

Propädeutikum: Programmierung in der Bioinformatik Dotplot

Thomas Mauermeier

10.12.2019

Ludwig-Maximilians-Universität München

	A	G	G	T	C	T	A	A	C
G		•	•						
G		•	•						
T				•		•			
A	•						•	•	
A	•						•	•	
C					•				•
A	•						•	•	
G		•	•						
G		•	•						

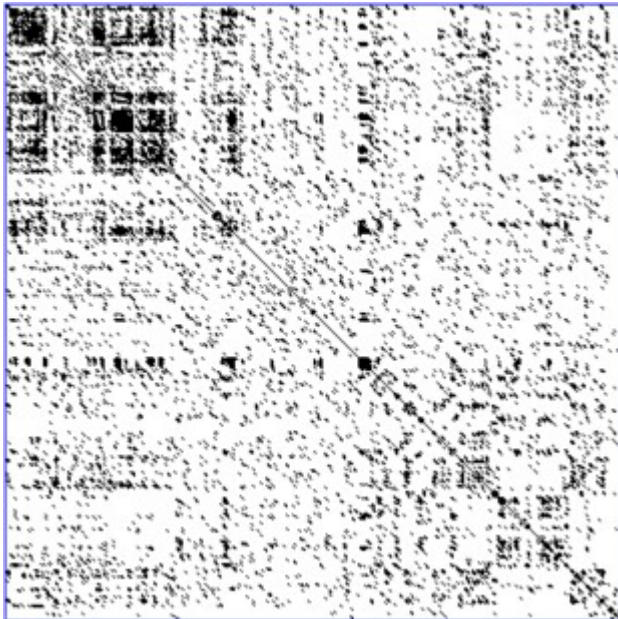


Plot
+
Dot

= Dotplot

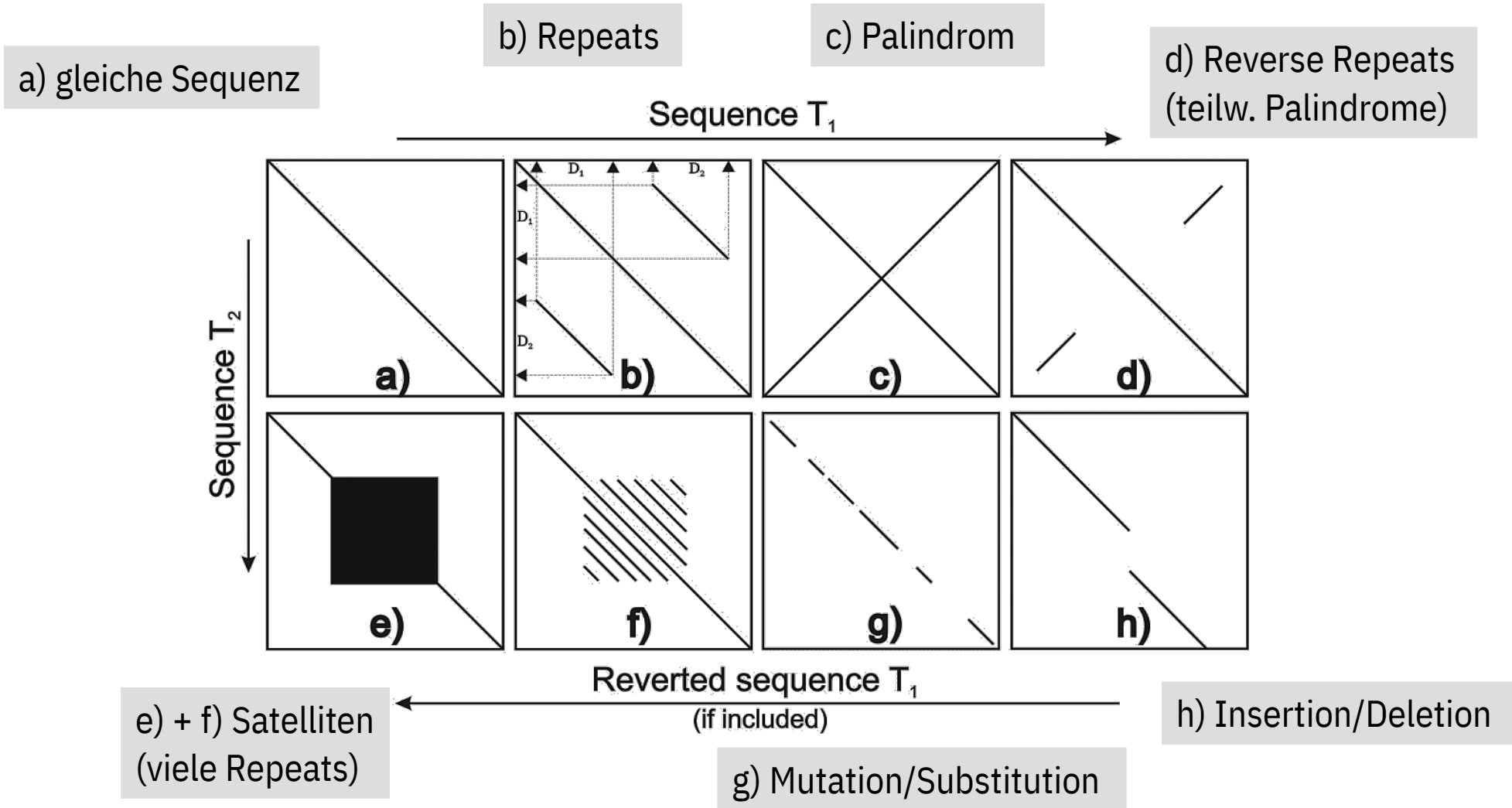


- Grafische Darstellung der Ähnlichkeit von Sequenzen
- Gleiche Sequenzabschnitte als Diagonale im Dotplot



<https://en.wikipedia.org/wiki/File:Zinc-finger-dot-plot.png>

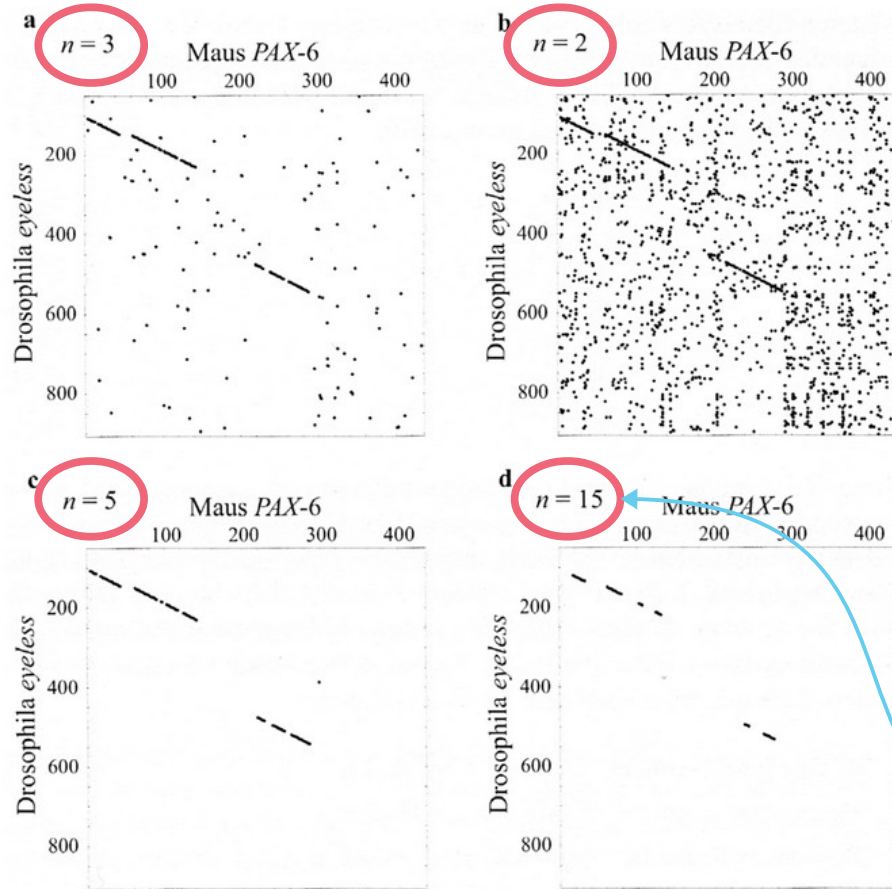
	A	G	G	T	C	T	A	A	C
G		•	•						
G		•	•						
T				•		•			
A	•						•	•	
A	•						•	•	
C					•				•
A	•						•	•	
G		•	•						
G		•	•						



Eure Aufgabe: Dotplot programmieren

- Unsere Vorgaben an euch:
 - **Input:** 2 Sequenzen (Aminosäuren oder Nukleotide – beides möglich)
 - **Output:** Dotplot auf dem Terminal ausgeben
- Ansonsten liegt die Implementierung bei euch
- Wer anspruchsvolleres möchte: Version die “fensterweise” vergleicht (folgende Folien)

Warum fensterweise Vergleichen?



“Herausfiltern” der Noise

n = Wort- bzw. Fenstergröße

A C C T T G T C C T C T T T G C C C

A																		
C																		
G																		
T																		
T																		
G																		
A																		
C																		
C																		
T																		
G																		
T																		
A																		
A																		
C																		
C																		
T																		
C																		

window = 9
step = 3
score = 4

	A	C	C	T	T	G	T	C	C	T	C	T	T	T	G	C	C	C
A																		
C																		
G																		
T																		
T																		
G																		
A																		
C																		
C																		
T																		
G																		
T																		
A																		
A																		
C																		
C																		
T																		
C																		

Vergleichen und Score berechnen: 7 Matches
 → 7 > score; also wird Punkt gesetzt!

window = 9
 step = 3
 score = 4

	A	C	C	T	T	G	T	C	C	T	C	T	T	T	G	C	C	C
A																		
C																		
G																		
T																		
T																		
G																		
A																		
C																		
C																		
T																		
G																		
T																		
A																		
A																		
C																		
C																		
T																		
C																		

Fenster um step-viel weiterschieben (also 3);
 Vergleich wiederholt sich, etc.

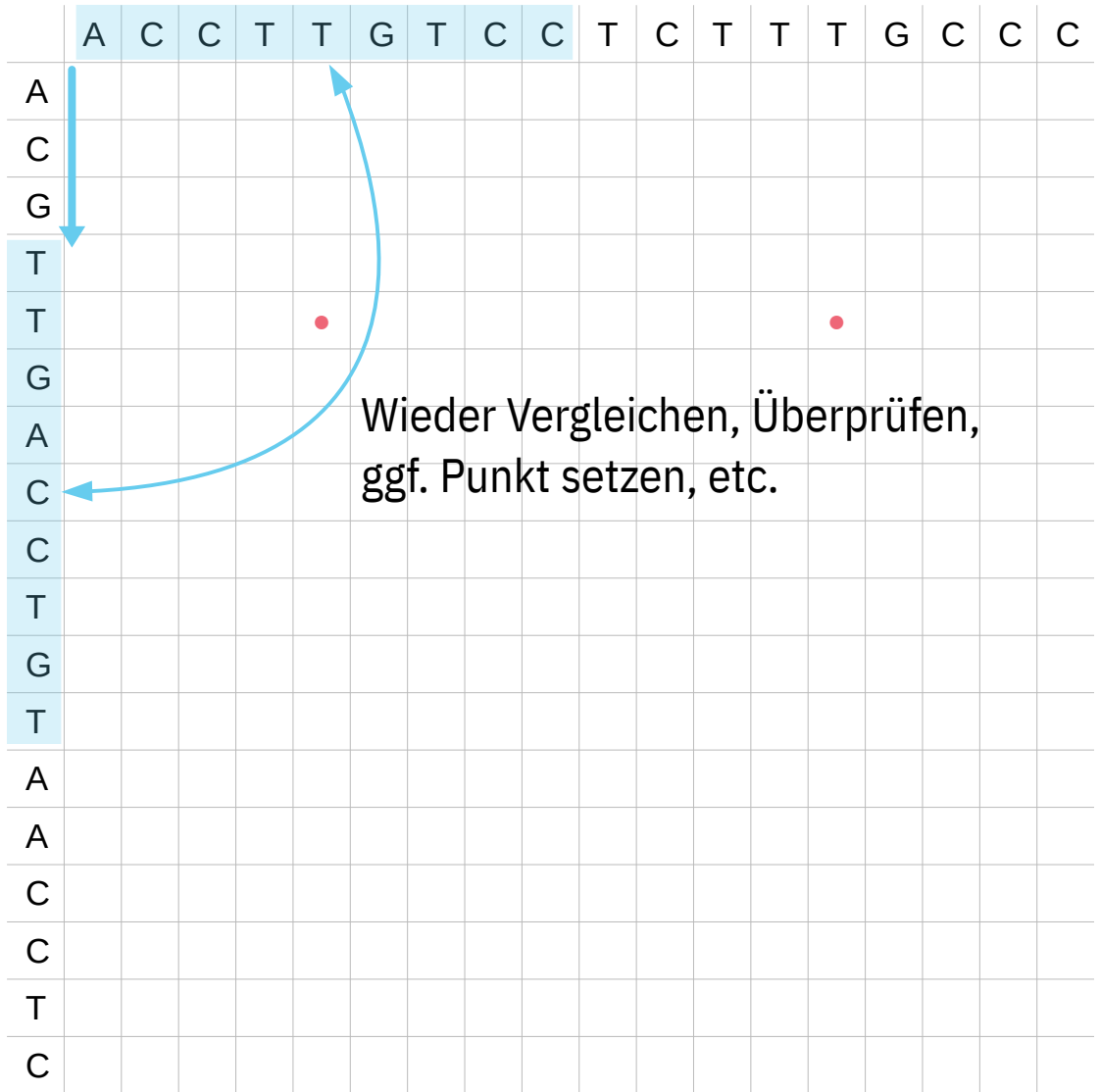


window = 9
 step = 3
 score = 4

	A	C	C	T	T	G	T	C	C	T	C	T	T	T	G	C	C	C
A																		
C																		
G																		
T																		
T																		
G																		
A																		
C																		
C																		
T																		
G																		
T																		
A																		
A																		
C																		
C																		
T																		
C																		

Am Ende?
 Fenster auf anderer
 Sequenz verschieben

window = 9
 step = 3
 score = 4



window = 9
step = 3
score = 4