

# BioinfoRmatyka nowotworów

Ewa Szczurek

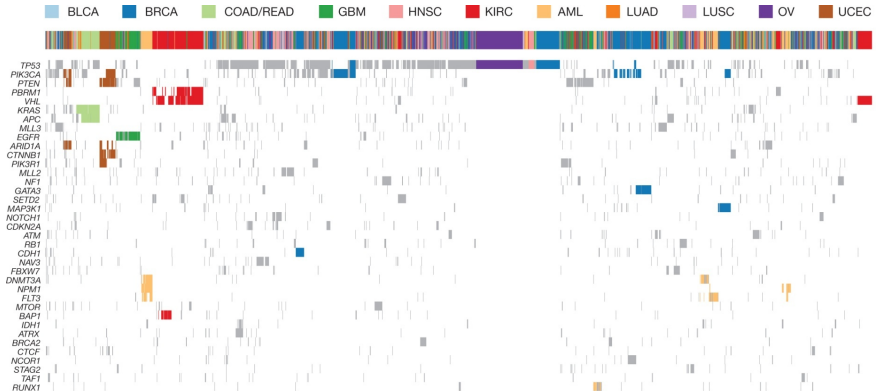
Universytet Warszawski

28 września 2017

# Rak to choroba genomu

- ❶ Dane o mutacjach w genomach komórek nowotworowych
  - Genomy komórek rakowych noszą liczne mutacje genów
  - Rak "wyłącza" bądź "aktywuje" geny
  - Komórki tracą/zyskują nową funkcjonalność
- ❷ Dane z experimentów perturbacyjnych RNA interference (RNAi)
  - pozwalają naukowcom "wyłączyć" geny w liniach komórkowych w laboratorium

# Dzũngła danych: mutacje genów w genomach rakowych



C Kandoth *et al. Nature*, (2013)

# Jak znaleźć drogę w tej dżungli danych?

- ❶ W jaki sposób wyłączanie genów zmienia funkcjonalność szlaków sygnalizacyjnych w raku?
  - Mutacje genów z tego samego szlaku wykluczają się wzajemnie.<sup>1,2</sup>
  - Modelowanie szlaków sygnalizacyjnych z danych perturbacji RNAi<sup>3</sup>
- ❷ Które mutacje możemy wykorzystać w terapii raka?
  - Syntetycznie letalni partnerzy genów zmutowanych<sup>4,5</sup>

---

<sup>1</sup> Szczurek & Beerenwinkel, *PLoS Comput. Biol.*, 2014

<sup>2</sup> Constantinescu et al. *Bioinformatics*, 2015

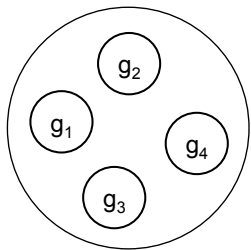
<sup>3</sup> Szczurek & Beerenwinkel, *Bioinformatics*, 2016

<sup>4</sup> Szczurek et al. *Int J Cancer*, 2013

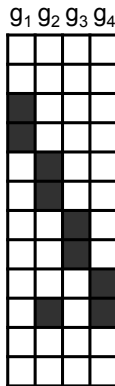
<sup>5</sup> Matlak & Szczurek, *PLoS Comput. Biol.*, 2017

# Wzorzec wzajemnego wykluczania się mutacji w raku a szlaki sygnałacyjne

**Cancer pathway**

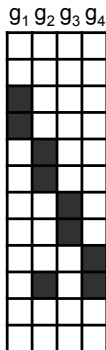


**Mutually exclusive pattern**

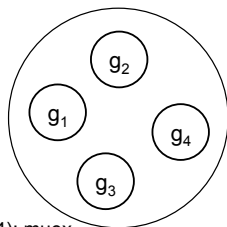


# Wzorzec wzajemnego wykluczania się mutacji w raku a szlaki sygnalizacyjne

## Mutually exclusive pattern



## Cancer pathway hypothesis



Vandin *et al.*, (2014)

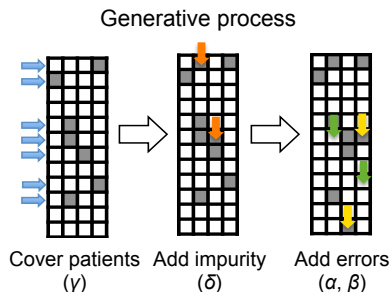
Kim *et al.*, (2015)

Szczurek & Beerenwinkel (2014): muex

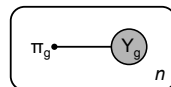
Constantinescu *et al.* (2015): TIMEx

Cristea *et al.* (2016)

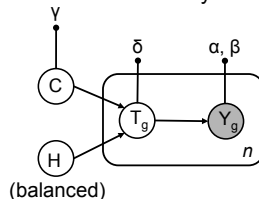
# Test ilorazu wiarygodności wykluczania się genów.



Null: Independence model

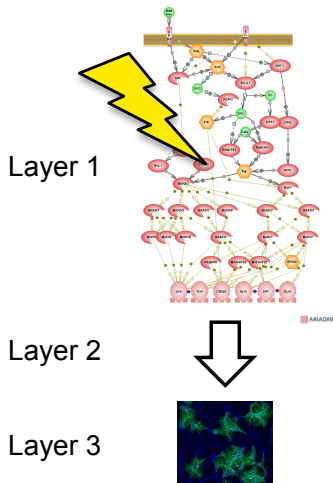


Mutual exclusivity model



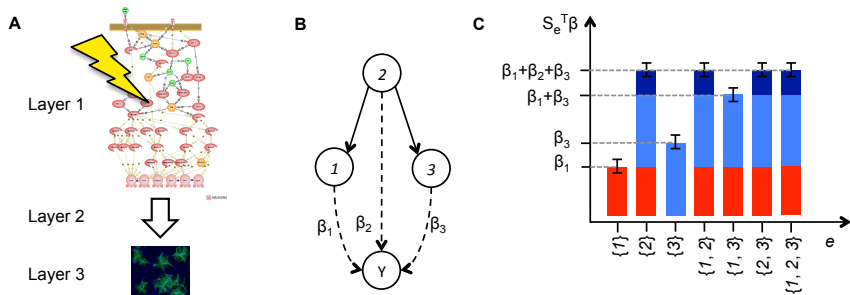
- Estymacja parametrów: algorytm expectation maximization
- Implementacja: Paczka R *muex*

# Co eksperymety RNAi mogą powiedzieć o szlakach sygnalizacyjnych



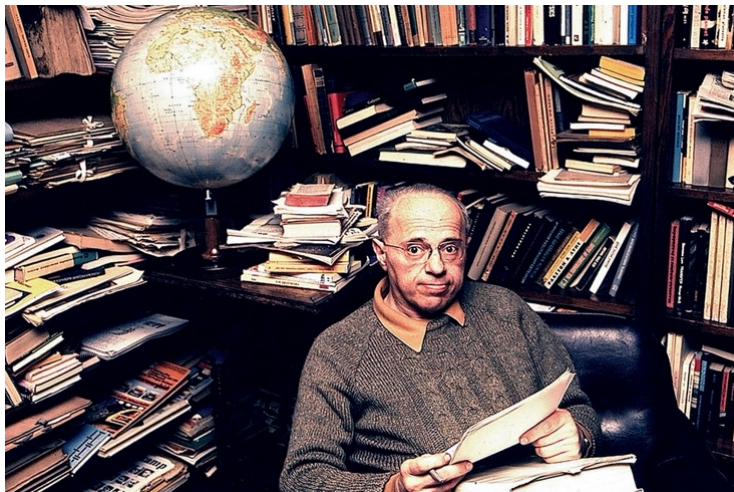


# Linear effects models



- Algorytmy przeszukiwania przestrzeni struktury modeli
- Inferencja modelu: regresja bayesowska
- Implementacja: Paczka R *lem*

# Linear effects models (paczka R *lme*)



# Jak znaleźć drogę w tej dżungli danych?

- ② Które mutacje możemy wykorzystać w terapii raka?
  - Syntetycznie letalni partnerzy genów zmutowanych <sup>4,5</sup>

---

4 Szczurek *et al.* *Int J Cancer*, 2013

5 Matlak & Szczurek, *PLoS Comput. Biol.*, 2017

# Odziaływanie genów

- Dla dwóch genów, ich genotyp to binarna para  $g \in \{0, 1\}^2$ , określająca który z genów jest zmutowany.
- Możliwe genotypy: 00, 10, 01, 11
- Fitness  $\Delta_g$  - miara sukcesu reprodukcyjnego i przeżycia komórki/organizmu o genotypie  $g$

# Oddziaływanie genów

- Dla dwóch genów, ich genotyp to binarna para  $g \in \{0, 1\}^2$ , określająca który z genów jest zmutowany.
- Możliwe genotypy: 00, 10, 01, 11
- Fitness  $\Delta_g$  - miara sukcesu reprodukcyjnego i przeżycia komórki/organizmu o genotypie  $g$

## Oddziaływanie genów

Zachodzi, gdy efekty właściwy i oczekiwany jednoczesnej mutacji genów są różne:  $\Delta_{00} \cdot \Delta_{11} \neq \Delta_{01} \cdot \Delta_{10}$

# Oddziaływanie genów

- Dla dwóch genów, ich genotyp to binarna para  $g \in \{0, 1\}^2$ , określająca który z genów jest zmutowany.
- Możliwe genotypy: 00, 10, 01, 11
- Fitness  $\Delta_g$  - miara sukcesu reprodukcyjnego i przeżycia komórki/organizmu o genotypie  $g$

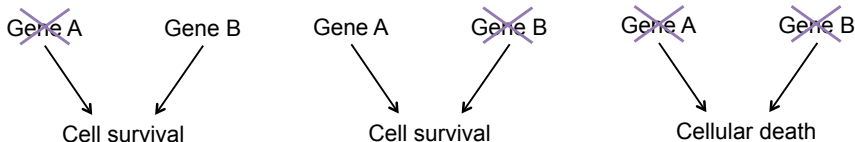
## Oddziaływanie genów

Zachodzi, gdy efekty właściwy i oczekiwany jednoczesnej mutacji genów są różne:  $\Delta_{00} \cdot \Delta_{11} \neq \Delta_{01} \cdot \Delta_{10}$

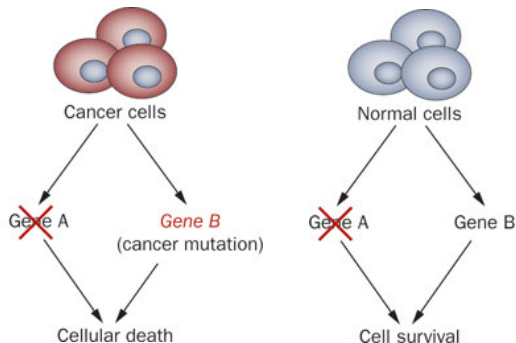
## Syntetyczna letalność: szczególny przypadek oddziaływania genów

Zachodzi, gdy  $\Delta_{00} \cdot \Delta_{11} < \Delta_{01} \cdot \Delta_{10}$ .

# Syntetyczna letalność: szczególny przypadek oddziaływania genów



# Wykorzystanie syntetycznej letalności w terapii nowotworów



- Eksperymentalnie nie jesteśmy w stanie zbadać
  - wszystkich możliwych par genów:  $\sim 200M$
  - tych par genów, dla których istnieją celujące w nie leki:  $\sim 500K$
- Potrzebne są metody analizy danych, wyszukujące pary genów będące obiecującymi kandydatami do syntetycznej letalności!



# Uogólnienie oddziaływań genów

## Oddziaływanie pary genów

Zachodzi, gdy efekty właściwy i oczekiwany jednoczesnej mutacji genów są różne:  $\Delta_{00} \cdot \Delta_{11} \neq \Delta_{01} \cdot \Delta_{10}$ , gdzie  $\Delta_g$  to fitness dla genotypu  $g \in \{0, 1\}^2$ .

# Uogólnienie oddziaływań genów

## Oddziaływanie pary genów

Zachodzi, gdy efekty właściwy i oczekiwany jednoczesnej mutacji genów są różne:  $\Delta_{00} \cdot \Delta_{11} \neq \Delta_{01} \cdot \Delta_{10}$ , gdzie  $\Delta_g$  to fitness dla genotypu  $g \in \{0, 1\}^2$ .

## Oddziaływanie trójki genów

Zachodzi, gdy  $\Delta_{g_0} \cdot \Delta_{g_3} \neq \Delta_{g_1} \cdot \Delta_{g_2}$ , gdzie  $g_0 + g_3 - g_1 - g_2 = 0$ , a  $\Delta_{g_i}$  to fitness dla genotypu  $g_i \in \{0, 1\}^3$ .

# Uogólnienie oddziaływań genów

## Oddziaływanie pary genów

Zachodzi, gdy efekty właściwy i oczekiwany jednoczesnej mutacji genów są różne:  $\Delta_{00} \cdot \Delta_{11} \neq \Delta_{01} \cdot \Delta_{10}$ , gdzie  $\Delta_g$  to fitness dla genotypu  $g \in \{0, 1\}^2$ .

## Oddziaływanie trójki genów

Zachodzi, gdy  $\Delta_{g_0} \cdot \Delta_{g_3} \neq \Delta_{g_1} \cdot \Delta_{g_2}$ , gdzie  $g_0 + g_3 - g_1 - g_2 = 0$ , a  $\Delta_{g_i}$  to fitness dla genotypu  $g_i \in \{0, 1\}^3$ .

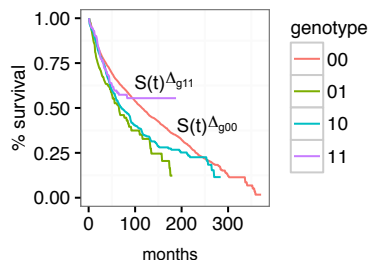
Przykład: syntetyczna letalność warunkowana brakiem mutacji w trzecim genie (biomarker)

Zachodzi, gdy  $\Delta_{000} \cdot \Delta_{110} < \Delta_{100} \cdot \Delta_{010}$ .

# Model oparty o *alternatywy Lehmana*

- $S(t)$  funkcja przeżycia u zdrowych,
- $S_g(t) = S(t)^{\Delta_g}$  -funkcja przeżycia dla pacjenta, którego komórki nowotworowe mają fitness  $\Delta_g$  i genotyp  $g$
- Ponieważ  $S(t) \in [0, 1]$ , im większy fitness nowotworu  $\Delta_g$ , tym niższe przeżycie pacjenta  $S_g(t)$ .
- Dla dwóch nieoddziałujących genów oczekujemy

$$\left(G(t)^{\Delta_{g00}}\right)^{\Delta_{g11}} = \left(G(t)^{\Delta_{g01}}\right)^{\Delta_{g10}}$$



Wykres Kaplana-Meiera dla praj syntetycznie letalnej

BRCA1-PARP1 i różnych genotypów

Test ilorazu wiarygodności dla oddziaływań genów przy

- hipotezie zerowej: brak oddziaływania  $\Delta_{g_0} \Delta_{g_3} = \Delta_{g_1} \Delta_{g_2}$
- hipotezie alternatywnej:  $\Delta_{g_0} \Delta_{g_3} \neq \Delta_{g_1} \Delta_{g_2}$

gdzie  $g_0 - g_1 - g_2 + g_3 = 0$ .

- wiarygodność (ang likelihood) obliczamy zakładając, że  $\Delta_{g_0}, \Delta_{g_3}, \Delta_{g_1}, \Delta_{g_2}$  są parametrami.
- parametry estymowane metodą największej wiarygodności.
- implementacja w paczce R. *survLRT*

# Dziękuję za uwagę!

## Szczególne podziękowania dla

- Niko Beerenwinkel, ETH Zurich,
- Dariusz Matlak, Uniwersytet Warszawski
- R
- Bioconductor

**Zapraszam do zadawania pytań!**



UNIVERSITY  
OF WARSAW



NATIONAL SCIENCE CENTRE  
POLAND