

Agnieszka Prudło

Cel projektu: Eksploracja czasoprzestrzennych danych pojedynczych komórek w kontekście aktywności szlaków ERK i AKT dla różnych linii komórkowych (WT, AKT1_E17K, PIK3CA_E545K, PIK3CA_H1047R, PTEN_del)

Wszystkie wyniki zostały uzyskane za pomocą programów i funkcji napisanych w języku Python.

Zadanie

1

Cel: Porównanie dynamiki aktywności szlaków ERK i AKT między liniami komórkowymi o różnych mutacjach.

Pytanie badawcze: Czy poszczególne mutacje wpływają na aktywność szlaków ERK i AKT w komórkach?

Wykresy aktywności ERK i AKT w czasie

Dla każdego punktu czasowego została policzona średnia wartość oraz odchylenie standardowe ERKKTR_ratio (ERK) oraz FoxO3A_ratio (AKT) dla każdej mutacji osobno.

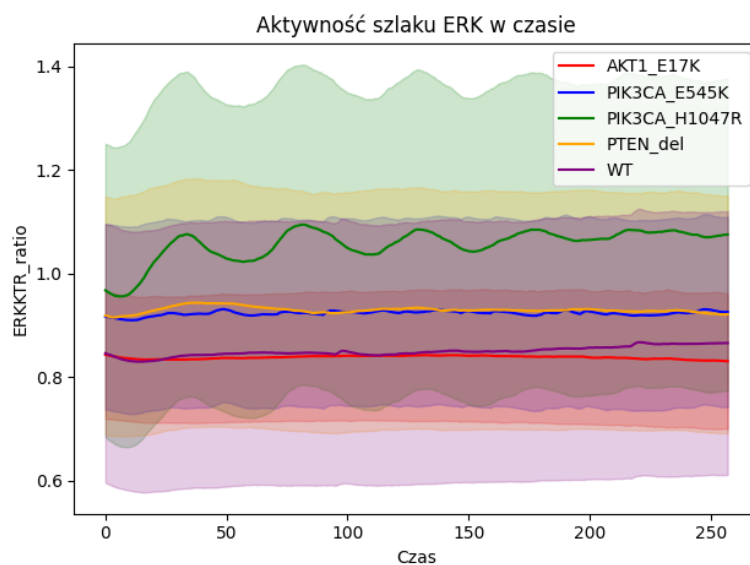


fig. 1) Średni przebiegi aktywności ERK wraz z zaznaczonymi odchyleniami standardowymi dla każdej mutacji w czasie. Kolorem czerwonym zaznaczona jest mutacja AKT1_E17K, niebieskim PIK3CA_E545K, zielonym PIK3CA_H1047R, żółtym PTEN_del, a fioletowym typ dziki (WT).

Na wykresie (fig. 1) widać że najbardziej zmienny charakter aktywności ERK występuje przy mutacji PIK3CA_H1047R. Szczególnie na początku widać okresowe wzrosty i spadki aktywności. Dla pozostałych mutacji ERK utrzymuje się na mniej więcej stałym poziomie. Dla wszystkich mutacji poza AKT1_E17K widać wzrost aktywności w stosunku do typu dzikiego, najwyższy dla mutacji PIK3CA_H1047R.

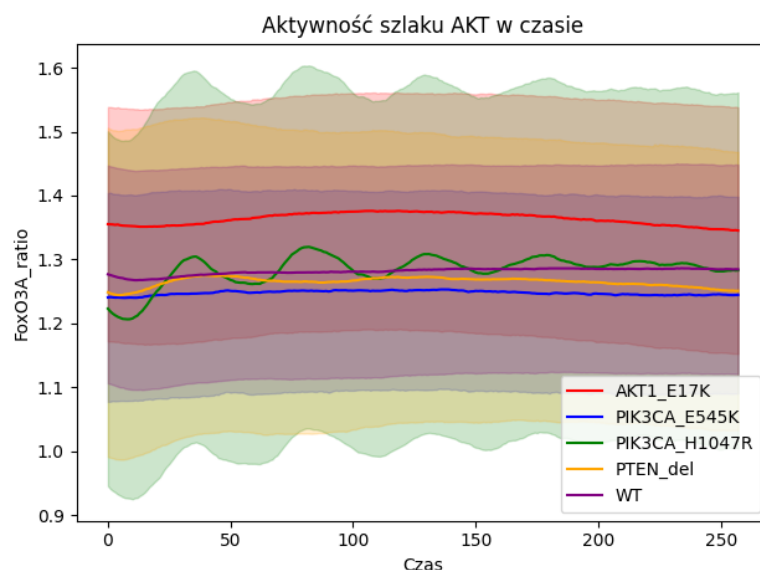


fig. 2) Średni przebiegi aktywności AKT wraz z zaznaczonymi odchyleniami standardowymi dla każdej mutacji w czasie. Kolorem czerwonym zaznaczona jest mutacja AKT1_E17K, niebieskim PIK3CA_E545K, zielonym PIK3CA_H1047R, żółtym PTEN_del, a fioletowym typ dziki (WT).

Na wykresie (fig. 2) również widoczny jest okresowy charakter aktywności w przypadku mutacji PIK3CA_H1047R. Dla pozostałych mutacji AKT utrzymuje się na mniej więcej stałym poziomie. Najwyższe wartości osiągane są dla mutacji AKT1_E17K. W stosunku do typu dzikiego niższa aktywność AKT prezentuje się w mutacjach PIK3CA_E545K i PTEN_del, a wartości dla PIK3CA_H1047R oscylują wokół wartości dla WT (szczególnie na początku, później PIK3CA_H1047R coraz bardziej przypomina WT).

Testy istotności statystycznej Manna–Whitneya

Dla każdego szlaku (ERK i AKT) przeprowadzono 4 testy statystyczne Manna–Whitneya, porównujące rozkłady wartości aktywności pomiędzy typem WT i każdą z mutacji (jeden test na mutację). Wybrany przedział czasowy to Image_Metatdata_T w granicach 60–180 (czyli 300–540 minut, ze względu na akwizycję obrazu co 5 minut). Porównano też średnie wartości w całym tym przedziale czasowym pomiędzy mutacjami a typem WT w celu oszacowania, czy mutacja wywołała spadek czy wzrost aktywności. Następnie przeprowadzono korektę Bonferroniego. Uzyskane wyniki prezentują się w tabeli poniżej:

Mutacja	Szlak	p_raw	U	Zmiana	p_bonferroni	Istotne (a=0.05)
AKT1_E17K	ERK	1.2466518066935373e-40	53.0	spadek	9.973214453548299e-40	PRAWDA
PIK3CA_E545K	ERK	3.360343522054828e-41	14641.0	wzrost	2.6882748176438622e-40	PRAWDA
PIK3CA_H1047R	ERK	3.360343522054828e-41	14641.0	wzrost	2.6882748176438622e-40	PRAWDA
PTEN_del	ERK	3.360343522054828e-41	14641.0	wzrost	2.6882748176438622e-40	PRAWDA
AKT1_E17K	AKT	3.360343522054828e-41	14641.0	wzrost	2.6882748176438622e-40	PRAWDA
PIK3CA_E545K	AKT	3.360343522054828e-41	0.0	spadek	2.6882748176438622e-40	PRAWDA
PIK3CA_H1047R	AKT	4.667843403553585e-07	10065.0	wzrost	3.734274722842868e-06	PRAWDA
PTEN_del	AKT	3.360343522054828e-41	0.0	spadek	2.6882748176438622e-40	PRAWDA

Tab. 1) Wyniki testów Manna–Whitneya. P_raw to surowa p-wartość, U to wartość statystyki U Manna–Whitneya, Zmiana to wpływ mutacji na aktywność szlaków oszacowana przez średnią w całym przedziale czasowym, p_bonferroni to poprawiona p-wartość korektą Bonferroniego, kolumna Istotne mówi, czy różnica jest istotna statystycznie z progiem alpha=0.05.

Z tabeli widać, że wszystkie różnice w aktywności szlaków ERK i AKT pomiędzy poszczególnymi mutacjami a typem dzikim są statystycznie istotne. Mutacje AKT1_E17K, PIK3CA_E545K oraz PTEN_del spowodowały odmienne zmiany aktywności: wzrost w jednym szlaku i spadek w drugim. Mutacja AKT1_E17K powodowała istotny wzrost aktywności szlaku AKT oraz spadek aktywności szlaku ERK. PIK3CA_E545K oraz PTEN_del wywołały wzrost aktywności szlaku ERK oraz spadek aktywności szlaku AKT. Z kolei mutacja PIK3CA_H1047R prowadziła do istotnego wzrostu aktywności obu szlaków jednocześnie, aczkolwiek efekt wzrostu dla AKT jest mniej mocny (p_wartość jest znacznie wyższa niż dla pozostałych, choć nadal < 0.05 po Bonferroni).

Wnioski

Analizowane mutacje w genach szlaku PI3K–AKT–mTOR istotnie wpływają na aktywność szlaków sygnałowych AKT i ERK.

Mutacja AKT1_E17K, prowadząca do konstytutywnej aktywacji białka AKT1 niezależnie od sygnałów zewnętrznych, skutkowała wzrostem aktywności szlaku AKT oraz spadkiem aktywności szlaku ERK. Być może sugeruje to wzajemną regulację pomiędzy tymi szlakami.

Mutacje PIK3CA_E545K i PTEN_del, mimo że prowadzą do nadaktywacji szlaku PI3K–AKT, powodowały wzrost aktywności szlaku ERK oraz spadek aktywności szlaku AKT. Może to świadczyć na przykład o przekierowaniu sygnałów na inne ścieżki przy zaburzonym działaniu PI3K–AKT.

Z kolei mutacja PIK3CA_H1047R prowadziła do wzrostu aktywności obu szlaków. Taki efekt może również być wynikiem wzajemnej regulacji tych szlaków, ale tym razem w taki sposób, że jeden napędza drugi.

Wszystkie zmiany były istotne statystycznie, co wzmacnia wiarygodność wyników. W sumie dane wskazują, że choć wszystkie mutacje powodowały nadaktywację PI3K–AKT, to prowadziły one do odmiennych efektów na poziomie dalszej sygnalizacji komórkowej.

Zadanie 2

Cel: Ustalenie przestrzennego kontekstu aktywności sygnałowej oraz możliwej propagacji sygnału w czasie

Pytanie badawcze: Jak aktywność przedstawia się w przestrzeni i czy propaguje się w czasie?

Wybrany wskaźnik: ERKKTR_ratio

Definicja stanu aktywnego komórki:

Aby komórkę uznać za aktywną, musi ona mieć w danym momencie czasowym wartość wskaźnika ERKKTR_ratio powyżej 90. percentyla wszystkich swoich wartości. Ponadto musi być lokalnym maksimum (wyszukiwanym za pomocą funkcji `scipy.signal.find_peaks()`), lub wartością o 2 klatki czasowe wcześniej lub później (ale dalej powyżej 90. percentyla). Czyli komórka jest aktywna w klatkach czasowych, w których sygnał przekracza 90. percentyl jej własnego sygnału i jest w pobliżu lokalnych maksimów (± 2 klatki wokół piku). Taka definicja jest bardziej biologicznie realistyczna niż same lokalne maksima, ponieważ sygnały są rozciągnięte w czasie, a nie pojedyncze.

Boxplot liczby pików w zależności od mutacji

Ze względu na definicję stanu aktywnego komórki i użycie funkcji `find_peaks()`, łatwo jest znaleźć liczbę pików (lokalnych maksimów powyżej 90 percentyla) w zależności od mutacji.

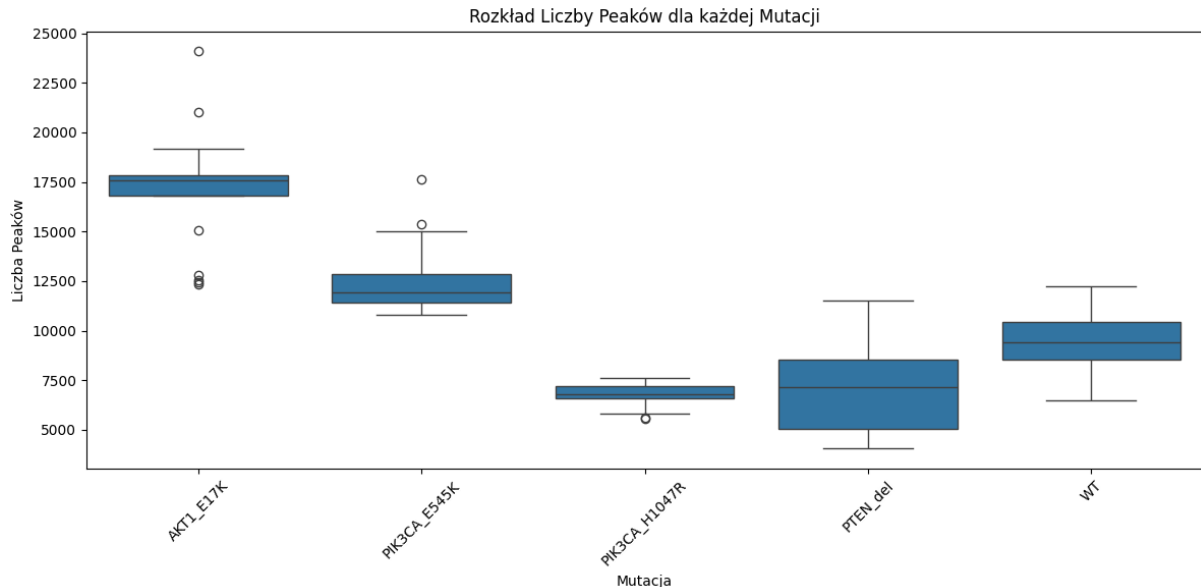


fig. 3) Rozkład liczby peaków w zależności od mutacji. Wykres uzyskano poprzez zliczanie pików dla każdego eksperymentu i site osobno.

Komórki z mutacją AKT1_E17K mają największą liczbę pików, co sugeruje dużą dynamikę sygnału i częstsze zmiany stanu aktywacji. Mutacja PIK3CA_E545K powoduje umiarkowane zwiększenie liczby pików w porównaniu do komórek typu dzikiego, co również może oznaczać dużą liczbę zmian stanu aktywacji. Komórki z mutacją PIK3CA_H1047R mają najmniejszą liczbę pików, co wskazuje na stabilniejszy, mniej zmienny przebieg sygnału ERKTR. Mutacja PTEN_del oraz komórki dzikie prezentują wartości pośrednie, przy czym komórki PTEN_del cechują się większym rozrzutem liczby pików, co może świadczyć o większej różnorodności komórek.

Znajdowanie liczby aktywnych sąsiadów:

Dla każdego eksperymentu, site i punktu czasowego zostało policzone osobne KDTree. Wtedy dla każdej komórki można było policzyć liczbę aktywnych sąsiadów w zadanym promieniu na podstawie definicji aktywności podanej powyżej. Promień dla każdego KDTree został ustalony oddzielnie jako średnia odległość do najbliższego sąsiada powiększona o odchylenie standardowe. Czyli promień = średnia odległość + odchylenie standardowe. Dodanie wartości odchylenia standardowego pozwala lekko rozszerzyć promień, aby złapać nie tylko najbliższych sąsiadów, ale też tych będących trochę dalej. Taka adaptacyjna wielkość promienia uwzględnia zmienności w gęstości upakowania komórek.

Dla każdego momentu czasowego policzono średnią liczbę aktywnych sąsiadów (osobno dla każdej z mutacji) i przedstawiono ją na wykresie poniżej wraz z odchyleniem standardowym.

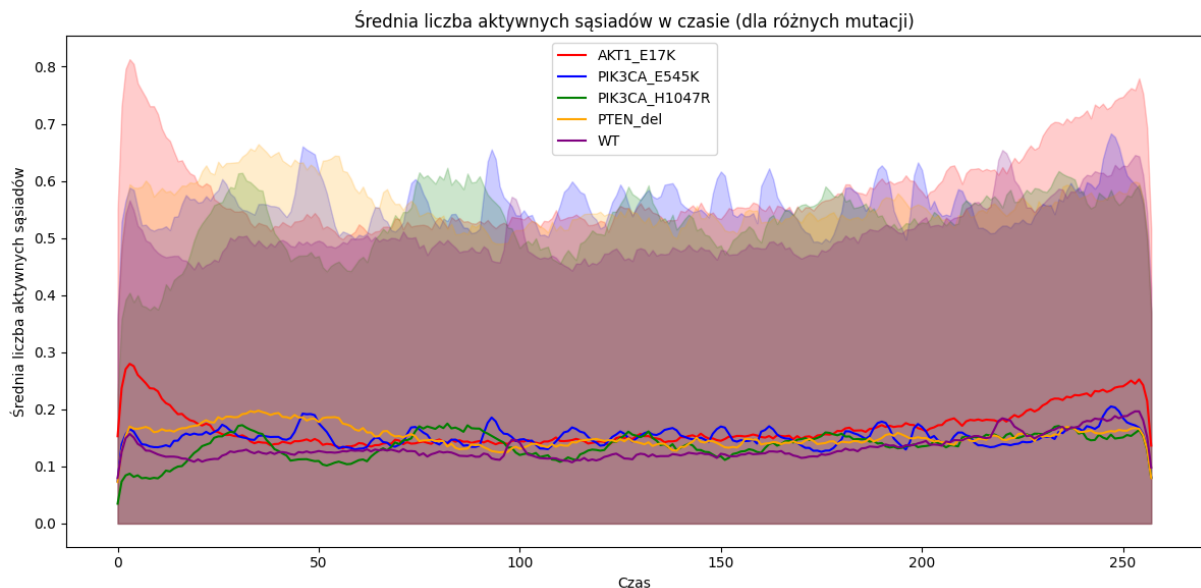


fig. 4) Średnia liczba aktywnych sąsiadów dla każdej mutacji w czasie wraz z odchyleniem standardowym. Wykres jest obcięty z dołu na wartości $y=0$, aby nie sugerować, że liczba sąsiadów może być ujemna, (pomimo że wartości średniej liczby sąsiadów minus odchylenie standardowe już mogą być ujemne). Kolorem czerwonym zaznaczona jest mutacja AKT1_E17K, niebieskim PIK3CA_E545K, zielonym PIK3CA_H1047R, żółtym PTEN_del, a fioletowym typ dziki (WT).

Wnioski

Wszystkie wykresy utrzymują się na mniej więcej tej samej wartości. Dla mutacji AKT1_E17K widać wyraźny skok na początku i na końcu. PIK3CA_E545K i PIK3CA_H1047R mają regularne wzrosty i spadki, dla PIK3CA_E545K są one częstsze. Dla sprawdzenia statystycznej istotności różnic w liczbie aktywnych sąsiadów w zależności od mutacji, przeprowadzono test Kruskala–Wallisa. Otrzymano p-wartość równą $4.575993976959647e-65$, co zdecydowanie jest poniżej progu $\alpha=0.05$. Różnice w wartościach są więc statystycznie istotne. Podsumowując, mutacje wpływają na liczbę pików i średnią liczbę aktywnych sąsiadów.

Analiza czasoprzestrzenna

Aby ocenić, czy aktywność sygnałowa ma charakter lokalny i czy może rozprzestrzeniać się w sposób skoordynowany w czasie i przestrzeni, wykonano wykresy na przestrzeni X-Y w klatkach czasowych równych Image_Metadata_T=[50, 100, 150, 200]. Zrobiono to dla eksperymentu 1 i jednego Site dla każdej z mutacji. Ze względu na pytanie badawcze postawione w Zadaniu 3 zostały także policzone komórki, które są aktywne stale w czasie (przynajmniej przez 10 punktów czasowych). Te zaznaczono na wykresach na zielono. Komórki „normalnie” aktywne zaznaczone są na czerwono, a komórki nieaktywne na szaro.

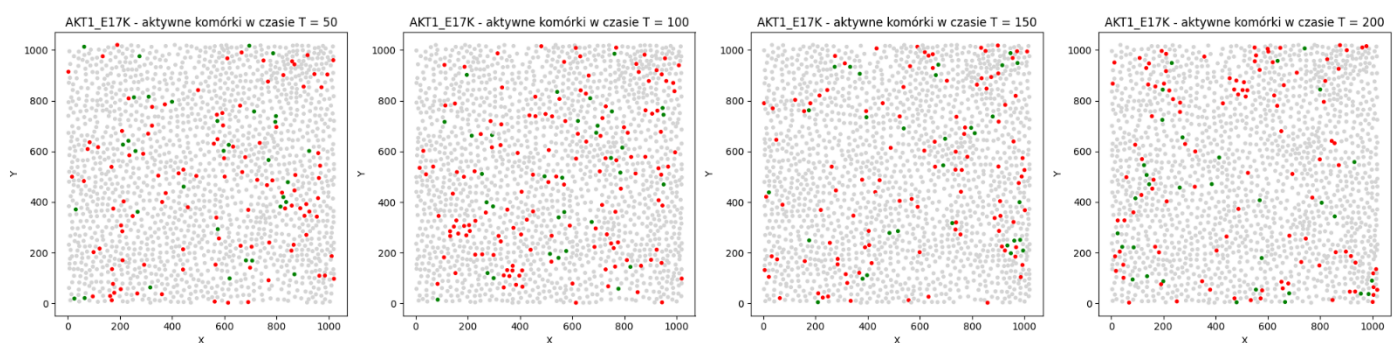


fig. 5) Wykresy płaszczyzny XY dla mutacji AKT1_E17K, (Exp_ID=1, Site=5) w punktach czasowych 50, 100, 150, 200. Komórki szare są nieaktywne, komórki czerwone aktywne, a zielone aktywne przez co najmniej 10 punktów czasowych.

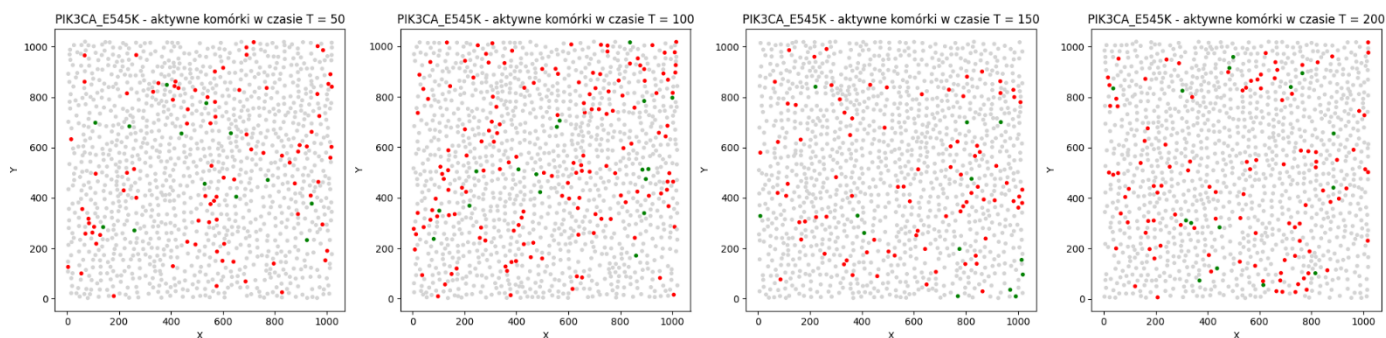


fig. 6) Wykresy płaszczyzny XY dla mutacji PK3CA_E54K, (Exp_ID=1, Site=9) w punktach czasowych 50, 100, 150, 200. Komórki szare są nieaktywne, komórki czerwone aktywne, a zielone aktywne przez co najmniej 10 punktów czasowych.

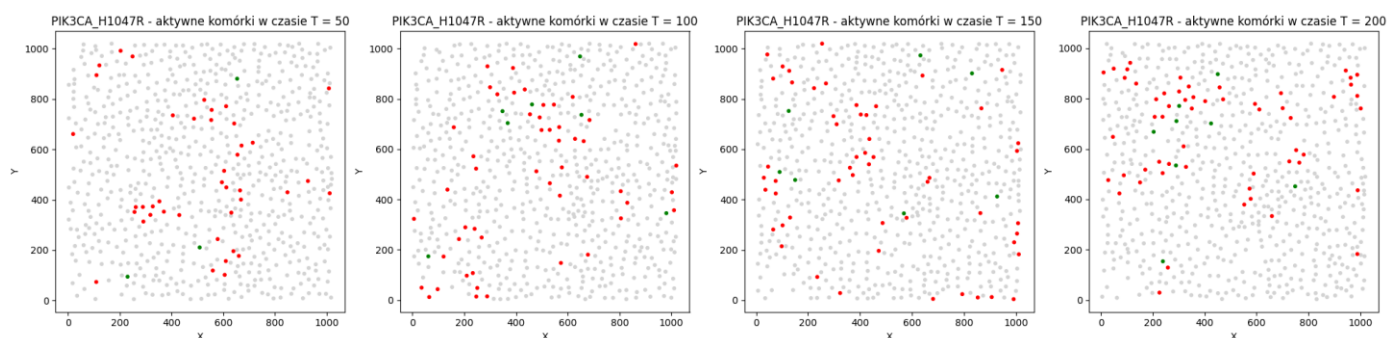


fig. 7) Wykresy płaszczyzny XY dla mutacji PK3CA_H1047R, (Exp_ID=1, Site=13) w punktach czasowych 50, 100, 150, 200. Komórki szare są nieaktywne, komórki czerwone aktywne, a zielone aktywne przez co najmniej 10 punktów czasowych.

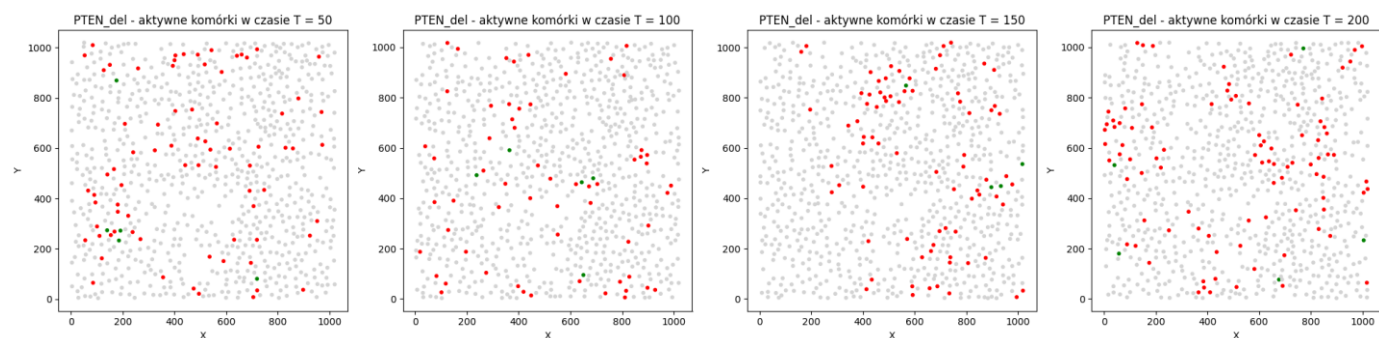


fig. 8) Wykresy płaszczyzny XY dla mutacji PTEN_del, (Exp_ID=1, Site=17) w punktach czasowych 50, 100, 150, 200. Komórki szare są nieaktywne, komórki czerwone aktywne, a zielone aktywne przez co najmniej 10 punktów czasowych.

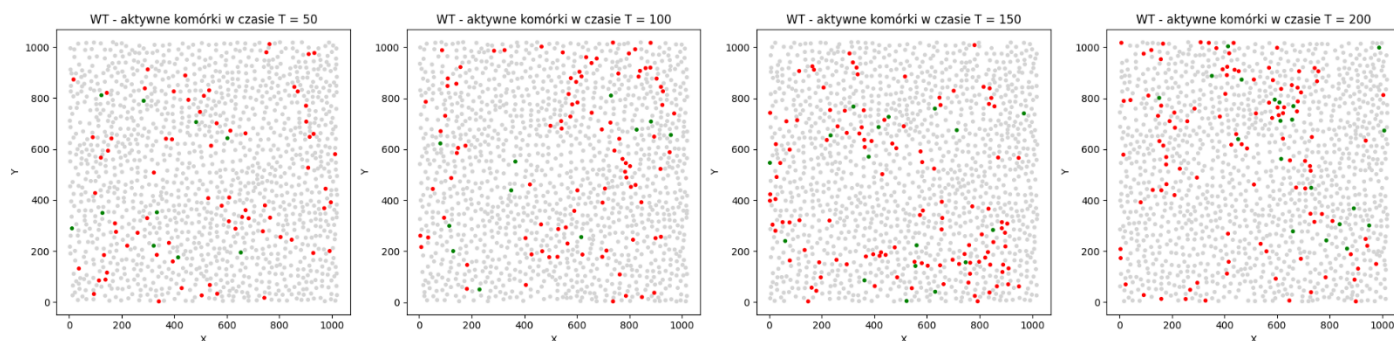


fig. 9) Wykresy płaszczyzny XY dla typu WT, (Exp_ID=1, Site=1) w punktach czasowych 50, 100, 150, 200. Komórki szare są nieaktywne, komórki czerwone aktywne, a zielone aktywne przez co najmniej 10 punktów czasowych.

Ze względu na to, że pojedyncze klatki nie oddają ruchu i ewentualnej propagacji w czasie, stworzono również gify dla punktów czasowych 50-130 dla każdej mutacji osobno, eksperymentu 1, jednego Site (tych samych, co przy figurach 5-9). Są one w plikach o nazwie `activity_animation_{mutacja}.gif`.

Wnioski

Na wykresach pojedynczych klatek można zauważyć, że dla mutacji PK3CA_H1047R i PTEN_del, komórki są rozmieszczone najrzadziej. Nas interesuje jednak to, czy tworzą się jakieś skupiska komórek aktywnych. Nie widać tego w sposób jednoznaczny, jednakże są miejsca w przestrzeni, gdzie zagęszczenie komórek aktywnych jest większe niż w innych oraz takie miejsca, gdzie komórek aktywnych nie ma wcale. Można więc przypuszczać, że istnieje lokalność w aktywności sygnałowej, jednak jest ona złożona i nie występuje zawsze.

Na podstawie gifów można wnioskować, że aktywność propaguje się w pewien falowy sposób. Da się zauważyć aktywowanie kolejnych komórek będących blisko siebie. Jednakże, dalej nie jest to bardzo regularne i często trudno wyznaczyć jednoznaczne ognisko i kierunek rozchodzenia się fal. Można jednak zakładać, że zachodzi propagacja sygnału.

Zadanie 3

Pytanie badawcze: Czy liczba komórek długotrwale aktywnych zależy od mutacji i czy w ich otoczeniu wartości wskaźnika ERKTR_ratio rozkładają się inaczej?

Definicja komórek długotrwale aktywnych: Takie komórki, które są aktywne przez przynajmniej 10 kroków czasowych na raz.

Analiza

Aby spróbować odpowiedzieć na pytanie badawcze, najpierw policzono (przy pomocy KDTree) średnią wartość wskaźnika ERKTR_ratio dla sąsiadów komórki (oraz samej siebie) w 1 klatce czasowej, a następnie oznaczono, które komórki są długotrwale aktywne. W celu sprawdzenia, czy liczba komórek długotrwale aktywnych różni się pomiędzy mutacjami, przygotowano wykres przedstawiający średnią liczbę komórek długotrwale aktywnych wśród wszystkich aktywnych w funkcji czasu.

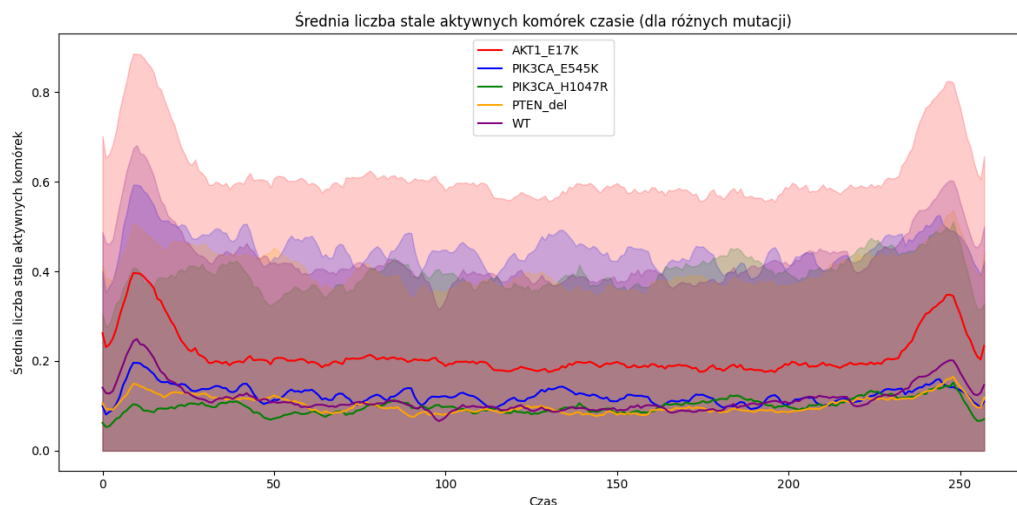


fig. 10) Średnia liczba komórek długotrwale aktywnych wśród komórek aktywnych wraz z odchyleniem standardowym. Wykres jest obcięty na wartości $y=0$, aby nie sugerować, że liczba komórek może być ujemna. Kolorem czerwonym zaznaczona jest mutacja AKT1_E17K, niebieskim PIK3CA_E545K, zielonym PIK3CA_H1047R, żółtym PTEN_del, a fioletowym typ dziki (WT).

Na wykresie widać, że dla mutacji AKT1_E17K wartości są zdecydowanie najwyższe. Wydaje się, że w stosunku do typu dzikiego wzrost występuje także dla mutacji PIK3CA_E545K. Aby potwierdzić te obserwacje, przeprowadzono test Mann-Whitneya w czasie 60-180. Porównywano tam typ WT z poszczególnymi mutacjami. Wyniki testu są w tabeli poniżej:

Mutacja	p_raw	U	Zmiana	p_bonferroni	Istotne (a=0.05)
AKT1_E17K	3.360343522054828e-41	14641.0	wzrost	1.3441374088219311e-40	PRAWDA
PIK3CA_E545K	2.021179929071789e-35	14084.0	wzrost	8.084719716287156e-35	PRAWDA
PIK3CA_H1047R	0.03064777467708333	6143.0	spadek	0.12259109870833332	FAŁSZ
PTEN_del	1.247578004636456e-10	3817.0	spadek	4.990312018545824e-10	PRAWDA

Tab. 2) Wyniki testów Manna–Whitneya dla komórek stale aktywnych. P_raw to surowa p-wartość, U to wartość statystyki U Manna–Whitneya, Zmiana to wpływ mutacji na liczbę komórek długotrwale aktywnych oszacowana przez średnią w całym przedziale czasowym, p_bonferroni to poprawiona p-wartość korektą Bonferroniego, kolumna Istotne mówi czy różnica jest istotna statystycznie z progiem $\alpha=0.05$.

Z tabeli można odczytać, że wszystkie mutacje poza PIK3CA_H1047R statystycznie różnią się od typu dzikiego. Tak jak wynika z wykresu, wzrost nastąpił dla mutacji AKT1_E17K i PIK3CA_E545K, a dla PTEN_del obserwujemy spadek.

Aby sprawdzić, czy rozkład wartości ERKKTR_ratio jest inny wokół komórek długotrwale aktywnych, wykonano dwa boxploty. Pierwszy prezentuje średnie wartości współczynnika ERKKTR_ratio osobno dla każdej mutacji i z podziałem na otoczenie komórek stale aktywnych i reszty komórek. Drugi wykres przedstawia odchylenia standardowych tych średnich.

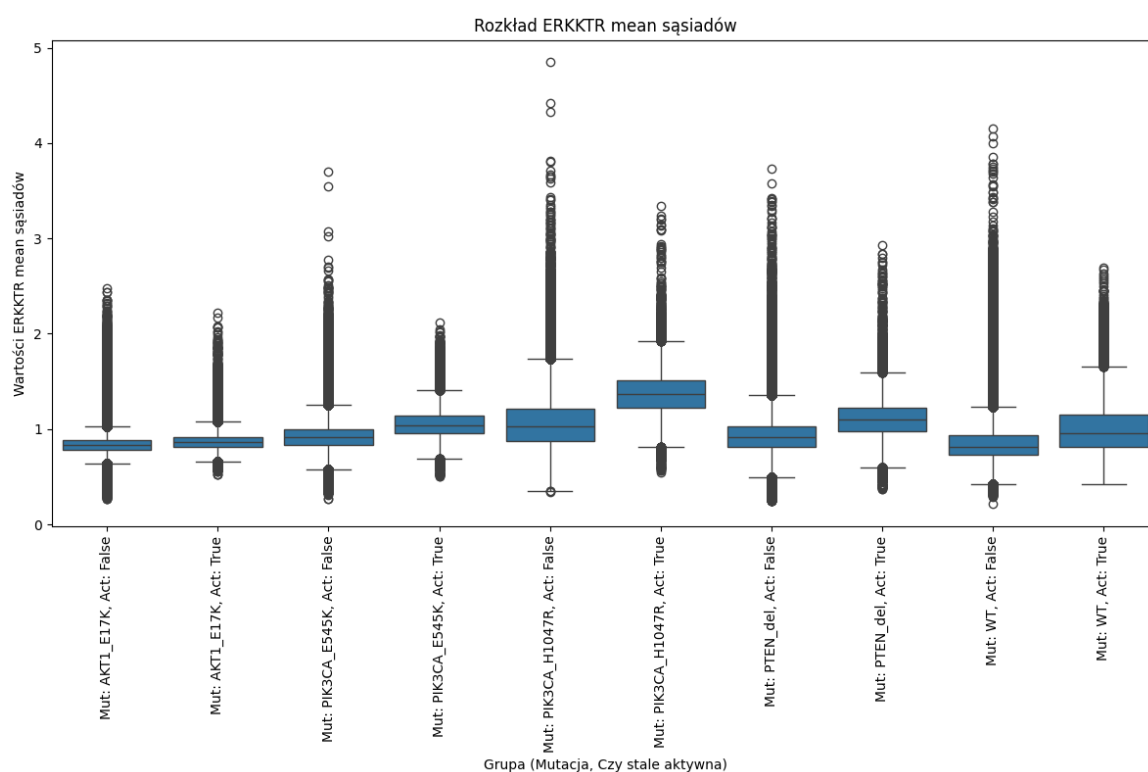


fig. 11) Rozkład średnich wartości ERK z podziałem na mutacje i informację, czy komórki są długotrwale aktywne.

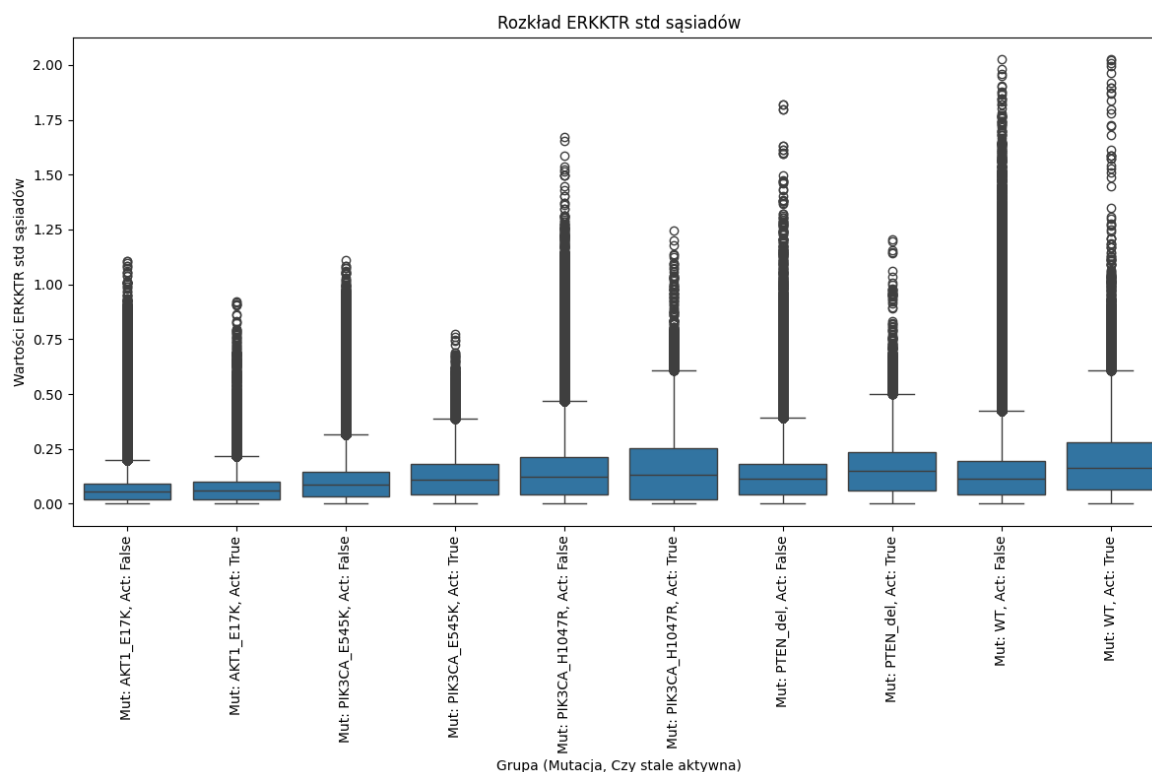


fig. 12) Rozkład odchyień standardowych wartości ERK z podziałem na mutacje informację, czy komórki są długotrwale aktywne.

Na wykresach widać, że średnie wartości ERK dla otoczenia komórek stale aktywnych są wyższe niż dla pozostałych komórek. Ponadto, w grupach komórek stale aktywnych obserwowaliśmy wyższe odchylenia standardowe, co sugeruje zwiększone zróżnicowanie odpowiedzi w mikrośrodowisku. Dla okolic komórek stale aktywnych pojawia się także mniej wartości odstających.

Dla statystycznego sprawdzenia różnic pomiędzy komórkami aktywowanymi długotrwale a pozostałymi przeprowadzono test Manna-Whitneya dla średnich i odchyień standardowych. Wyniki są w tabelach poniżej:

Mutacja	p_raw	U	Zmiana	p_bonferroni	Istotne (α=0.05)
AKT1_E17K	0.0	163518593419.0	wzrost	0.0	PRAWDA
PIK3CA_E545K	0.0	124968551478.5	wzrost	0.0	PRAWDA
PIK3CA_H1047R	0.0	34093980098.5	wzrost	0.0	PRAWDA
PTEN_del	0.0	41642476421.0	wzrost	0.0	PRAWDA
WT	0.0	70512971000.5	wzrost	0.0	PRAWDA

Tab. 3) Wyniki testów Manna-Whitneya dla **średnich** komórek stale aktywnych i pozostałych. P_raw to surowa p-wartość, U to wartość statystyki U Manna-Whitneya, Zmiana to wpływ otoczenia komórki stale aktywnej na ERK oszacowana przez średnią w całym przedziale czasowym, p_bonferroni to poprawiona p-wartość korektą Bonferroniego, kolumna Istotne mówi czy różnica jest istotna statystycznie z progiem α=0.05

Mutacja	p_raw	U	Zmiana	p_bonferroni	Istotne (a=0.05)
	2.7070290911353782e-			1.353514545567689e-	
AKT1_E17K	15	134473950420.0	wzrost	14	PRAWDA
PIK3CA_E545K	0.0	91725531062.5	wzrost	0.0	PRAWDA
	5.907969323575404e-			2.953984661787702e-	
PIK3CA_H1047R	19	21260337807.0	wzrost	18	PRAWDA
PTEN_del	0.0	31764198397.0	wzrost	0.0	PRAWDA
WT	0.0	57778802073.5	wzrost	0.0	PRAWDA

Tab. 4) Wyniki testów Manna–Whitneya dla **odchyłeń standardowych** komórek stale aktywnych i pozostałych. P_raw to surowa p-wartość, U to wartość statystyki U Manna–Whitneya, Zmiana to wpływ otoczenia komórki stale aktywnej na odchylenie standardowe ERK oszacowane przez średnią w całym przedziale czasowym, p_bonferroni to poprawiona p-wartość korektą Bonferroniego, kolumna Istotne mówi czy różnica jest istotna statystycznie z progiem $\alpha=0.05$.

Dla wszystkich mutacji i typu dzikiego nastąpił statystycznie istotny wzrost średnich i odchyłeń standardowych pomiędzy sąsiedztwem komórek stale aktywnych i pozostałych komórek, co potwierdza wnioski wyciągnięte z boxplotów.

Wnioski:

Liczba komórek stale aktywnych zależy od mutacji (szczególnie mutacja AKT1_E17K wykazuje wzrost w stosunku do WT). Większa długotrwała aktywność może znaczyć, że jest wysoka aktywność ale mała zmienność tej aktywności- komórki są po prostu cały czas mocno aktywne. Mała długotrwała aktywność sugeruje albo niską aktywność w sumie, albo częste zmiany w aktywności, w taki sposób, że mało komórek może być długo aktywnych.

Zaobserwowaliśmy, że stała aktywność komórek wpływa na ich mikrośrodowisko. Wokół nich średnie wartości ERK są wyższe, co może to sugerować tworzenie się skupisk, w których aktywność jest wysoka. Również odchylenie standardowe wzrosło dla okolic komórek stale aktywnych. Wskazuje to, że stała aktywacja sygnału może prowadzić do bardziej zróżnicowanych zachowań sąsiadujących komórek. Wysoka liczba wartości odstających dla komórek nieaktywnych długotrwale sugeruje, że w wielu przypadkach są pojedyncze komórki lub grupy komórek o nietypowo wysokim lub niskim poziomie ERK. Mogą to być na przykład lokalne nisze aktywności sygnałowej.