

# Analiza danych czasoprzestrzennych ze szlaków sygnałowych ERK i AKT

Aleksander Janowiak  
ćwiczenie 2 z Biologii systemów

28 kwietnia 2025

## 1 Wstęp

W tym ćwiczeniu przeprowadzono analizę eksploracyjną danych czasoprzestrzennych o ekspresji białek ERK i AKT w komórkach ludzkiego nabłonka z różnymi mutacjami.

Szlaki MAPK/ERK i PI3K-AKT odgrywają kluczową rolę w regulacji cyklu komórkowego. Tym samym, od ich funkcjonowania zależy sposób w jaki komórka reaguje na stres i adaptuje się do zmian środowiska. Naturalnie, białka, których dotyczą pomiary w analizowanym eksperymencie biorą udział w regulacji podziałów komórkowych i procesu różnicowania, co jest szczególnie istotne dla osób badających nowotworzenie.

W omawianym eksperymencie podjęto analizę komórek z linii MCF 10A (ludzki nabłonek gruczołu mlekowego) w pięciu wariantach mutacyjnych. Oprócz typu dzikiego (WT), badano jeszcze 4 zestawy komórek o następujących mutacjach: AKT1\_E17K, PIK3CA\_E545K, PIK3CA\_H1047R, PTEN\_del. Wszystkie z nich dotyczą szlaku PI3K-AKT.

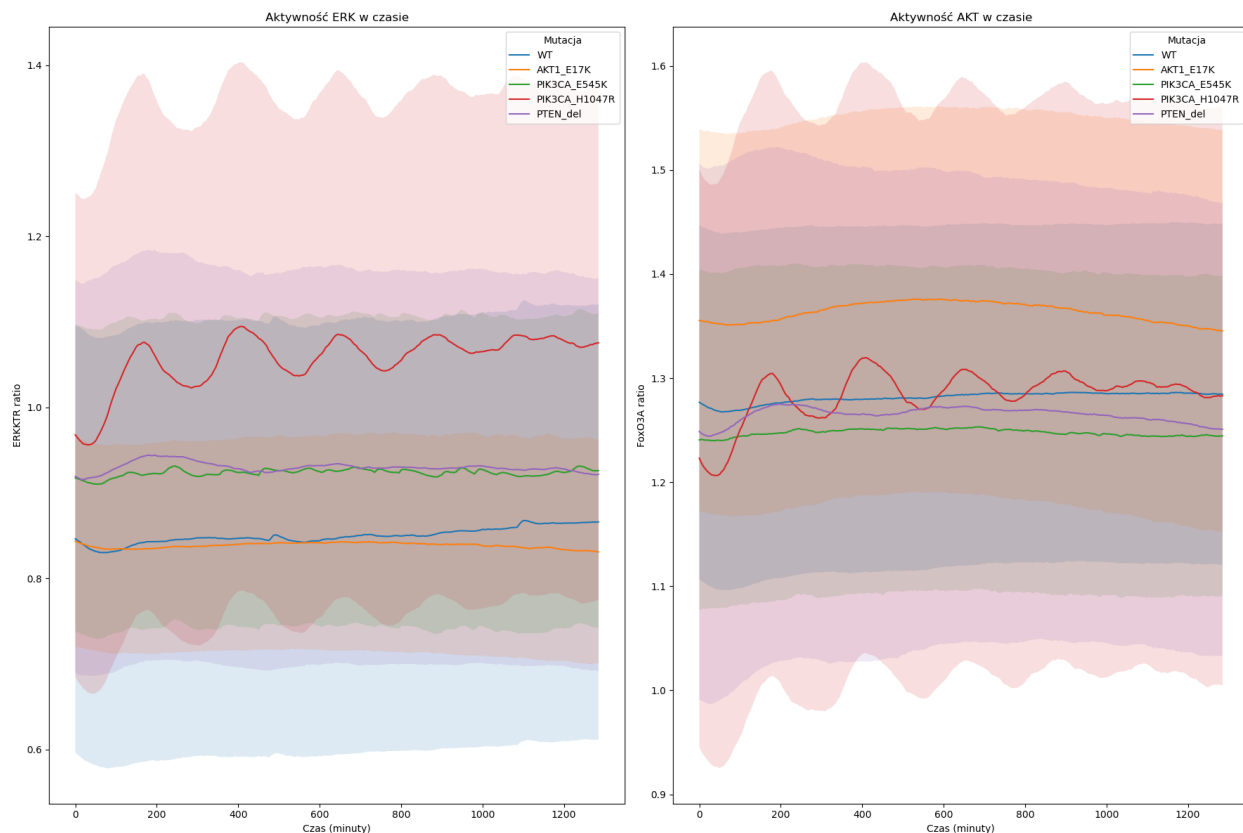
Do pomiaru poziomu aktywności białek ERK i AKT wykorzystano biosensory fluorescencyjne ErkKTR-mTurq.2 (dla ERK) i FoxO-mNeonGreen (dla AKT). Dane zbierano co 5 minut przez niecałe 24 godziny. Eksperymenty powtarzano wielokrotnie, około 20 razy dla każdej mutacji.

W następnych sekcjach opisano przeprowadzone analizy, mające na celu zbadanie poziomów ekspresji białek ERK i AKT oraz charakterystykę rozchodzenia się sygnału w czasie i przestrzeni.

## 2 Porównanie aktywności szlaków sygnałowych między mutacjami

### 2.1 Średni poziom ekspresji w czasie

Dla każdej mutacji obliczono średnie poziomy ERK i AKT w danym punkcie czasowym. Obliczono także odchylenie standardowe dla poziomu ekspresji każdego z białek w danym punkcie czasowym, w zależności od mutacji. Krzywe średnich ekspresji oraz obszary w mieszczące się od nich w odległości  $\pm$  jednego odchylenia standardowego przedstawiono na *Rysunku 1*.



Rysunek 1: Średnie poziomy ERK i AKT w czasie w zależności od mutacji  $\pm$  jedno odchylenie standardowe

Jak widać na *Rysunku 1*, średni poziom ERK różni się między komórkami o różnych mutacjach, ale nie wykazuje bardzo znaczących zmian w czasie. Wyjątkiem od tej reguły jest mutacja PIK3CA\_H1047R. Na odpowiadającej tym komórkom części wykresu widać wyraźne oscylacje poziomu ERK - co około 4 godziny poziom tego białka szybko rośnie, a następnie, szybko spada. To też dla tej mutacji poziom ERK jest najwyższy i ma największą wariancję.

Na wykresie średniego poziomu AKT w czasie, podobnie jak w przypadku ERK, komórki PIK3CA\_H1047R wykazują oscylacje w czasie oraz dużą wariancję. Jednak najwyższym bezwzględnym poziomem białka, w tym przypadku, charakteryzują się komórki AKT1\_E17K.

Taki rezultat nie jest niczym zaskakującym: mutacja AKT1\_E17K polega na substytucji glutaminianu na lizynę w pozycji 17 białka AKT co skutkuje jego konstytutywną aktywacją niezależnie od okoliczności.

Spodziewano się również, że komórki typu dzikiego będą wykazywać najniższy poziom AKT, okazało się to jednak nieprawdą. W omawianym eksperymencie, najniższy poziom AKT prezentują komórki z mutacją PIK3CA\_E545K. To zaskakujący wynik, ponieważ mutacja ta powinna pobudzać szlak PI3K-AKT.

Na podstawie tych wyników, mimo że są bardzo ciekawe, należy ostrożnie wyciągać wnioski, gdyż wariancja wewnątrz próbek jest większa niż wariancja między próbkami o różnych mutacjach.

Najciekawszym ze wszystkich wyników są opisane wcześniej oscylacje poziomów obu białek w komórkach PIK3CA\_H1047R. Niestety, nie potrafię biologicznie wyjaśnić tego zjawiska. Być może jest ono związane z zaburzeniem systemu sprzężeń zwrotnych pomiędzy szlakami MAPK/ERK i PI3K-AKT wywołanych wprowadzoną mutacją.

## 2.2 Wnioskowanie statystyczne o różnicach między mutacjami

### 2.2.1 Wstępne testy

W celu bardziej gruntownego uargumentowania zaobserwowanych pomiędzy mutacjami różnic w poziomach ERK i AKT przeprowadzono testy Kruskala-Wallisa. Test ten porównuje jednocześnie wszystkie mutacje, aby zdecydować czy któraś próbka pochodzi z innego rozkładu niż pozostałe. Wyniki testów Kruskala-Wallisa, dla poziomów ERK i AKT przedstawiono w Tabeli 1.

Testy przeprowadzono na ekspresji ERK i AKT w zdefiniowanym oknie czasowym, w poniższym przykładzie było to 60-180 minut. W celu przeprowadzenia testów uśredniono (niezależnie od siebie) poziomy białek w danej komórce, po wszystkich punktach czasowych zdefiniowanego okna. Jest to podejście odmienne od poprzedniej analizy, w której uśredniano wszystkie komórki w danym punkcie czasowym.

testowane białko \ wynik testu	wartość H-statystyki	p-wartość
ERK	15815.0	0.0
AKT	7742.5	0.0

Tabela 1: Wyniki testów Kruskala-Wallisa dla średnich poziomów ERK i AKT pomiędzy wszystkimi mutacjami

Bardzo niskie p-wartości (dużo mniejsze niż 0.05) w obu testach wskazują na to, że zarówno w przypadku ERK, jak i AKT, średni poziom ekspresji białka w przynajmniej jednej próbce

pochodzi z innego rozkładu niż pozostałe. Wynik ten jest zgodny z przeprowadzoną analizą *Rysunku 1*, w której wskazano na wyraźne różnice pomiędzy komórkami o różnych mutacjach.

Wartości H-statystyki będącej wynikiem testu, nie jest dla nas w tym przypadku informatywna. Podano je z przyzwoitości.

### 2.2.2 Dokładniejsze testy

Chcąc dokładniej zbadać rozbieżności w poziomach ERK i AKT przeprowadzono parami test Manna-Whitneya pomiędzy WT i pozostałymi mutacjami. Test Manna-Whitneya, tak samo jak test Kruskala-Wallisa, sprawdza czy próbki pochodzą z tego samego rozkładu prawdopodobieństwa; robi to jednak dla dwóch próbek, nie dla wielu. Wyniki testów Manna-Whitneya dla ERK i AKT przedstawiono odpowiednio w *Tabeli 2* i *Tabeli 3*. Ze względu na wielokrotne testowanie, zastosowano korektę Bonferroniego.

testowana mutacja \ wynik testu	wartość U-statystyki	p-wartość po korekcie
AKT1_E17K	1.083788e+09	0.0
PIK3CA_E545K	4.963474e+08	0.0
PIK3CA_H1047R	1.029544e+09	0.0
PTEN_del	6.458919e+08	0.0

Tabela 2: Wyniki testów Manna-Whitneya poziomow ERK przeprowadzonych pomiędzy WT i pozostałymi mutacjami

testowana mutacja \ wynik testu	wartość U-statystyki	p-wartość po korekcie
AKT1_E17K	392649178.5	0.0
PIK3CA_E545K	611300621.5	7.623906e-118
PIK3CA_H1047R	325795432.5	3.955092e-09
PTEN_del	367240844.5	1.091781e-02

Tabela 3: Wyniki testów Manna-Whitneya poziomow AKT przeprowadzonych pomiędzy WT i pozostałymi mutacjami

Bardzo niskie, nawet po korekcie Bonferroniego, p-wartości wskazują na statystyczną istotność zaobserwowanych różnic w poziomach ERK i AKT. Z przeprowadzonych testów Manna-

Whitneya wynika, że poziomy zarówno ERK jak i AKT we wszystkich mutacjach istotnie odbiegają od odpowiednich poziomów w komórkach typu dzikiego.

### 3 Analiza koordynacji przestrzennej i dynamiki sygnału

Do kolejnych analiz konieczne było zdefiniowanie stanu aktywnego komórki. Zdecydowano się na następującą definicję wykorzystującą wartości ERKTR\_ratio:

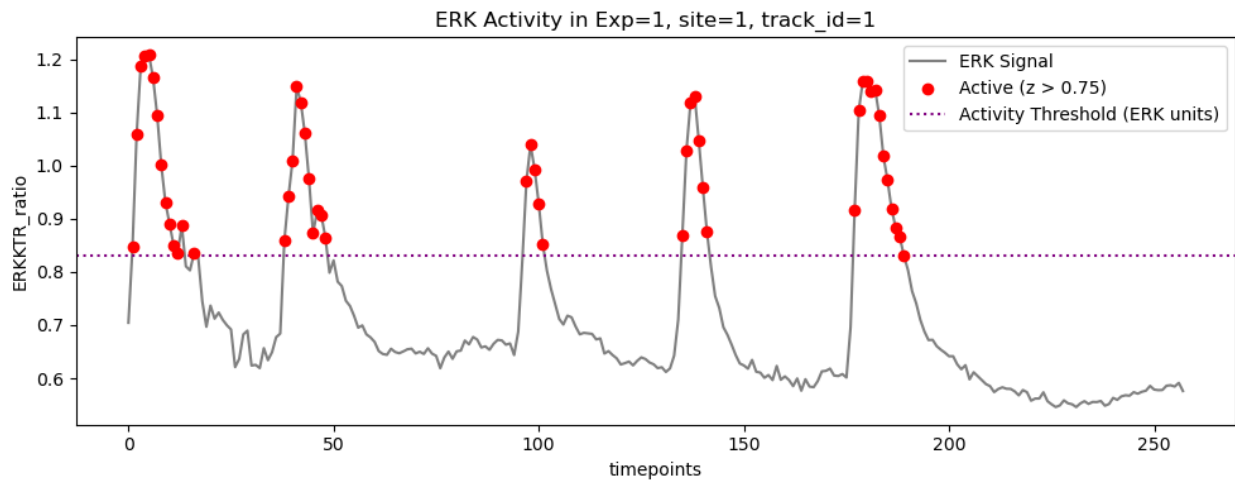
Komórka  $i$  jest uznawana za aktywną w chwili  $t$  gdy odpowiadający jej  $z_{score}$  przekracza pewien zdefiniowany próg:  $z_{i,t} > threshold$  (ustawiony na wartość 0.75). Z kolei  $z_{score}$  jest zdefiniowany jako:

$$z_{i,t} = \frac{x_{i,t} - \mu_i}{\sigma_i}$$

Gdzie:

- $x_{i,t}$  jest sygnałem ERK komórki  $i$  w czasie  $t$
- $\mu_i$  jest średnim sygnałem ERK komórki  $i$  w czasie
- $\sigma_i$  jest odchyleniem standardowym sygnału ERK komórki  $i$  w czasie

Zdecydowano się na wyznaczanie aktywności komórki względem poziomu ERK, gdyż jest to ważniejsze z dwóch białek z punktu widzenia podziałów komórkowych i proliferacji - a to właśnie z tymi procesami kojarzy mi się pojęcie "komórka aktywna". Wybrana metoda wnioskowania o aktywności komórkowej oparta na ustaleniu progu do przekroczenia przez  $z_{score}$  jest elegancka i wygodna, gdyż eliminuje ona komplikacje wynikające z różnych wielkości odchylenia i różnych bazowych poziomów ERK między komórkami. Poniżej, na *Rysunku 2* przedstawiono przykładowy wykres poziomu ERK w komórce wraz z zaznaczonymi punktami, w których komórka jest aktywna i progiem odcięcia.

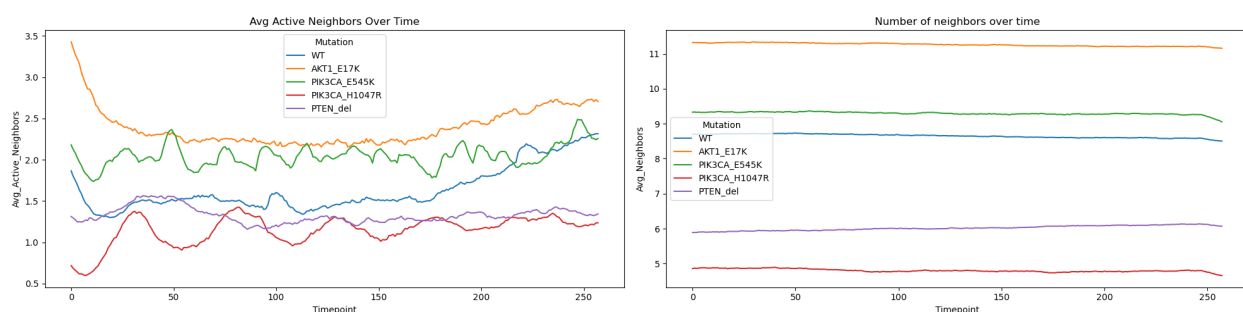


Rysunek 2: Poziom ERK w pierwszej komórce pierwszego eksperymentu wraz z progiem odcięcia

### 3.1 Liczba aktywnych sąsiadów w czasie

Dla każdego typu komórek obliczono średnią liczbę aktywnych sąsiadów w danym punkcie czasowym, wyniki tego obliczenia pokazano na *Rysunku 3*. Liczba aktywnych sąsiadów komórki jest w oczywisty sposób związana z zagęszczeniem komórek. Z tego powodu, jako kontekst, umieszczono na *Rysunku 3* również średnią liczbę sąsiadów (niezależnie od aktywności) komórek każdego typu.

Dla obu wykresów, kryterium uznania danych dwóch komórek za sąsiadów była odległość pomiędzy ich jądrami. Przyjęto  $50\ \mu\text{m}$  jako wartość graniczną - komórki oddalone od siebie o więcej niż  $50\ \mu\text{m}$  nie są sąsiadami w przyjętym modelu. Sąsiedztwo komórek wyznaczono korzystając ze struktury KDTree.



Rysunek 3: Średnia liczba aktywnych sąsiadów oraz wszystkich sąsiadów dla komórek o różnych typach mutacji

W przypadku większości mutacji, średnia liczba aktywnych sąsiadów nie wykazuje bardzo wyraźnych trendów. Ciekawy jest obserwowany wzrost tej wartości w grupie WT pod koniec eksperymentu, jak i fakt, że dla komórek o mutacji PIK3CA\_H1047R na wykresie średniej liczby aktywnych sąsiadów występują podobne oscylacje co na wykresach pokazanych na *Rysunku 1*. Godna uwagi wydaje się również stosunkowo mała średnia aktywnych sąsiadów dla komórek AKT1\_E17K w porównaniu do ich zagęszczenia.

### 3.2 Rozprzestrzenianie się aktywności sygnałowej

W celu wizualnej oceny sposobu rozprzestrzeniania się aktywności sygnałowej przygotowano 5 animacji. Dla każdej mutacji wybrano losowo jedno powtórzenie eksperymentu. Następnie, na podstawie danych pochodzących z wylosowanego powtórzenia przygotowano wizualizację przestrzenną komórek aktywnych i nieaktywnych. Gify będące wynikiem tej analizy można obejrzeć w tym samym repozytorium GitHub, w którym znajduje się ten raport. Na *Rysunku 4* umieszczono po jednej przykładowej klatce z każdej animacji:

Na *Rysunku 4* wyraźnie widać różnice w zagęszczeniu komórek pomiędzy różnymi mutacjami. Różnice te są zgodne z drugim wykresem *Rysunku 3*: komórki typu AKT1\_E17K są rozmieszczone najgęściej, a PIK3CA\_H1047R - najmniej gęsto. Na pierwszej klatce wizualizacji dla komórek AKT1\_E17K widać również bardzo dużo komórek aktywnych - odpowiada

to wysokiej wartości na początku pierwszego wykresu *Rysunku 3*.

Nawet po kilkukrotnym obejrzeniu wszystkich wizualizacji trudno jest wywnioskować cokolwiek o sposobie propagacji aktywności sygnałowej między komórkami. W niektórych regionach obrazka, komórki aktywne wydają się nie mieć ze sobą nic wspólnego. W innych, sygnał wygląda jakby rozchodził się w sposób kierunkowy (w kształcie przypominającym błyskawicę). Wynik ten mnie zaskoczył, gdyż spodziewałem się efektu bardziej podobnego do fali centralnej, czyli promienistego rozchodzenia się sygnału z jednej komórki we wszystkich kierunkach.

W następnej sekcji zajęto się matematycznym poszukiwaniem rzeczonego "efektu fali centralnej", a gdy go nie znaleziono, postawiono następujące pytanie: "Czy komórki aktywne w ogóle mają ze sobą cokolwiek wspólnego?".

## 4 Poszukiwanie efektu falowego

Postawiono następującą hipotezę: **Sygnał rozchodzi się w sposób falowy, we wszystkich kierunkach, niezależnie od mutacji.**

W celu ilościowego zbadania zjawiska rozchodzenia się sygnału ERK pomiędzy komórkami przeprowadzono następującą analizę:

1. Wyznaczono wszystkie (charakteryzowane identyfikatorem komórki i punktem w czasie) przypadki aktywacji komórki. Aktywacja komórki  $i$  w momencie  $t$  następuje wtedy gdy jest ona aktywna w momencie  $t$  i nieaktywna w momencie  $t - 1$ .
2. Dla każdej komórki  $i$ , dla której w czasie  $t$  zaszła aktywacja:
  - (a) Wyznaczono sąsiadów w momencie  $t$  za pomocą struktury KDTree
  - (b) Dla wyznaczonych sąsiadów sprawdzono poziomy ERK w czasie  $t - dt$ ,  $t$  oraz  $t + dt$
3. Zebrano wyniki dla  $dt = 1$  i przedstawiono je na wspólnym wykresie skrzypcowym (*Rysunek 5a*) oraz na oddzielnych wykresach, z rozróżnieniem na mutacje (*Rysunki 5b-5f*).

Niestety nie znaleziono istotnego skoku w poziomie ERK u sąsiadów danej komórki po jej aktywacji - na *Rysunku 5* wszystkie wykresy z danego zestawu wyglądają praktycznie identycznie. Stanowi to podstawę do odrzucenia hipotezy o falowej naturze rozchodzenia się aktywności sygnałowej pomiędzy komórkami. Należy również wspomnieć, że taki wynik nie jest spowodowany zbyt małą wartością parametru  $dt$ , gdyż dla większego przedziału czasowego wykresy wyglądają tak samo.

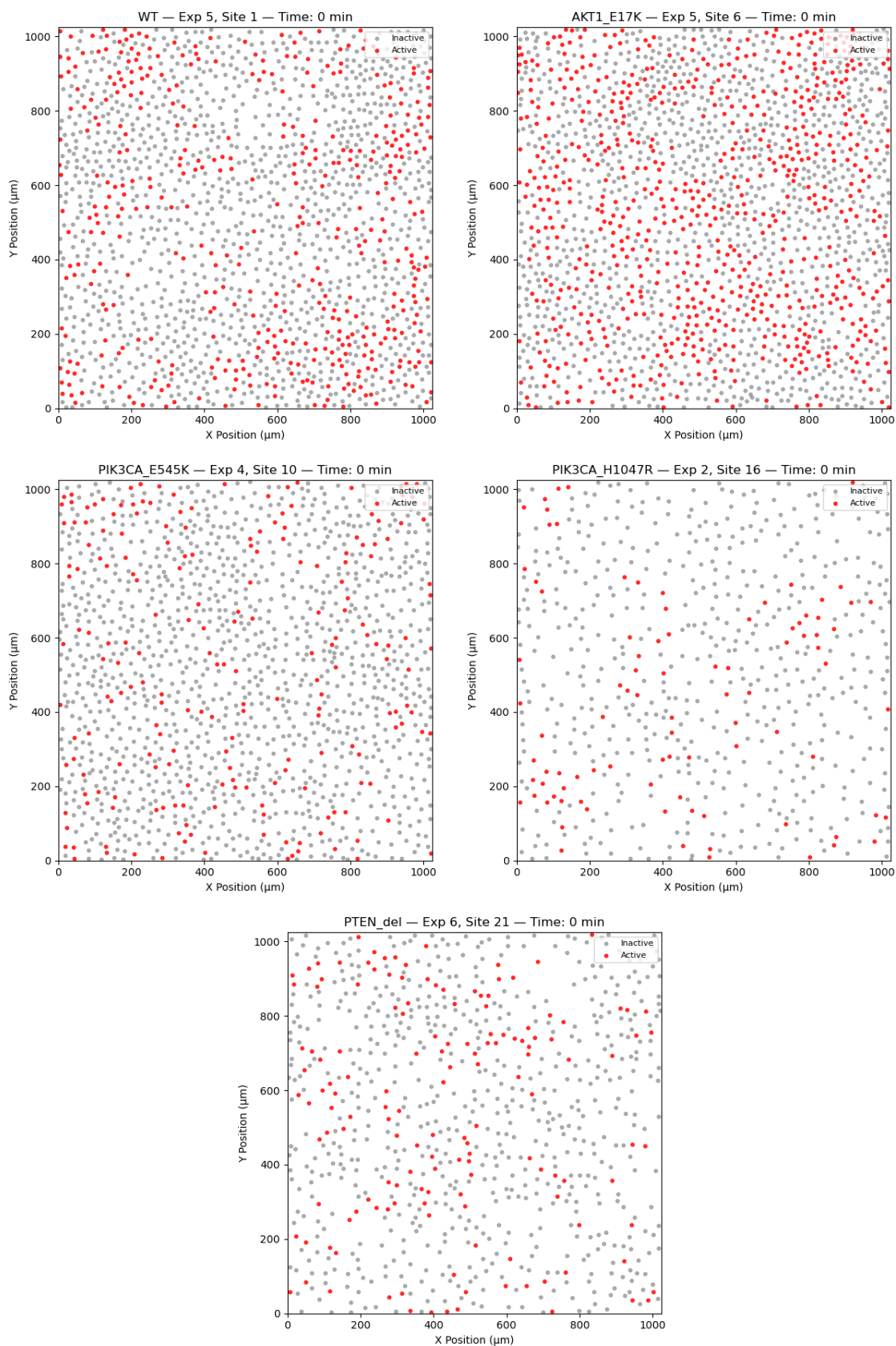
Drobnych różnic pomiędzy trzema rozkładami można doszukiwać się w przypadku mutacji WT, PTEN\_del oraz PIK3CA\_H1047R, jednak nawet tam rozbieżności są dosyć subtelne.

Co ciekawe, tym trzem mutacjom odpowiadają eksperymenty o najmniej gęstym rozmieszczeniu komórek. A najważniejszą obserwacją jest to, że mutacja PIK3CA\_H1047R, w której widać zdecydowanie największe różnice pomiędzy trzema wykresami skrzypcowymi, charakteryzuje się zdecydowanie najmniejszym zagęszczeniem komórek.

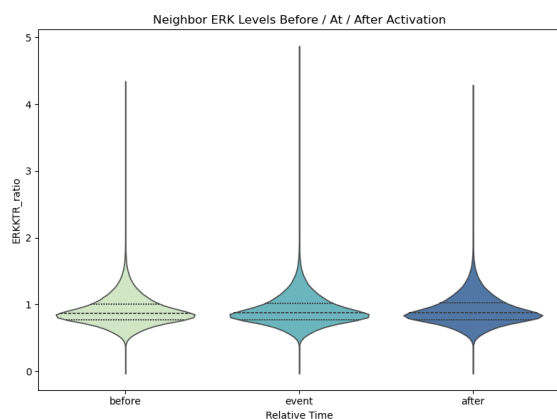
Możliwe jest zinterpretowanie uzyskanych wyników jako przesłanki za tym, że sygnał rozchodzi się w sposób kierunkowy. Jest tak, ponieważ gdyby sygnał rozchodził się tylko w jednym kierunku, np. zawsze był przekazywany do tylko jednego sąsiada (zamiast do wszystkich naokoło) to wykres skrzypcowy poziomu ERK u wszystkich sąsiadów i tak powinien takie zjawisko wychwycić. Jednak wtedy, efekt byłby lepiej widoczny w populacjach o mniejszym zagęszczeniu: mniej sąsiadów wiąże się z większym wpływem jednostki na kształt rozkładu.

Dokładnie taką zależność między gęstością populacji i różnicami w wykresach skrzypcowych można dostrzec na *Rysunku 5*. Jest to jednak bardzo odważna interpretacja wyników. Niestety, bez dalszych analiz próba powiedzenia czegokolwiek więcej o naturze rozchodzenia się sygnału byłaby bardziej snuciem domysłów niż naukowym wnioskowaniem.

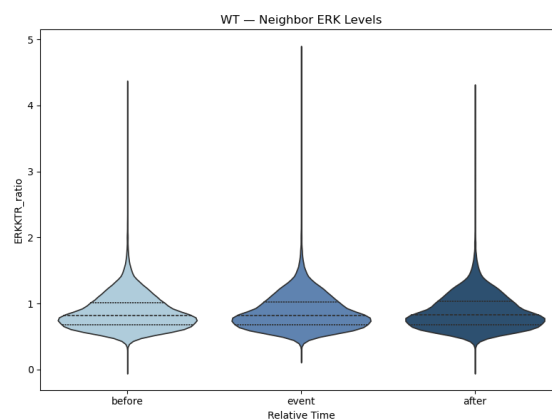




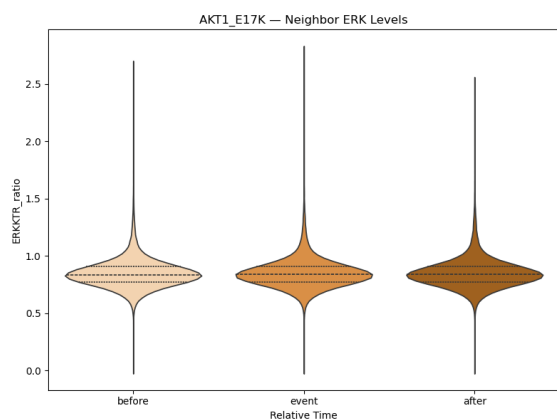
Rysunek 4: Pierwsza klatka każdej z pięciu wizualizacji.



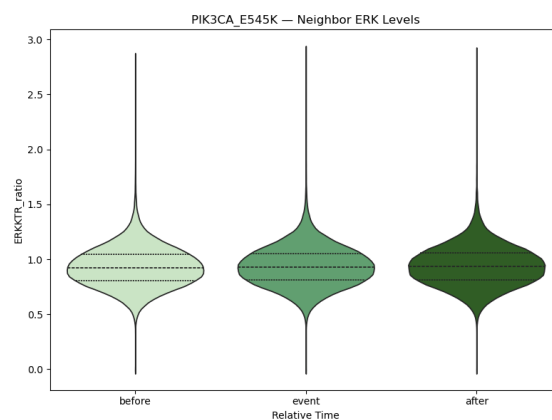
(a) Wspólny wykres dla wszystkich mutacji



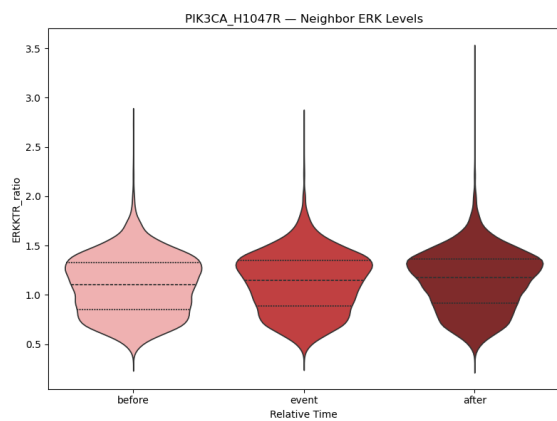
(b) Wykresy dla WT



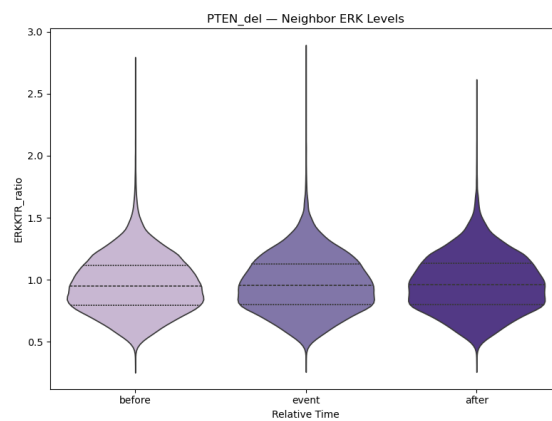
(c) Wykresy dla AKT1\_E17K



(d) Wykresy dla PIK3CA\_E545K



(e) Wykresy dla PIK3CA\_H1047R



(f) Wykresy dla PTEN\_del

Rysunek 5: Wykresy szkrzypcowe poziomów ERK sąsiadów komórek aktywowanych dla chwil  $t - 1$ ,  $t$  oraz  $t + 1$ .