplot ved at sammenligne sekvenser på op til 20 aminosyrer.

4. Hvor mange punktforskelle er der mellem de to sekvenser i den bedst mulige alignment?
5. Vurder om de to sekvenser er homologe.

Indelmutationer

I dette eksempel skal du lave en dotplot-alignment af udvalgte nucleotider fra genet for ficolin-3 fra en rask og syg person.

Allel fra rask person GAGGGCAGGGCCCTCCCAGTCTTT
Allel fra syg person GAGGGCAGGGCCTCCCAGTCTTT

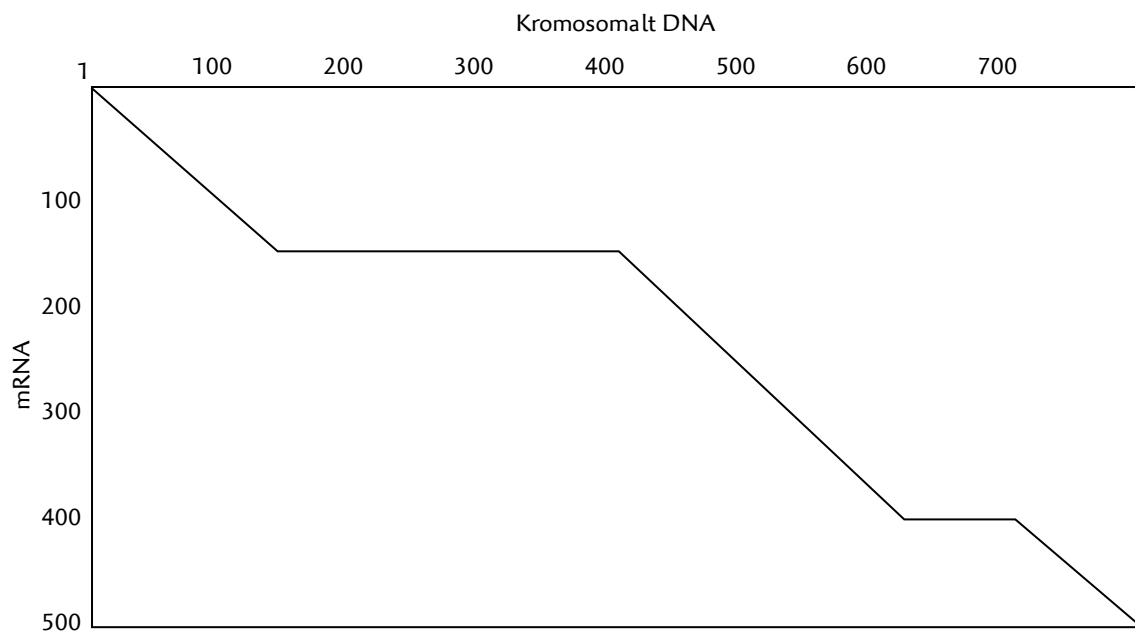
Data er fra en tidligere eksamensopgave i Biologi A. Opgave 1 fra sættet den 25. maj 2010.

1. Lav et dotplot af de to sekvenser, indtegn de to længste diagonaler og forbind dem.
2. Forklar ud fra dit dotplot hvilken type mutation der er tale om, og hvor i sekvensen denne mutation mest sandsynligt er opstået hos den syge person.

Exon-intron-struktur

Dotplots kan bruges til at sammenligne mRNA med kromosomt DNA og på den måde finde exons og introns i kromosomt DNA. I mange tilfælde er sekvenserne man ønsker at sammenligne, så lange at det er urealistisk at håndtere dem i hånden. Til dette formål findes der computerprogrammer der kan sammenligne sekvenser af vilkårlig længde og udskrive dotplots.

I denne opgave skal du ikke selv lave et dotplot, men i stedet analysere og fortolke et dotplot. For at gøre dotplottet lettere at fortolke, er det renset for tilfældige ligheder så det der vises, kun er den dominerende sti igennem dotplottet.



1. Marker exon-intron-strukturen på dotplottet og angiv hvor mange exons der er i det undersøgte gen.
2. Hvad er den totale længde af henholdsvis exons og introns i det undersøgte gen ca.?