



BIOINFORMATYKA

WPROWADZENIE

PIOTR WAWRZYNIAK

INSTYTUT INFORMATYKI, POLITECHNIKA POZNAŃSKA

PIOTR.WAWRZYNIAK@CS.PUT.POZNAN.PL

WWW.PIOTR.E.WAWRZYNIAK.DOCTORATE.PUT.POZNAN.PL

ORGANIZACJA ZAJĘĆ

Wprowadzenie

Definicja problemu

Algorytm dokładny

Algorytm przybliżony (wielomianowy)

ZALICZENIE

Grupy 2 osobowe

Sprawozdanie na ostatnich zajęciach

Sprawozdanie w PDF-ie plus kod

Oprogramowanie w trybie wsadowym

Termin oddawania sprawozdania – ostatnie zajęcia

KWAS DEOKSYRYBONUKLEINOWY

Puryny

Adenina – Ade – A
Guanina – Gua – G

Pirymidyny

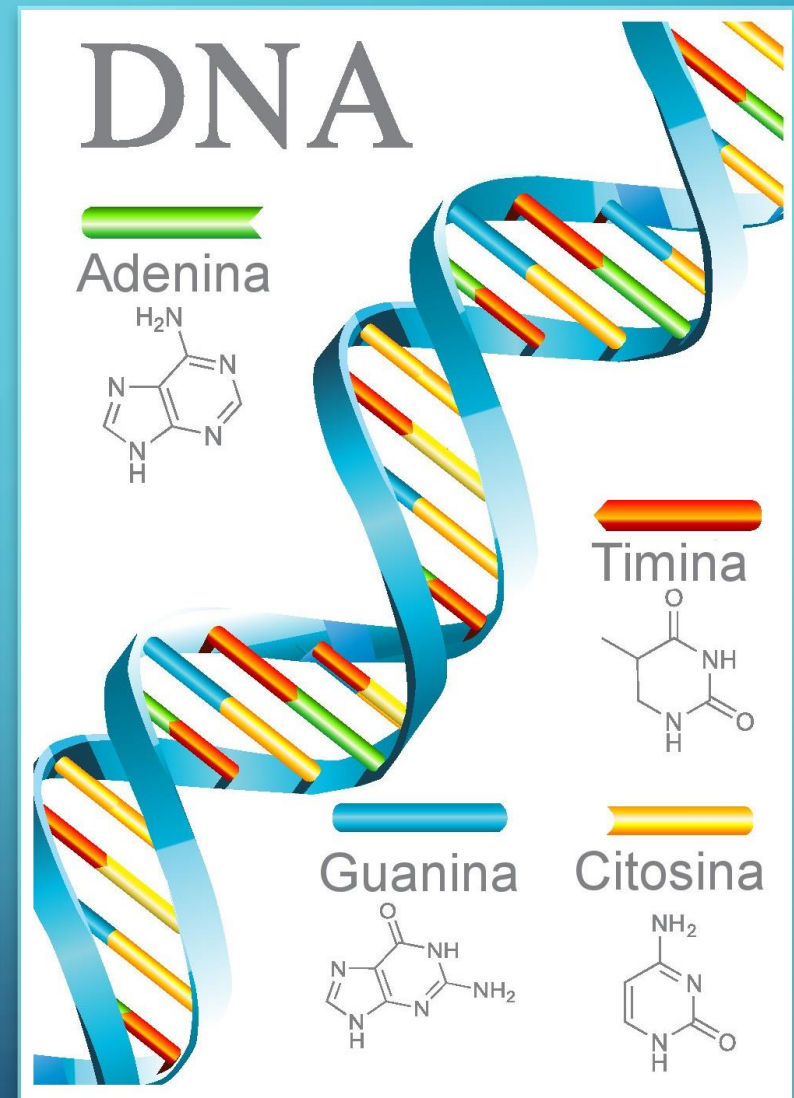
Cytozyna – Cyt – C
Tymina – Thy – T

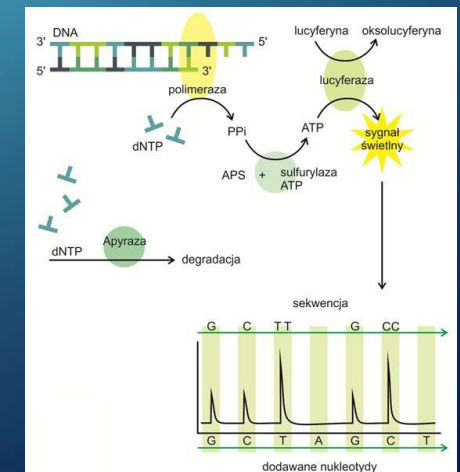
Słabe

Adenina – Ade – A
Tymina – Thy – T

Silne

Guanina – Gua – G
Cytozyna – Cyt – C





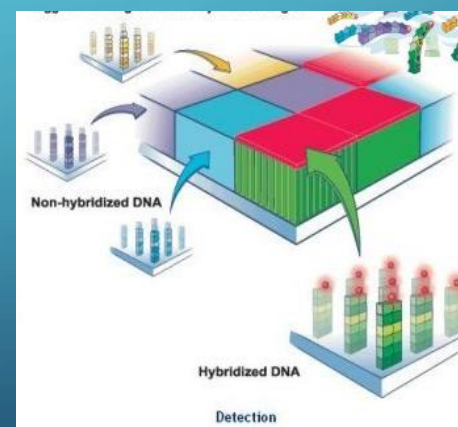
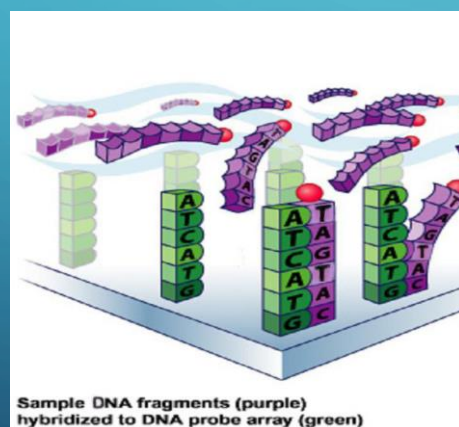
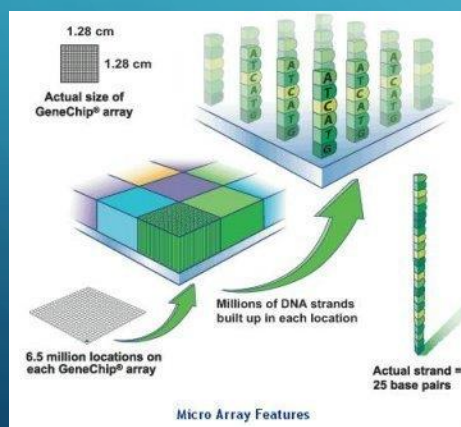
SEKWENCJONOWANIE PRZEZ HYBRYDYZACJĘ

Microchip DNA z
naniesionymi
oligonukleotydami

Hybrydyzacji
badanej cząsteczki
do poszczególnych
oligonukleotydów

Rozpoznanie
sparowanych
oligonukleotydów

Analiza wykrytych
sekwencji
oligonukleotydowych



BŁĘDY

Pozytywne

Hybrydyzacja błędnych
(niekomplementarnych)
oligonukleotydów

Błędne rozpoznanie
obrazu mikrochipu

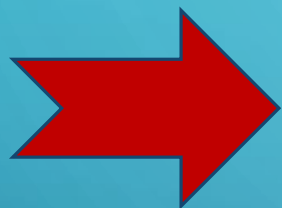
Negatywne

Błędne rozpoznanie
obrazu mikrochipu – za
słaby sygnał

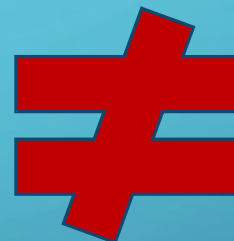
Występowanie
rozpoznanej sekwencji
więcej niż raz

POPRAWNOŚĆ ROZWIĄZANIA

1 ACG
2 CGA
3 CGT
4 GAC
5 GTA
6 TAC
7 TCG



7 TCG
3 CGT
5 GTA
6 TAC
1 ACG
2 CGA
4 GAC
TCGTACGAC



7 TCG
2 CGA
4 GAC
1 ACG
3 CGT
5 GTA
6 TAC
TCGACGTAC

SEKWENCJONOWANIA PRZEZ HYBRYDYZACJĘ PODEJŚCIA NIEKLASYCZNE

PSBH – z dodatkową informacją o położeniu

z dodatkową informacją o ilości powtórzeń

z uniwersalnymi nukleotydami

- gapped chip – NNNN + NNNXXXN
- alternating chip – NXNXNXN + NXNXNN

ze zdegenerowanymi nukleotydami

- binary chip - ZZZN + PPPN
Z = {W, S} – słabe W = {A, T} / silne S = {C, G}
P = {R, Y} – puryny R = {A, G} / pirymidyny Y = {C, T}

WYBÓR PROBLEMÓW

SBH z informacją o położeniu (PSBH)

- dodatkowa informacja p – położenie oligonukleotydu w części sekwencji
- błędy: pozytywne lub negatywne i w informacji o położeniu

SBH z informacją o powtórzeniach

- dodatkowa informacja r – o liczbie powtórzeń oligonukleotydu w sekwencji
- błędy: pozytywne lub negatywne, dotyczące informacji o powtórzeniach

Uniwersalne nukleotydy

- struktura sond oligonukleotydowych
- błędy: pozytywne lub negatywne

Zdegenerowane nukleotydy (binary chip)

- struktura sond oligonukleotydowych
- błędy: pozytywne lub negatywne

DANE WEJŚCIOWE

- n – długość badanej sekwencji
- k – długość oligonukleotydu chipu
- Początkowa sekwencja długości k
- Spektrum
 - struktura oligonukleotydu, np.: CATGGGCC, ACANNNC
 - ilość powtórzeń
 - położenie w sekwencji