

# Entwurf eines DNA- Algorithmus zur Lösung des Maximum Clique Problems

Sarah Carolin Kempf  
Friedrich-Schiller-Universität Jena  
Molekulare Algorithmen SS 2016

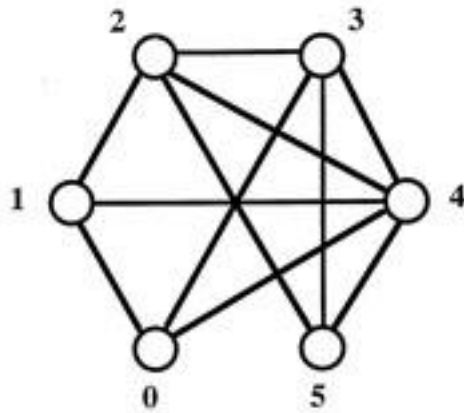
# Inhalt

- Einleitung
- Idee
- Algorithmus
- Ouyang et al.
- Bewertung des Algorithmus

# EINLEITUNG

# Einleitung

- Graph G:
  - ungerichtet



- Clique: Knoten in einem Graph die alle untereinander verbunden sind
- Maximum Clique: Größte Clique im Graph

Bildquelle: [1]

# Einleitung

- NP vollständiges Problem
- Nutzen in der Biologie
  - Genexpressionsnetzwerke (Regulatorische Systemeoft Cliques)
  - Beziehungen zwischen Chemikalien durch Cliques zwischen den Substanzen zeigen (Pharmakologie)
  
- Frage: Wie viele Knoten sind in der größten Clique enthalten?

IDEE

# Idee

- Knoten: DNA-Einzelstrang mit 2 Bausteinen

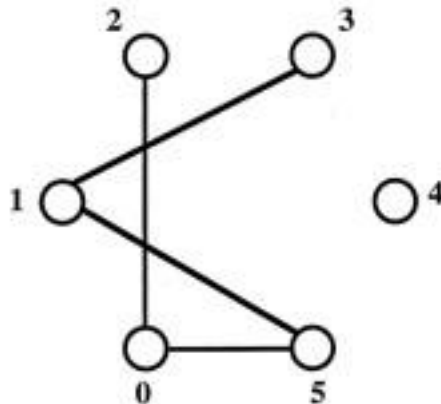
Knoten	Bausteine
0	ATTA
1	GCCG
2	AGGA

- Kante: DNA Einzelstrang mit 4 Bausteinen

Knoten	Bausteine
1-2	GCTC
1-2-4	GCTCGG
2-5	CTTT

# Idee

- Alle möglichen Kombinationen von Cliques erzeugen
- Falsche Cliques herausziehen
- Übrig gebliebene Cliques sortieren
- Falsche Cliques:
  - Cliques die in DNA zusammenfinden aber im Graphen nicht existieren



Bildquelle: [1]



# ALGORITHMUS

# Algorithmus

- Schritt 1
  - Erzeuge alle möglichen Cliquen

- Annealing

- Doppelstränge erzeugen



- Ligation

- Verkettung der Stränge



Bildquelle: [3]

# Algorithmus

- Schritt 2
  - Subtrahiere falsche Cliquen

- Melting



- Biotin-Avidin-Purification:

- DNA-Einzelstrang „0-2“ biotinylieren
- Anlagern komplementärer DNA-Einzelstränge an DNA-Einzelstränge, die mit Biotin und Avidin fixiert sind.

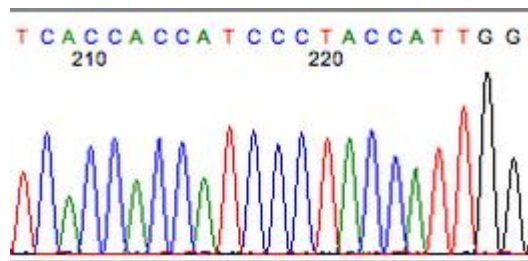
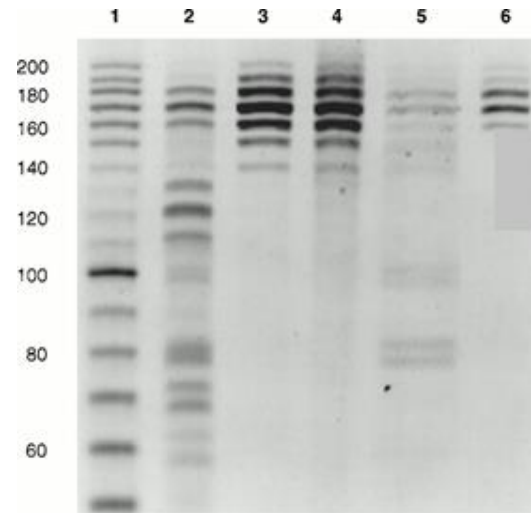


- Ausgewaschene Stränge in Reagenzglas sammeln
- Annealing
- für alle falschen Cliquen wiederholen

Bildquelle: [3]

# Algorithmus

- Schritt 3
  - Nach Länge sortieren und Auslesen
- Agarose Gel Elektrophorese
  - Längenbestimmung
- Sequencing
  - Bestimmung der Clique



Bildquellen: [1], [2]

OUYANG ET AL.

# Ouyang et al.

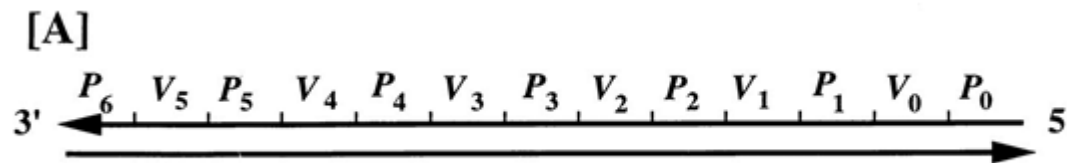
- 1997 Lösung für das Maximum Clique Problems
- DNA Computer
- veröffentlicht in Science
- Findet maximale Clique bei bis zu 27 Knoten

# Ouyang et al.

- 4 Schritte:
  - 1) Jeder Graph mit  $N$  Knoten wird durch eine Binärzahl mit  $n$ -Ziffern dargestellt.  
1 = in der Clique  
2 = nicht in der Clique
  - 2) Verbindungen aus dem komplementären Graph werden aus dem Originalgraphen entfernt
  - 3) Aus den Daten werden alle Zahlen entfernt, die eine Verbindung im komplementären Graphen haben
  - 4) Daten werden nach Größe sortiert.

# Ouyang et al.

- Jeder Knoten Wert  $V_i$  und Position  $P_i$
- Eine Clique:  $P_0V_0 P_1V_1 P_2V_2 P_3V_3 P_4V_4P_5V_5 P_6$
- Aufbau durch 12 Oligonukleotide
- $V_i = 0$  bp, wenn  $V_i = 1$  und 10 bp, wenn  $V_i = 0$ .



Bildquelle: [1]



# Ouyang et al.

- Entfernung der illegalen Cliques durch Restriktionsenzyme
- Ergebnis durch Agarose Gel Elektrophorese
  - Unterstes Band = kürzeste DNA= Größte Clique
- Genaue Knoten in der Clique
  - Durch Klonierung der DANN
  - Bacteriophage -> E. coli -> Klonierung -> Sequenzierung

```

6400  ACTCACTATAGGGC GAATTGGGTA CCCTGGATCCCGCCCCTCTC
                               tail                               P6

          GAGTCTATTTACACGTAGG CCTTTATCTCAAANACTGCA
                P5                               P4

          TGCGCAGGCTTGTTAGCACT AGTAATCTCTCTTCACGAAG
                P3                               P2

          CGGTTATGAA CTTACTGATAGGTGACTCTT CTGCCGCGTC
                V10                               P1                               V00

          AAGGTCGCANAATTCTACG TGGAGCTCCAGC TTTTGTTC      6260
                P0                               tail
    
```

# BEWERTUNG DES ALGORITHMUS

# Bewertung des Algorithmus

- Viele Knoten theoretisch umsetzbar
  - Zeitverlust durch Ansetzen der Bausteine
- Massiv parallelisierbar

# Vergleich Ouyang et al.

- Ouyang et al. braucht weniger Vorbereitung
  - Bsp.: 12 Oligonukleotide  $\leftrightarrow$  20 Bausteine
- Ouyang et al. durch Restriktionsenzyme anfälliger für Fehler
- Beide:
- je größer die Menge der Untersuchung desto mehr Fehler durch biochemische Reaktionen
- Biotin-Avidin Purification sehr langsam
- Elektrophorese unruhig

Vielen Dank für Ihre  
Aufmerksamkeit!

# Quellen

- [1] „DNA Solution of the Maximal Clique Problem“ Qi Ouyang et al., Science 17 Oct 1997: Vol. 278, Issue 5337, pp. 446-449 DOI: 10.1126/science.278.5337.446
- [2] Sequence Analysis of your Addgene Plasmid URL: <https://www.addgene.org/recipient-instructions/sequence-analysis/>
- [3] Vorlesungsskript Dr. Thomas Hinze