

AUTOREFERAT

OPIS DOROBKU I OSIĄGNIĘĆ NAUKOWYCH

Dr n. med. Paweł Karpiński
Katedra i Zakład Genetyki
Uniwersytet Medyczny we Wrocławiu
ul. Marcinkowskiego 1, 50-368 Wrocław
E-mail: pawel.karpinski@umw.edu.pl

Wrocław, maj 2023

1. Imię i nazwisko: Paweł Karpiński

2. Posiadane dyplomy, stopnie naukowe/ artystyczne – z podaniem nazwy, miejsca i roku ich uzyskania oraz tytułu rozprawy doktorskiej

2010 Doktor nauk medycznych – Wydział Lekarski, Akademia Medyczna we Wrocławiu (obecnie Uniwersytet Medyczny we Wrocławiu)

Tytuł rozprawy doktorskiej: Rola niestabilności metylacyjnej DNA w etiologii sporadycznego raka jelita grubego

Promotor: prof. dr hab. n. med. Maria Małgorzata Sasiadek

2004 Magister Mikrobiologii - Uniwersytet Wrocławski, kierunek Biologia, specjalizacja Mikrobiologia

2002 Licencjat Mikrobiologii - Uniwersytet Wrocławski, kierunek Biologia,

3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych

2005-2009– studia doktoranckie, Katedra i Zakład Genetyki, Akademia Medyczna we Wrocławiu (obecnie Uniwersytet Medyczny we Wrocławiu)

10.2009- obecnie – adiunkt, Katedra i Zakład Genetyki, Uniwersytet Medyczny we Wrocławiu

2019-2021 – adiunkt, Laboratorium Genomiki i Bioinformatyki, Instytut Immunologii i Terapii Doświadczalnej im. Ludwika Hirszfelda Polskiej Akademii Nauk we Wrocławiu

2023 – obecnie, specjalista bioinformatyk, Laboratorium Genomiki i Bioinformatyki, Instytut Immunologii i Terapii Doświadczalnej im. Ludwika Hirszfelda Polskiej Akademii Nauk we Wrocławiu

4. Wskazanie osiągnięcia* wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. 2016 r. poz. 882 ze zm. w Dz. U. z 2016 r. poz. 1311.)

Osiągnięcie naukowe stanowi cykl publikacji składających się z 6 prac oryginalnych i 1 przeglądowej o łącznej punktacji: IF: **30,942**; Pkt. MNiSW/KBN: **440**.

a) Tytuł osiągnięcia naukowego:

Zastosowanie metod bioinformatycznych w badaniach chorób o heterogennej etiologii molekularnej

b) (autor/autorzy, tytuł/tytuły publikacji, rok wydania, nazwa wydawnictwa, recenzenci wydawniczy)

1. Reduced number of peripheral natural killer cells in schizophrenia but not in bipolar disorder. **PAWEŁ KARPIŃSKI**, DOROTA FRYDECKA, MARIA M. SĄSIADK, BŁAŻEJ MISIAK. Brain Behav.Immun. 2016 Vol.54 s.194-200.

IF: 5,964 / IF_{5-letni}: 6,458

Pkt. MNiSW: 40.000

Liczba cytowań wg WoS:25

udział własny: 60%

2. Further evidence for depletion of peripheral blood natural killer cells in patients with schizophrenia: a computational deconvolution study. **PAWEŁ KARPIŃSKI**, JERZY SAMOCHOWIEC, DOROTA FRYDECKA, MARIA M. SĄSIADK, BŁAŻEJ MISIAK. Schizophr.Res. 2018 Vol.201 s.243-248.

IF: 4,569 / IF_{5-letni}: 3,597

Pkt. MNiSW: 40.000

Liczba cytowań wg WoS:11

udział własny: 60%

3. Immunological landscape of consensus clusters in colorectal cancer. **PAWEŁ KARPIŃSKI**, JOANNA ROSSOWSKA, MARIA M. SĄSIADK. Oncotarget, 2017, vol. 8, nr 62, s. 105299-105311

IF: 0 / IF_{5-letni}: 0 **

Pkt. MNiSW: 40

Liczba cytowań wg WoS:33

udział własny: 80%

** IF czasopisma *Oncotarget* wynosił 5,168 w 2017 roku. W styczniu 2018 czasopismo zostało usunięte, bez podania wyraźnej przyczyny, z listy Web of Science.

4. Major histocompatibility complex genes as therapeutic opportunity for immune cold molecular cancer subtypes. **PAWEŁ KARPIŃSKI**, ŁUKASZ ŁACZMAŃSKI, MARIA M. SĄSIADK. *J.Immunol.Res.* 2020 Vol.2020 art.8758090 [9 s.]

IF: 4,818 / IF_{5-letni}: 4,261

Pkt. MNiSW: 140

Liczba cytowań wg WoS:2

udział własny: 80%

5. Pan-cancer analysis reveals presence of pronounced DNA methylation drift in CpG island methylator phenotype clusters. **PAWEŁ KARPIŃSKI**, KAROLINA PESZ, MARIA M. SĄSIADK. *Epigenomics* 2017 Vol.9 no.11 s.1341-1352.

IF: 4,979 / IF_{5-letni}: 4,061

Pkt. MNiSW: 40.000

Liczba cytowań wg WoS:12

udział własny: 80%

6. Multilevel omic data clustering reveals variable contribution of methylator phenotype to integrative cancer subtypes. **PAWEŁ KARPIŃSKI**, ARPAD V. PATAI, WOJCIECH HAP, WOJCIECH KIELAN, IZABELA ŁACZMAŃSKA, MARIA MAŁGORZATA SĄSIADK. *Epigenomics* 2018 Vol.10 no.10 s.1289-1299.

IF: 4,404 / IF_{5-letni}: 4,061

Pkt. MNiSW: 40.000

Liczba cytowań wg WoS:2

udział własny: 80%

7. CpG island methylator phenotype - a hope for the future or a road to nowhere?. **PAWEŁ KARPIŃSKI**, MARIA MAŁGORZATA SĄSIADK. *Int.J.Mol.Sci.* 2022 Vol.23 no.2 art.830 [21 s.], DOI: 10.3390/ijms23020830

IF: 6,208 / IF_{5-letni}: 6,628

Pkt. MNiSW: 140

Liczba cytowań wg WoS:4

udział własny: 75%

c) omówienie celu naukowego/artystycznego ww. pracy/prac i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania

W swojej pracy naukowej od lat zajmuję się badaniami nad genetyczną i epigenetyczną etiologią schorzeń charakteryzujących się heterogennością zaburzeń molekularnych, z wykorzystaniem różnych, biostatystycznych metod.

Cykl publikacji stanowiący osiągnięcie naukowe poświęcony jest badaniom transkryptomiki i epigenetyki w schorzeniach człowieka, charakteryzujących się wysoką heterogennością molekularną, takich jak schizofrenia i nowotwory, których wyniki analizowałem wykorzystując starannie dobrane do analizowanego zagadnienia, metody obliczeniowe.

Wyniki prowadzonych przeze mnie i moich współpracowników badań nad mechanizmami epigenetycznymi w etiologii nowotworów skłoniły mnie do zastanowienia się nad pytaniem, czy te mechanizmy mają rzeczywiście tak istotne znaczenie, jak jest im przypisywane i tym samym do podjęcia analizy istniejącej literatury pod kątem znaczenia wyników badań nad fenotypem metylatora, uzyskanych przez różnych autorów. W cyklu zamieściłem więc również publikację przeglądową, będącą wynikiem tych dociekań.

Wstęp

Obecnie obserwuje się bardzo szybki rozwój metod badań genomu i transkryptomu, które przybliżają nas do zrozumienia mechanizmów odpowiedzialnych za powstawanie i/lub przebieg wielu schorzeń człowieka [1].

Metody molekularne pozwalają obecnie na jednoczesną analizę, w skali genomowej, ogromnej liczby różnorodnych danych, takich jak występowanie nie zrównoważonych addycji i delecji chromosomowych (w tym mikro-zmian), identyfikację wariantów nukleotydowych w DNA (o różnym znaczeniu klinicznym, od zmian patogennych, poprzez łagodne, do zmian o nieznanej patogenności), określenie ekspresji lub metylacji kilkudziesięciu tysięcy genów jednocześnie, ekspresji wszystkich znanych genów kodujących białka i długich, niekodujących RNA (*long-noncoding RNA*; lncRNA) w wielu tysiącach pojedynczych komórek jednocześnie (*single cell RNA-seq*) [2].

Rozwój metod molekularnych zaowocował pojawieniem się ogromnej liczby danych genomowych gromadzonych i udostępnianych w repozytoriach takich jak np. Gene Expression Omnibus (GEO) zawierającym dane w najróżniejszych kontekstach medycznych lub Genomic Data Commons (GDC) specjalizującym się w chorobach nowotworowych [3, 4]. Ta dostępność danych molekularnych jest przyczyną znacznego przyspieszenia rozwoju bioinformatyki, które przejawia się intensywnym rozwojem i dostępnością algorytmów umożliwiających analizę wysoce zróżnicowanych danych molekularnych.

Coraz bardziej szczegółowa i wielokierunkowa analiza genetyczna różnych schorzeń człowieka zaowocowała odkryciem zjawiska zmienności (heterogenności) genetycznej badanych tkanek. Heterogenność genetyczna *sensu stricto* oznacza, że to samo schorzenie u różnych pacjentów może mieć różnorodną etiologię genetyczną. Okazało się, że wiele chorób o takich samych objawach klinicznych jest spowodowanych mutacjami w różnych genach (heterogenność niealleliczna, jak np. polineuropatie, zespół dziedzicznego, zwiększonego ryzyka zachorowania na raka piersi lub/i jajnika) lub licznymi mutacjami w pojedynczym genie (heterogenność alleliczna, np. mukowiscydoza) [5].

Obok heterogenności genetycznej, jednym z najważniejszych czynników zakłócających analizę danych w skali genomowej jest wysoka heterogenność (zmienność/niejednorodność) analizowanych próbek. Jest to zjawisko wielowymiarowe dotyczące zarówno zróżnicowania kompozycji komórkowej próbek (np. różnych populacji limfocytów z których izolowane jest DNA, różnych komórek pobranych z pozornie jednolitej tkanki jak np. mieszanina komórek nowotworowych i prawidłowych w masie guza) jak i indywidualnego zróżnicowania patologicznych tkanek w pozornie jednolitej grupie pacjentów (odmienne podtypy molekularne nowotworów klasyfikowanych histopatologicznie do tej samej grupy rozpoznania), a także odmienności technik stosowanych podczas przygotowania próbek do analizy [6].

Przykładem danych genomowych, w których spotkamy wszystkie wyżej wymienione typy heterogenności są dane pochodzące z guzów nowotworowych [7]. Wysoka heterogenność guzów nowotworowych powstaje w wyniku niejednorodnej akumulacji zmian genetycznych i epigenetycznych w komórkach nowotworowych oraz zmienności wzajemnych proporcji komórek nowotworowych i prawidłowych (tzw. komórek mikrośrodowiska) w guzach. W związku z tym próbki pobrane z różnych regionów tego samego nowotworu mają najczęściej różny skład komórkowy. Zjawisko takie nazywamy heterogennością w obrębie guza (*intra-tumor heterogeneity*) [8].

Większość danych genomowych nowotworów, dostępnych w bazach jest wygenerowane z relatywnie dużego, mrożonego wycinka guza (tzw. *tumor bulk*), któremu najczęściej nie towarzyszy ocena proporcji między komórkami nowotworowymi i prawidłowymi (m.in.: komórki podścieliska i komórki nacieku zapalnego). Tymczasem wzajemne proporcje komórek prawidłowych/nowotworowych w badanej próbce mają ważki wpływ na wyniki np. przykład, profilu ekspresji genów czy metylacji DNA. Problem niepełnej wiarygodności tak uzyskanych wyników zaczął być zauważany około 2009 roku i od tego czasu opracowano dziesiątki algorytmów, które umożliwiają tzw. dekonwolucję czyli rozplot sygnału ekspresji lub metylacji [9]. Metody te opierają się na założeniu, że profil ekspresji lub metylacji genów mierzony z wycinka tkanki jest mieszaniną profili ekspresji lub metylacji genów pochodzących z mieszaniny różnych komórek. Najskuteczniejszy rozplot sygnału jest możliwy dzięki znajomości genów markerowych, które ulegają specyficznej ekspresji/metylacji w specyficznych rodzajach komórek. Dużym ograniczeniem przy tym typie rozplotu jest mała dostępność profili ekspresji/metylacji pochodzących z czystych linii komórkowych [10].

Kolejną przyczyną niepełnej wiarygodności wyników badań w skali genomowej jest fakt, że guzy klasyfikowane jako ten sam typ histopatologiczny, cechują się znaczną odmiennością molekularną. Ta zmienność molekularna stała się obecnie podstawą do opracowania klasyfikacji molekularnej lub histopatologiczno-molekularnej, dzielącej guzy kwalifikowane histopatologicznie do określonej grupy, na podtypy molekularne, jak np. guzy jelita grubego dzielone na podgrupy konsensusowe [11]. Zjawisko to określamy heterogennością pomiędzy guzami (*Inter-Tumor Heterogeneity*). Analizę danych wykazujących ten typ heterogenności umożliwia analiza skupień (*cluster analysis*), której celem jest znalezienie/wyróżnienie takich grup (klastrów), które wykazują wysoką spójność wewnętrzną i są jak najbardziej zróżnicowane pomiędzy sobą [12]. Tak przeprowadzona analiza skupień umożliwia wyodrębnienie w badanej chorobie podgrup molekularnych opartych np. o analizę ekspresji genów (mRNA) w skali genomowej. Tym samym pozwala na precyzyjną charakterystykę oraz na ustalenie różnic pomiędzy odrębnymi podgrupami. Analiza skupień odegrała i odgrywa bardzo ważną rolę w identyfikacji podtypów molekularnych różnych typów nowotworów [13].

Ponieważ analiza skupień rozwija się od kilku dekad, opracowano w tym czasie wiele różnych metod realizujących to zadanie, jednak w ostatnich latach obserwuje się dwojaki rozwój podejść do klastrowania. Pierwsze to agregowanie analizy skupień (*ensemble clustering*), którego ideą jest łączenie wyników uzyskanych przy pomocy różnych metod

klastrowania oraz metod uzyskiwania optymalnych klastrów. Podejście to ma na celu usunięcie błędów wynikających ze specyficznych cech metod klastrowujących [14]. Drugie podejście to integracyjna analiza skupień (*integrative clustering*), która odbywa się w oparciu o zbiór zmiennych pochodzący z co najmniej dwóch „warstw” molekularnych tej samej próbki np. mRNA i metylacji DNA [15]. Integracyjna analiza skupień pozwala nie tylko na precyzyjne określenie podtypów molekularnych różnych typów nowotworów ale także na ocenę interakcji pomiędzy różnymi zmianami molekularnymi co ma ważne implikacje dla badań nad rozwojem i progresją chorób nowotworowych.

Na obecność niebiologicznej zmienności mającej techniczne pochodzenie (*technical variability*) zwrócono uwagę na początku XXI wieku wraz z pojawieniem się technik opartych o mikromacierze [16]. Jest obecna w eksperymentach opierających się na użyciu jednocześnie wielu mikroczipów lub komórek przepływowych (w przypadku RNAseq), dlatego też określana jest czasem mianem efektu serii (*batch effect*). Badania wykazały, że zmienność techniczna może mieć znaczny wpływ na wyniki eksperymentu. Jak wykazano, najskuteczniejsze niwelowanie wpływu tej zmienności na uzyskiwane wyniki opiera się na projektowaniu eksperymentu w taki sposób, aby poszczególne fenotypy (np. tkanka zdrowa/tkanka zmieniona patologicznie) były równomiernie rozdystrybuowane na mikroczipach/komórkach przepływowych. Istnieją także techniki biosatystyczne mające na celu korektę danych, tak by zminimalizować wpływ tej zmienności [6].

Celami poniższej rozprawy habilitacyjnej są:

- 1) Wykorzystanie metod rozplotu (dekownolucji) profili ekspresji lub metylacji genów we krwi obwodowej do oceny składu komórkowego krwi w chorobach związanych z zaburzeniami psychicznymi.
- 2) Wykorzystanie metod rozplotu (dekownolucji) profili metylacji DNA i profili ekspresji wybranych genów oraz agregowanie analizy skupień (*ensemble clustering*) do oceny immunogenności w ujęciu *Pan-Cancer* (wielu typów nowotworów).
- 3) Badanie przyczyn, heterogenności oraz znaczenia fenotypu metylatora wsyp CpG (CIMP) jako ekstremalnego przykładu zaburzeń metylacji DNA w ujęciu *Pan-Cancer* (wielu typów nowotworów) przy użyciu metod opartych o agregowanie analizy skupień (*ensemble clustering*) oraz integracyjną analizę skupień (*integrative clustering*).

Omówienie wyników

Ad.1. Pierwszy artykuł stanowiący element rozprawy habilitacyjnej pt. „*Reduced number of peripheral natural killer cells in schizophrenia but not in bipolar disorder*” jest jedną z pierwszych publikacji, w której zająłem się wykorzystaniem metod rozplotu (dekownolucji) profili ekspresji genów we krwi obwodowej do oceny składu komórkowego krwi, z której izolowano mRNA do badań transkryptomu. Podstawą prowadzonych badań była immunologiczno-zapalna koncepcja powstawania schorzeń psychicznych, w szczególności schizofrenii. Liczne badania epidemiologiczne potwierdziły znacząco wysoką częstość występowania chorób autoimmunologicznych u pacjentów ze schizofrenią. Ponadto udowodniono, że prenatalne infekcje wirusem grypy, opryszczki czy wirusem cytomegalii zwiększają ryzyko wystąpienia schizofrenii u potomstwa. Podobnie, w przypadku choroby dwubiegunowej potwierdzono udział w etiologii schizofrenii zaburzeń neuroimmunologicznych, autoimmunologicznych oraz zaburzeń stężenia cytokin. Od wielu lat uważano, że pacjenci ze schizofrenią charakteryzują się względną przewagą limfocytów Th1 nad limfocytami Th2 co potwierdziła meta-analiza oparta o 378 pacjentów z SCZ i 358 osób zdrowych, opublikowana przez Guo i wsp. [17]. Natomiast szerszą perspektywę na poziomy limfocytów we krwi obwodowej u osób ze schizofrenią rzuciła meta-analiza opublikowana przez Miller i wsp. [18]. Autorzy stwierdzili znaczny wzrost odsetka limfocytów Th (CD4+) i komórek NK (CD56+) u pacjentów hospitalizowanych z ostrym nawrotem choroby. Ponadto meta-analiza opublikowana przez Miller et al. wykazała znaczące podwyższenie limfocytów T CD3+, limfocytów Th (CD4+) oraz stosunku limfocytów Th CD4+ do limfocytów T CD8+ u pacjentów ze schizofrenią. Ponadto wykazano znaczący wzrost poziomu limfocytów NK podczas leczenia przeciwpsychotycznego. Różne zmiany w proporcjach leukocytów we krwi były również stwierdzane u pacjentów z chorobą afektywną dwubiegunową w tym zwiększone proporcje monocytów, zmniejszone proporcje limfocytów T CD3+ i cytotoksycznych limfocytów T CD8+, wyższy odsetek aktywowanych limfocytów T CD4+CD25+, zmiany stosunku Th1 do Th2 oraz niższe procent komórek Treg [19].

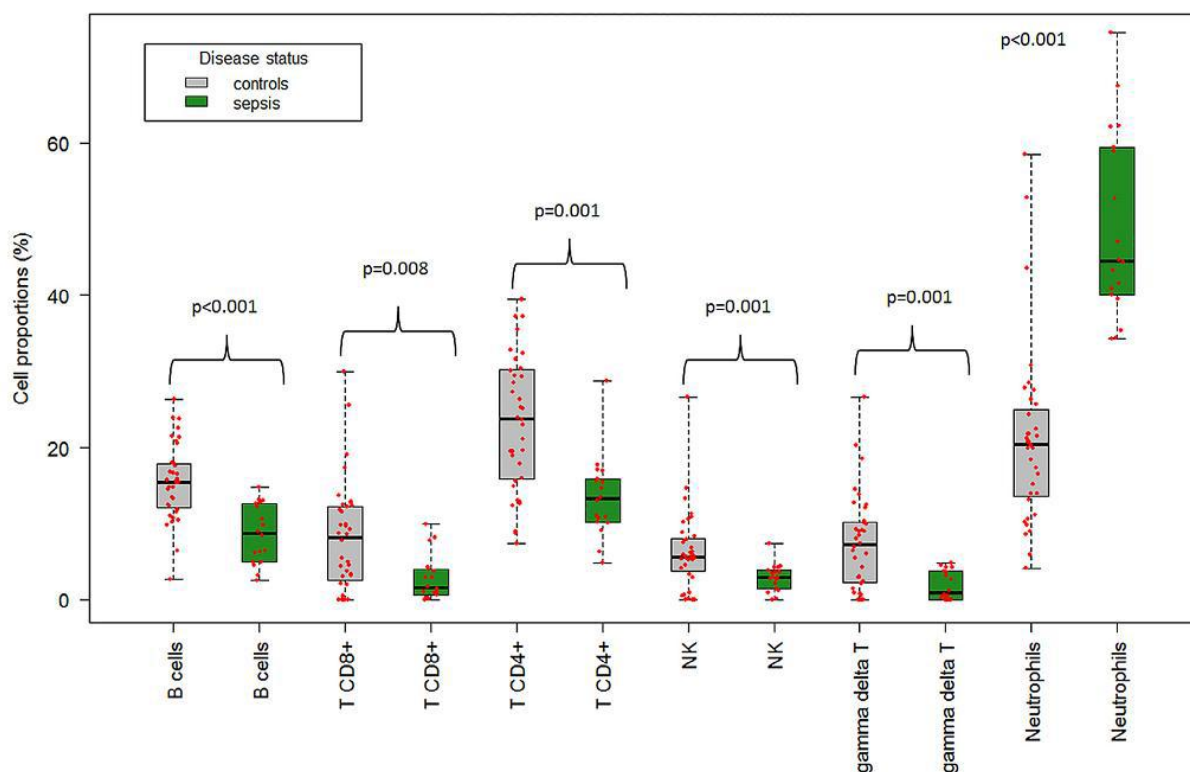
Należy podkreślić, że większość opublikowanych do tej pory pojedynczych badań była skoncentrowana na jednym, dwóch podtypach komórek krwi co powodowało brak obserwacji dla całego wachlarza podtypów komórek jednocześnie.

Celem pracy było wykorzystanie metod rozplotu (dekownolucji) profili ekspresji genów dla oceny proporcji komórek krwi i ich porównania pomiędzy osobami zdrowymi (n=218), pacjentami ze schizofrenią (n=199) oraz pacjentami z chorobą afektywną dwubiegunową (n=49). Z uwagi na dobrze opisane w literaturze zmiany wzajemnych proporcji komórek krwi do walidacji metody rozplotu użyto zestawu danych zawierających profile ekspresji pacjentów chorych na sepsę (n=18) i osób zdrowych (n=36) [20]. Wszystkie siedem zestawów danych zawierające profile ekspresji pochodziły z bazy Gene Expression Omnibus (GEO) i zostały otrzymane za pomocą technik macierzowych. Ponieważ analiza obejmowała kilka zestawów danych wygenerowanych w różnym czasie, różnymi technikami macierzowymi stwierdzono bardzo silny efekt serii (batch effect), który usunięto algorytmem ComBat [16]. Jako metody rozplotu użyto algorytm CIBERSORT do oceny wzajemnych proporcji następujących komórek: limfocytów B, limfocytów T CD4+, limfocytów T CD8+, limfocytów $\gamma\delta$ (gamma-delta), komórek NK oraz neutrofili [10]. Pod uwagę brano takie czynniki zakłócające analizę (*confounding factors*) jak wiek i płeć.

Za pomocą przyjętej metodologii wykazano znaczące statystycznie spadki limfocytów B, limfocytów T CD4+, limfocytów T CD8+, limfocytów $\gamma\delta$ i komórek NK oraz wzrost poziomu neutrofili u pacjentów z sepsą w porównaniu do osób z grupy kontrolnej (**Rycina 1**). Są to wyniki zgodne z opisywanymi w literaturze [21].

Pośród badanych komórek krwi tylko komórki NK wykazały statystycznie znaczące różnice pomiędzy grupą kontrolną a grupą pacjentów ze schizofrenią. We wszystkich zestawach danych stwierdzono znaczący statystycznie spadek proporcji NK w grupie pacjentów ze schizofrenią. U pacjentów z chorobą afektywną dwubiegunową nie stwierdzono znaczących statystycznie różnic w proporcjach wybranych komórek krwi w porównaniu do grupy kontrolnej.

Podsumowując, przedstawione powyżej wyniki były jednymi z pierwszych analiz wielu podgrup komórek krwi jednocześnie w kontekście zaburzeń psychicznych. Co więcej, prezentowane wyniki, mogą mieć znaczenie dla badań nad patofizjologią schizofrenii biorąc pod uwagę związek pomiędzy układem nerwowym i układem odpornościowym. W szczególności zaburzenia przekąźnictwa dopaminergicznego, obserwowane w schizofrenii, mogą mieć związek ze spadkiem poziomu komórek NK ponieważ receptory dopaminy ulegają ekspresji w krążących komórkach układu immunologicznego w tym w komórkach NK [22].



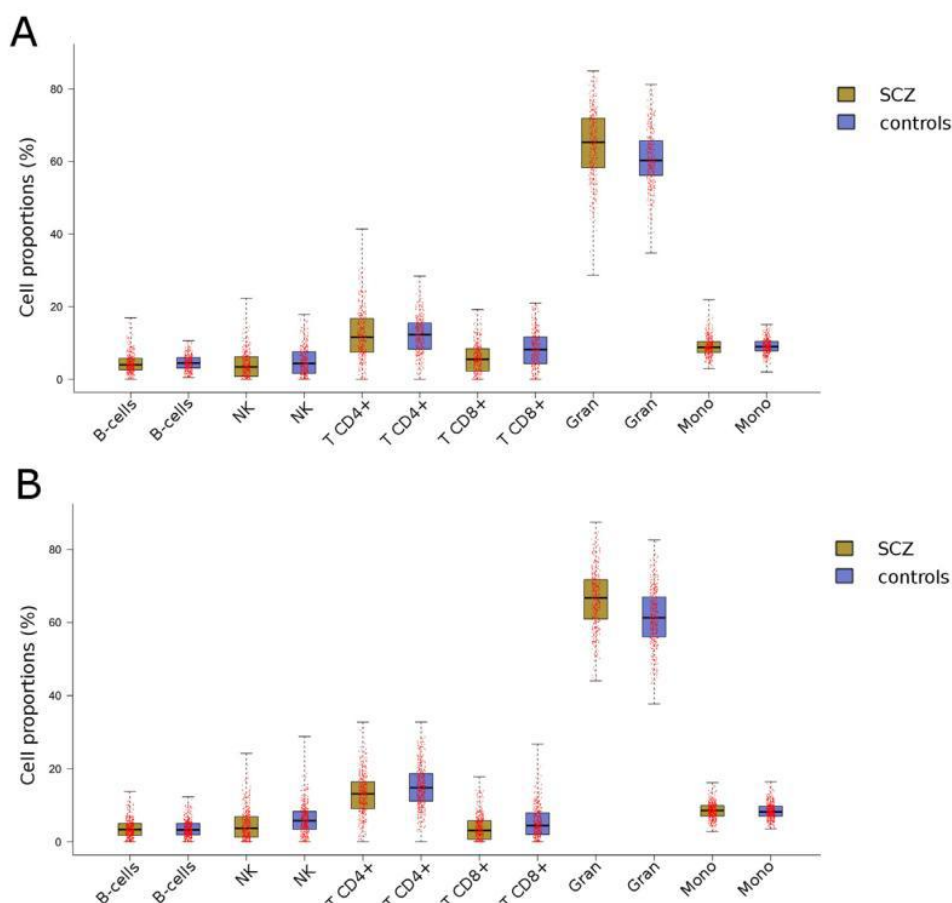
Rycina 1. Walidacja algorytmu CIBERSORT (Karpinski i wsp., 2016; [23]). Wykres pudełkowy przedstawiający różnice we wzajemnych proporcjach wybranych komórek krwi u pacjentów z sepsą (kolor zielony) i bez sepsy (kolor szary). Wzajemne proporcje zostały wyliczone na podstawie algorytmu CIBERSORT. Wartości p (obliczone testem Kruskala–Wallisa) wskazują znaczące statystycznie różnice pomiędzy badanymi grupami czyli znaczące spadki limfocytów B, limfocytów T CD4+, limfocytów T CD8+, limfocytów $\gamma\delta$ i komórek NK oraz wzrost poziomu neutrofilii u pacjentów z sepsą w porównaniu do osób z grupy kontrolnej.

Ad.2. Kolejna praca pt. „*Further evidence for depletion of peripheral blood natural killer cells in patients with schizophrenia: a computational deconvolution study*” była kontynuacją badań podjętych w pierwszej pracy. Celem pracy było wykorzystanie metod rozplotu (dekownolucji) profili metylacji DNA do oceny proporcji komórek krwi i ich porównania pomiędzy osobami zdrowymi (n=713), pacjentami ze schizofrenią (n=711) [24]. Wszystkie dwa zestawy danych zawierające profile metylacji DNA z krwi obwodowej pochodziły z bazy Gene Expression Omnibus (GEO) i zostały otrzymane za pomocą macierzy Infinium HumanMethylation450 BeadChip (Illumina) zawierającą 480 tysięcy sond. Podjęcie kontynuacji badań było motywowane kilkoma czynnikami: możliwością weryfikacji wyników otrzymanych na macierzach transkryptomicznych, możliwością wykonania badań na znacznie większych niż uprzednio grupach, możliwością zastosowania techniki dekonwolucji o znacznie większej rozdzielczości (w sensie liczby badanych markerów), która

była zoptymalizowana przez Teschendorff i wsp. na ~600 próbkach pod kątem wykrywania wzajemnych proporcji komórek we krwi obwodowej [25]. Jako metody rozplotu użyto metody RPC (*Robust Partial Correlations*) wbudowanej w algorytm EpiDISH do oceny wzajemnych proporcji następujących komórek: limfocytów B, limfocytów T CD4+, limfocytów T CD8+, komórek NK monocytów i granulocytów [25]. Otrzymane wyniki zostały skorygowane pod kątem występowania takich czynników zakłócających analizę jak wiek i płeć. Ponieważ analiza obejmowała dwa zestawy danych, dla których profile metylacji DNA wygenerowane zostały na kilkuset mikroczipach, wykryto bardzo silny efekt serii (*batch effect*), który został usunięty algorytmem ComBat [16].

Spośród badanych komórek krwi komórki NK i granulocyty wykazały statystycznie znaczące różnice pomiędzy grupą osób zdrowych a grupą pacjentów ze schizofrenią w obydwu zestawach danych jednocześnie (**Rycina 2**). W szczególności, pacjenci ze schizofrenią charakteryzowali się niższym odsetkiem komórek NK i oraz wyższym odsetkiem granulocytów.

Podsumowując, przedstawione powyżej wyniki potwierdziły nasze wcześniejsze rezultaty otrzymane na danych opartych o mikromacierze transkryptomiczne (**publikacja nr 1**, [23]) wskazujące na znacząco niższe proporcje komórek NK u pacjentów ze schizofrenią. Ponadto zmieniony odsetek komórek NK i granulocytów może wskazywać na zaburzenia odporności wrodzonej u osób ze schizofrenią na co wskazują także inne, niezależne doniesienia oparte o badania transkryptomiczne mózgu (*post mortem*) i krwi opublikowane przez Hess i wsp., 2016 [26].



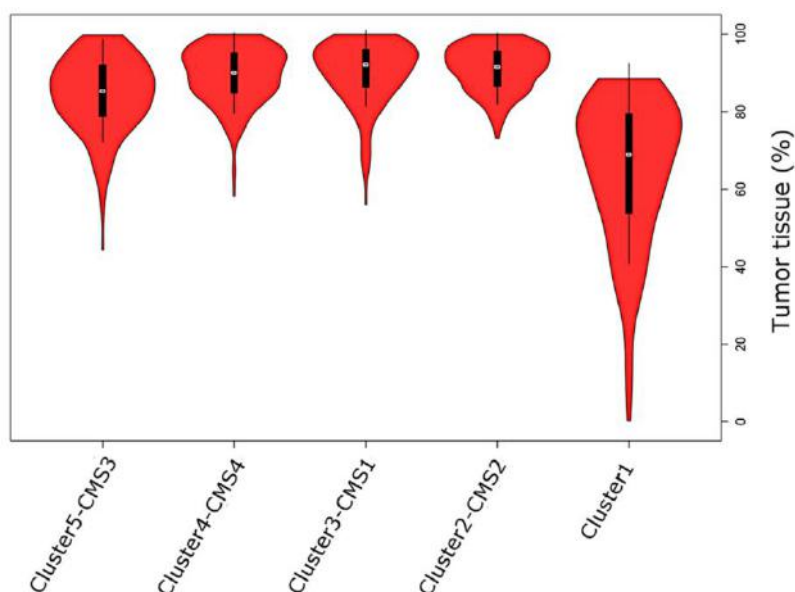
Rycina 2. Wyniki oceny wzajemnych proporcji wybranych komórek krwi wykonanych przy pomocy rozplotu metylacji DNA we krwi pełnej (Karpinski i wsp., 2017; [27]). Wykres pudełkowy przedstawiający różnice we wzajemnych proporcjach wybranych komórek krwi u pacjentów z schizofrenią (Scz, kolor ciemno-żółty) i osób zdrowych (controls, kolor niebieski). Wzajemne proporcje zostały wyliczone na podstawie algorytmu EpiDISH. **A)** Zestaw danych GSE80417 zawierający profile metylacji DNA dla 326 chorych na schizofrenię oraz 306 osób zdrowych. **B)** Zestaw danych GSE84727 zawierający profile metylacji DNA dla 385 chorych na schizofrenię oraz 407 osób zdrowych.

Ad3. W latach 2000-2010 wraz z rozwojem technik molekularnych o zasięgu genomowym takim jak macierze transkryptomiczne czy sekwencjonowanie RNA oraz technik biostatystycznych, przede wszystkim nienadzorowanego grupowania, rozwija się koncepcja choroby nowotworowej jako zespołu heterogennych chorób o różnym podłożu molekularnym (ang. *cancer molecular subtypes*) [28]. Zwieńczeniem kilkuletniego wyodrębnienia i systematyzacji typów molekularnych w raku jelita grubego jest klasyfikacja opublikowana przez Guinney i wsp. w 2015 roku, w której zostały wyróżnione podtypy: CMS1 — immunologiczny, CMS2 — kanoniczny, CMS3 — metaboliczny i CMS4 — mezenchymalny [11]. Podtypy CMS są znacząco zróżnicowane pod względem cech klinicznych oraz zmian molekularnych. Podtypy CMS, w badaniach naukowych, są do dnia dzisiejszego wiodącą

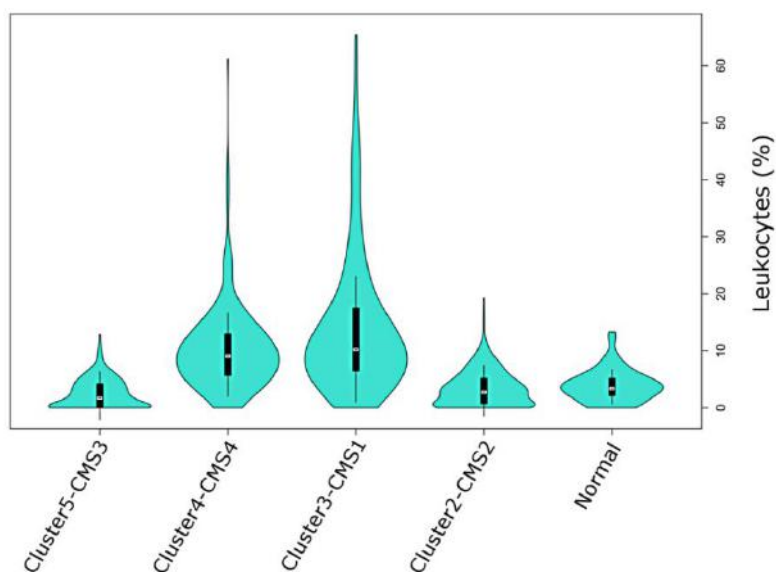
klasyfikacją molekularną raka jelita grubego. Równolegle do badań nad podtypami molekularnymi w różnych typach raków rozwija się zainteresowanie charakterystyką mikrośrodowiska guzów nowotworowych w szczególności frakcją komórek immunologicznych. Artykuł pt. „*Immunological landscape of consensus clusters in colorectal cancer [29]*” wpisuje się w wyżej opisane trendy, ponieważ charakteryzowano w nim, na tle podtypów molekularnych raka jelita grubego, poziomy komórek immunologicznych w guzach jak i ekspresję genów modulujących frakcję immunologiczną mikrośrodowiska (16 aktywatorów, 19 supresorów oraz 16 genów odpowiedzialnych za prezentację antygenów limfocytom T (MHC)). Do badań użyto zestawów danych transkryptomicznych pochodzących z 15 eksperymentów mikromacierzowych zawierających 1597 guzów jelita grubego oraz 125 tkanek makroskopowo prawidłowych, zlokalizowanych w pobliżu guza. Ponieważ analiza obejmowała wiele zestawów danych, wykryto bardzo silny efekt serii (*batch effect*), który został usunięty algorytmem ComBat [16]. Przed etapem grupowania danych osunięto próbki odstające (tzw. *outlier'y*). Ponadto metodą dekonwolucji oszacowano procentowy udział komórek prawidłowych w materiale z guzów oceniając tym samym wzbogacenie komórkami nowotworowymi. Do nienadzorowanego grupowania użyto agregowanej analizy skupień (*ensemble clustering*) będącej wypadkową użycia pięciu algorytmów klasteryzujących [14]. Rezultatem grupowania było 5 klastrów. Porównanie wzbogacenia komórek nowotworowych pomiędzy klastrami wykazało, że jeden z klastrów ma statystycznie znacząco niższy odsetek komórek nowotworowych (**Rycina 3**). Guzy zgrupowane w tym klastrze wyłączono z dalszych analiz. Następnie klasterzy przyporządkowano metodą ORA (*over-representation analysis*) do podtypów molekularnych CMS (*consensus molecular subtypes*) wyróżnionych przez Guinney i wsp. [11]. Kolejno, za pomocą algorytmu CIBERSORT przeprowadzono dekonwolucję profili ekspresji guzów uzyskując szacunkowe poziomy wzajemnych proporcji 22 rodzajów komórek immunologicznych [10]. Podtyp CMS1 wykazywał wzrost odsetka zarówno leukocytów związanych z odpornością nabytą (T follicular helper, Th i TCD8), jak i odpornością wrodzoną (komórki NK, komórki T gamma-delta, makrofagi M1, komórki dendrytyczne i neutrofile). CMS4 charakteryzował się najwyższym odsetkiem leukocytów związanych z aktywnością pronowotworową (eozynofile, monocyty, makrofagi M2 i limfocyty T regulatorowe (Treg)). CMS2 wyświetlał najwyżej wzbogacenie w komórki Th (CD4) i komórki pamięci B. CMS3 wykazywał najniższe wzbogacenie komórkami immunologicznymi. Podsumowując, mikrośrodowiska CMS1 i CMS4 charakteryzowały się wysokim wzbogaceniem komórkami układu immunologicznego choć o innym składzie

komórkowym. Natomiast CMS2 i CMS3, z uwagi na relatywnie niskie wzbogacenie leukocytami ich mikrośrodowiska, określa się podtypami immunologicznie zimnymi (*immune cold*) (**Rycina 4**). Porównanie poziomu ekspresji genów modulujących frakcję immunologiczną mikrośrodowiska pomiędzy podtypami i tkanką zdrową jelita wykazało na znacząco wyższą ekspresję aktywatorów, supresorów i genów z rodziny MHC CMS1 i CMS4 w porównaniu do CMS2 i CMS3. W tej pracy skupiliśmy się również na poszukiwaniu cech uniwersalnych dla wszystkich grup CMS gdyż mogą one stanowić potencjalnie uniwersalne cele terapeutyczne. Wśród stymulatorów układu immunologicznego, geny takie jak *IL6R* (interleukin-6 receptor), *BCMA* (B cell maturation antigen), *TACI* (Cyclophilin ligand interactor) oraz *CD27* wykazywały znacząco obniżoną ekspresję w co najmniej 3 grupach CMS. Ponadto, takie geny mające działanie supresorowe na układ immunologiczny miały znacząco wyższą ekspresję w większości grup CMS: *CXCL1*, *CXCL2*, *CXCR2*, *MIF* (macrophage migration inhibitory factor), *IDO1* (indoleamine 2,3-dioxygenase 1), *TDO2* (TRP-2,3- dioxygenase 2), *ARG2* (arginase 2), i *VEGFA* (vascular endothelial growth factor A).

Podsumowując, wyniki otrzymane w tej pracy wskazały na wysoką heterogenność immunologiczną mikrośrodowisk kłastrów CMS. W związku z tym skuteczna immunoterapia przeciwnowotworowa będzie wymagała nie tylko znajomości statusu CMS guza, ale także kombinatorycznych (wielolekowych) strategii. Pomimo wysokiej heterogenności immunologicznej mikrośrodowiska raka jelita grubego, powyższa praca wyłoniła kilka genów-kandydatów (takich jak *MIF*, *IDO1*, *TDO2* czy *ARG2*) mogących stać się celem uniwersalnych immunoterapii niezależnych od statusu CMS.



Rycina 3. Wykres wiolinowy przedstawiający odsetek komórek nowotworowych w guzach jelita grubego oszacowanych metodą dekonwolucji transkryptomu (Karpinski i wsp., 2017; [29]). Próbkę zostały podzielone na pięć grup metodą nienadzorowanego grupowania (oś X). Oś Y - szacowany odsetek (%) komórek nowotworowych w próbce. Grupa guzów opisana jako „Cluster 1” wyróżnia się znacząco niższym odsetkiem komórek nowotworowych. Guzy zgrupowane w tym klastrze wyłączone z dalszych analiz.



Rycina 4. Wykres wiolinowy przedstawiający odsetek leukocytów w guzach jelita grubego oszacowanych metodą dekonwolucji transkryptomu (Karpinski i wsp., 2017; [29]). Guzy zostały podzielone na cztery grupy metodą nienadzorowanego grupowania, dodano również profile ekspresji prawidłowych tkanek jelita (oś X). Oś Y - szacowany odsetek (%) leukocytów w próbkach. Na rycinie można zauważyć wyraźnie wyższe wzbogacenie w leukocyty guzów z grup CMS1 i CMS4. Natomiast wzbogacenie grup CMS2 i CMS3 w leukocyty było porównywalne do wzbogacenia tkanki normalnej.

Ad4. Odkrycie w latach 90tych ubiegłego wieku i zbadanie szlaków związanych z białkami PD-1 i CTLA4 w pierwszej dekadzie XXI wieku doprowadziło do odkrycia mechanizmów wyciszania odpowiedzi immunologicznych w mikrośrodkowisku guza poprzez indukcję ekspresji tych białek [30]. W konsekwencji w ostatnich 10 latach FDA zatwierdziło kilka leków przeciwnowotworowych celujących w ten jak i inne szlaki, odpowiedzialne za wyciszanie odpowiedzi limfocytów w mikrośrodkowiskach guzów. Pomimo tego, że powyższe leki zrewolucjonizowały współczesne strategie terapeutyczne, według danych z 2019 roku zaledwie u 12% pacjentów notuje się pozytywną odpowiedź na powyższe terapie, uśredniając wyniki z różnych typów nowotworów [30]. Przyczyn tych niepowodzeń można upatrywać w wysokiej heterogenności mikrośrodkowisk guzów. Z grubsza, mikrośrodkowiska guzów można podzielić na wysoko wzbogacone komórkami układu immunologicznego (*immune hot*) i nisko wzbogacone komórkami układu immunologicznego (*immune cold*). Wymienione wyżej strategie terapeutyczne są skutecznie jedynie w guzach z mikrośrodkowiskiem typu *immune hot*, w których lek odblokowuje wyciszony, przeciwnowotworowy potencjał nagromadzonych limfocytów T. Nie ma obecnie wypracowanych podejść terapeutycznych, opartych o modulację układu immunologicznego w guzach z mikrośrodkowiskiem typu *immune cold*. W pracy pt. „*Major Histocompatibility Complex Genes as Therapeutic Opportunity for Immune Cold Molecular Cancer Subtypes*”, bazując na danych transkryptomicznych, testowaliśmy teoretyczne możliwości zastosowania immunomodulacji przeciwnowotworowej w klastarach molekularnych z mikrośrodkowiskiem typu *immune cold* [31]. W tym celu z bazy The Cancer Genome Atlas (TCGA) pobrano dane otrzymane z macierzy metylacji DNA oraz sekwencjonowania DNA i RNA [32]. Analizowany zbiór danych składał się z 5651 guzów pierwotnych reprezentujących 13 podtypów nowotworów. Tło badania stanowiło nienadzorowane grupowanie guzów na podstawie ekspresji 14 genów z rodziny MHC (*major histocompatibility complex*), które pozwoliło podzielić guzy na 3 grupy: z wysoką ekspresją genów MHC (MHC-high), średnią ekspresją genów MHC (MHC-intermediate) oraz z niską ekspresją genów MHC (*MHC-low*). Otrzymane klastery MHC scharakteryzowano następującymi zmiennymi:

- poziomami limfocytów T CD4⁺ i limfocytów T CD8⁺ otrzymanymi przez dekonwucję profili metylacji DNA w guzach granulocytów metodą RPC (*Robust Partial Correlations*) wbudowanej w algorytm EpiDISH
- obciążeniem mutacyjnym guza (TMB - *Tumor mutation burden*) liczona jako średnia liczba mutacji somatycznych na guz.

- ekspresją 25 genów odpowiedzialnych za wyciszenie odpowiedzi limfocytów w mikrośrodowiskach guzów
- ekspresją 105 antygenów rakowo-jądrowych (*cancer germline antigens* (CAGs)). Ekspresja CAGs, za wyjątkiem jąder, nie występuje w prawidłowych tkankach ludzkich lecz często występuje w guzach nowotworowych. W związku z tym CAGs mają bardzo wysoki potencjał indukowania odpowiedzi immunologicznej co czyni je szczególnie atrakcyjnymi celami dla immunoterapii.
- przynależnością do jednego z 55 podtypów molekularnych guzów

Połowa (n=26) z 55 badanych podtypów molekularnych została przyporządkowana do klastra z niską ekspresją MHC (*MHC-low*). Charakteryzował je również najniższy poziom ekspresji genów 25 genów odpowiedzialnych za wyciszenie odpowiedzi immunologicznej oraz najniższe poziomy limfocytów T CD4+ i limfocytów T CD8+. Mutacje somatyczne w 14 genach MHC były zjawiskiem relatywnie rzadkim. Wysokie obciążenie mutacyjne guza stwierdzono w 3 podtypach molekularnych należących do grupy *MHC-low*: w raku płuca (podtyp Ad-1), w raku głowy i szyi (podtyp NSD1) oraz w raku pęcherza (podtyp LumU). W grupie *MHC-low* stwierdzono podwyższoną ekspresję kilku genów z grupy CAG m.in *TSGA10* i *CEP290*.

Otrzymane wyniki pozwoliły sformułować następujące wnioski:

- niska ekspresja genów odpowiedzialnych za wyciszenie odpowiedzi immunologicznej i niskie poziomy limfocytów T są przyczyną braku odpowiedzi guzów *MHC-low* na współczesne immunoterapie przeciwnowotworowe wymienione we wstępie paragrafu
- wysoka ekspresja niektórych genów CAG w grupie *MHC-low* wskazuje na to, że potencjał immunologiczny mikrośrodowiska skierowany na te antygeny, mógłby być wyzwolony pod warunkiem zwiększenia ekspresji genów MHC.
- wysokie obciążenie mutacyjne w 3 podtypach molekularnych należących do grupy *MHC-low* że potencjał immunologiczny mikrośrodowiska skierowany przeciwko tym guzom mógłby być wyzwolony pod warunkiem zwiększenia ekspresji genów MHC.

Konkludując, w niniejszej pracy stwierdziliśmy, że istnieje przeciwnowotworowy potencjał immunologiczny mikrośrodowiska guzów *MHC-low* jednak może on zostać wyzwolony jedynie w przypadku zwiększenia ekspresji genów z rodziny MHC.

Ad5. Nieprawidłowa metylacja DNA należy do najlepiej scharakteryzowanych zmian epigenetycznych w obrębie ludzkiego genomu i odnosi się to do metylacji cytozyny w dinukleotydach CpG, których liczbę w ludzkim genomie szacuje się na 30 milionów [33].

Duża część dinukleotydów CpG wykazuje wysoki stopień metylacji, jak na przykład zlokalizowane w sekwencjach powtarzających się (LINE, ALU) lub w regionach międzygenowych. Regiony bogate w dinukleotydy CpG, długości minimum 200 par zasad, nazwa się wyspami CpG. Około 60% genów kodujących białka zawiera wyspę CpG w regionie 5' promotora i/lub regionie startu transkrypcji zwykle o niskim poziomie metylacji [34]. Wysoki stopień metylacji wysp CpG związanych z promotorami genów prowadzi do ustalenia nieaktywnego transkrypcyjnie stanu chromatyny i wyciszenia genu. Należy zaznaczyć, że stan wysokiej metylacji genu nie jest stanem permanentnym i może zostać odwrócony/usunięty przez demetylazy z rodziny TET (ang. ten-eleven translocation (TET) methylcytosine dioxygenases). W trakcie starzenia się tkanek oraz w procesie transformacji nowotworowej oraz w obserwuje się dwa przeciwstawne zjawiska: tendencję do wysokiej metylacji wysp CpG w promotorach genów i jednoczesną demetylację sekwencji powtarzających się i regionów między-genowych. Zjawisko to określa się mianem dryfu metylacji (ang. methylation drift) lub dryfu epigenetycznego [35]. W związku z pierwszym zjawiskiem (starzenia się tkanek) od roku 2011 opracowano kilka algorytmów (tzw. zegarów epigenetycznych) do szacowania wieku tkanek na podstawie stopnia metylacji wybranych sond na macierzy Infinium HumanMethylation450 BeadChip [36]. Zegary epigenetyczne mają potencjalne zastosowanie do oszacowania ryzyka zachorowania na nowotwory jak również w kryminalistyce do oszacowania wieku zwłok. W 1999 opisano po raz pierwszy w raku jelita grubego zjawisko, polegające na predyspozycji do występowania, w niektórych guzach, bardzo częstej hipermetylacji regionów promotorowych wielu genów jednocześnie [37]. Fenomen ten, nazwany fenotypem metylatora (CIMP), wzbudza kontrowersje bowiem tylko przez niektórych autorów jest uznawany za osobny podtyp molekularny wielu typów nowotworów. Podłoże molekularne powstawania fenotypu metylatora zostało opisane jedynie dla glejaka wielopostaciowego, który wiąże się z mutacjami w genach kodujących dehydrogenazy izocytrynianowe (*IDH1* lub *IDH2*). Mutacje dehydrogenaz izocytrynianowych powodują hamowanie aktywności genów z rodziny TET a co za tym idzie spadku demetylacji w genomie i w konsekwencji akumulację hipermetylacji wysp CpG [38]. Od odkrycia CIMP podjęto kilka prób znalezienia wspólnych cech tego fenotypu w ujęciu wielu typów nowotworów jednocześnie (ang. *pan-cancer*) m.in Sanchez-Vega i wsp. (2015) oraz Gevaert i wsp. (2015) [39, 40]. Wyniki tych prac wskazują na bardzo wysoką heterogenność molekularną guzów CIMP występującymi w różnych typach nowotworów. Artykuł pt. „*Pan-cancer analysis reveals presence of pronounced DNA methylation drift in CpG island methylator phenotype clusters*” wpisuje się w trend poszukiwania wspólnych

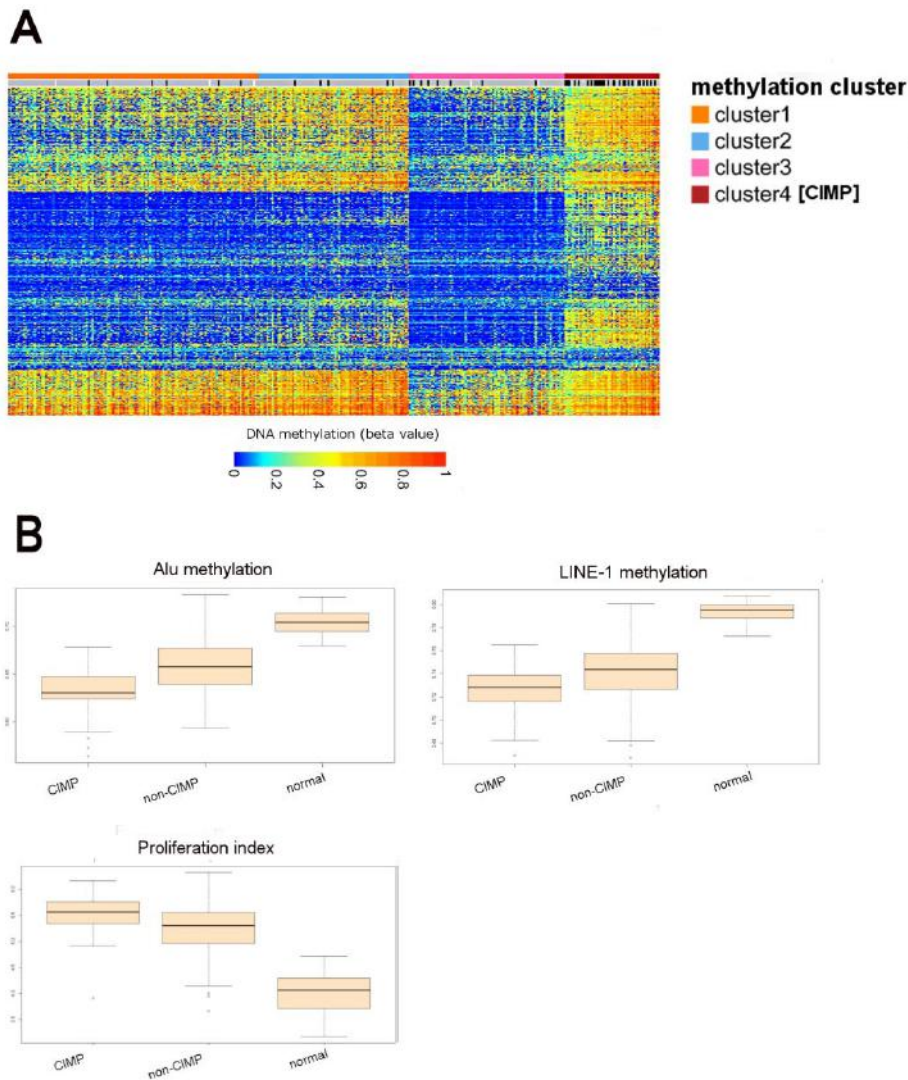
cech CIMP w ujęciu *pan-cancer* [41]. Celem pracy było poszukiwanie wspólnych cech molekularnych guzów z CIMP w każdym z typów nowotworów. W tym celu z bazy The Cancer Genome Atlas (TCGA) pobrano dane otrzymane za pomocą mikromacierzy Infinium HumanMethylation450 BeadChip (Illumina) pozwalające zbadać metylację dla około 480 tysięcy dinukleotydów CpG oraz profile ekspresji genów otrzymane za pomocą techniki RNA-seq. Analizowany zbiór danych po filtracji zawierał profile metylacji i ekspresji genów dla 6570 guzów pierwotnych reprezentujących 23 podtypy nowotworów. Ponieważ analiza obejmowała 23 zestawy danych, dla których profile metylacji DNA lub wygenerowane zostały na kilkuset mikroczipach, wykryto bardzo silny efekt serii (*batch effect*), który został usunięty algorytmem *ComBat* [16].

Do wyróżnienia klastrów CIMP użyto agregowanej analizy skupień (ensemble clustering) będącej wypadkową użycia pięciu algorytmów klasterujących [14]. Poza identyfikacją guzów z CIMP w każdym z typów nowotworów, za pomocą technik nienadzorowanego klusterowania danych macierzowych, zbadano szereg zmiennych m.in:

- oceniono prędkość podziałów mitotycznych komórek na podstawie metylacji wybranych sond Infinium HumanMethylation450 za pomocą algorytmu EPIToc [36]
- zbadano ekspresję genów związanych z proliferacją (*MKI67 (Ki-67)*, *TOP2A*, *PCNA* oraz *MCM2*).
- zbadano poziom metylacji obszarów między-genowych
- zbadano metylację w sekwencji powtarzających się (Alu and LINE-1)

Przeprowadzone obliczenia wskazały na istnienie cech wspólnych dla 21 z 23 badanych klusterów CIMP: przyspieszone podziały mitotyczne, zwiększona ekspresja genów związanych z proliferacją oraz obniżony poziom metylacji obszarów między-genowych oraz sekwencji powtarzających się (Alu and LINE-1). Analizę niektórych zmiennych, na przykładzie raka jelita grubego, przedstawiono na **rycinie 5**.

Otrzymane wyniki pozwoliły sformułować hipotezę dotyczącą możliwego powstawania CIMP. Czynnikiem inicjującym fenotyp metylatora są znacząco przyspieszone podziały komórek nowotworowych, co ma prawdopodobnie miejsce na początkowych etapach transformacji nowotworowej i czego wynikiem jest wyróżniający się w guzach CIMP dryf metylacji. Tłumaczyłoby to również brak molekularnych cech wspólnych dla fenotypu metylatora występującego w różnych typach nowotworów i jego wysoce tkankową specyficzność.



Rycina 5. Identyfikacja klasteru CIMP jego asocjacji z niektórymi zmiennymi na przykładzie raka jelita grubego (Karpinski i wsp., 2017; [41]). Do analizy włączono 351 guzów pierwotnych jelita grubego oraz 42 tkanki makroskopowo prawidłowe, przystające do guza. **A)** Za pomocą agregowanej analizy skupień wyróżniono 4 klastry metylacyjne. Klaster o najwyższej, uśrednionej metylacji wysp CpG oznaczono jako klaster CIMP, pozostałe trzy klastry w analizach asocjacyjnych połączono razem w jedną grupę opisaną jako non-CIMP. **B)** Wykresy pudełkowe ilustrujące poziomy odpowiednio: metylacji dinukleotydów CpG w sekwencjach Alu, metylacji dinukleotydów CpG w sekwencjach LINE-1 oraz indeks proliferacji (uśredniona ekspresja genów związanych z proliferacją (*MKI67* (*Ki-67*), *TOP2A*, *PCNA* oraz *MCM2*)) w guzach z CIMP, w guzach bez CIMP (non-CIMP) oraz w tkance normalnej (normal). Należy zauważyć, że zarówno guzy CIMP jak i non-CIMP wykazują obniżenie stopnia metylacji Alu oraz LINE-1 i podwyższenie indeksu proliferacyjnego w porównaniu do tkanki normalnej. Należy jednak podkreślić, że guzy CIMP od guzów non-CIMP różniły znacząco niższe poziomy metylacji Alu oraz LINE-1 oraz znacząco wyższy poziom.

Ad.6. Kolejna praca pt. „*Multilevel omic data clustering reveals variable contribution of methylator phenotype to integrative cancer subtypes*” była kontynuacją badań podjętych w publikacji omówionej jako piąta. W guzach nowotworowych większość zmian na poziomie molekularnym ma charakter zmian ubocznych (*passengers*), nie wpływających na fenotyp guza, w przeciwieństwie do zmian napędzających (*drivers*). Celem tej pracy było zbadanie czy i w jakich typach nowotworów fenotyp metylatora ma udział w tworzeniu specyficznych integracyjnych klusterów. Statystycznie znaczące wzbogacenie jednego klusterów integracyjnych w guzy z fenotypem metylatora mogłoby świadczyć o jego charakterze napędzającym w danym typie nowotworu. Z bazy The Cancer Genome Atlas (TCGA) pobrano: mikromacierzowe profile metylacji DNA, profile ekspresji wszystkich genów kodujących białka otrzymane za pomocą techniki RNA-seq oraz mikromacierzowe dane obejmujące zaburzenia liczby kopii genów kodujących białka. Analizowany zbiór danych po filtracji zawierał dane dla 4688 guzów pierwotnych reprezentujących 17 podtypów nowotworów. Z wszystkich danych został usunięty efekt serii. W badaniach zwrócono uwagę na obecność próbek o niskiej czystości (*low purity*) rozumiane jako próbki z niskim odsetkiem komórek nowotworowych. Próbki o niskiej czystości są jednym z najistotniejszych z czynników zakłócających analizę nienadzorowanego grupowania. Do obliczenia czystości próbek użyto algorytmu zaproponowanego przez Aran i wsp. opartego o analizę metylacji wybranych sond mikromacierzy Infinium HumanMethylation450 BeadChip (Illumina) [41]. W wyżej opisanym kontekście analizowano również wpływ klastrów integracyjnych związanych z CIMP na całkowite przeżycie. Przeprowadzono także analizy transkryptomu w celu poszukiwania molekularnych cech wspólnych dla klastrów CIMP. Do wyróżnienia klastrów CIMP użyto agregowanej analizy skupień (*ensemble clustering*) będącej wypadkową użycia pięciu algorytmów klasterujących [14]. Do badań integracyjnych wykorzystano integracyjną analizę skupień (*integrative clustering*) opartą o jednoczesne klasterowanie 3 rodzajów danych molekularnych (metylacji DNA, ekspresji genów oraz zaburzeń liczby kopii genów) metodą fuzji sieci podobieństw (Similarity network fusion - SNF) połączoną z grupowaniem konsensusowym (consensus clustering - CC) [15]. **Rycina 6** ilustruje dwa przykładowe wyniki klasterowania integracyjnego [42]. Na podstawie przeprowadzonych analiz stwierdzono, że:

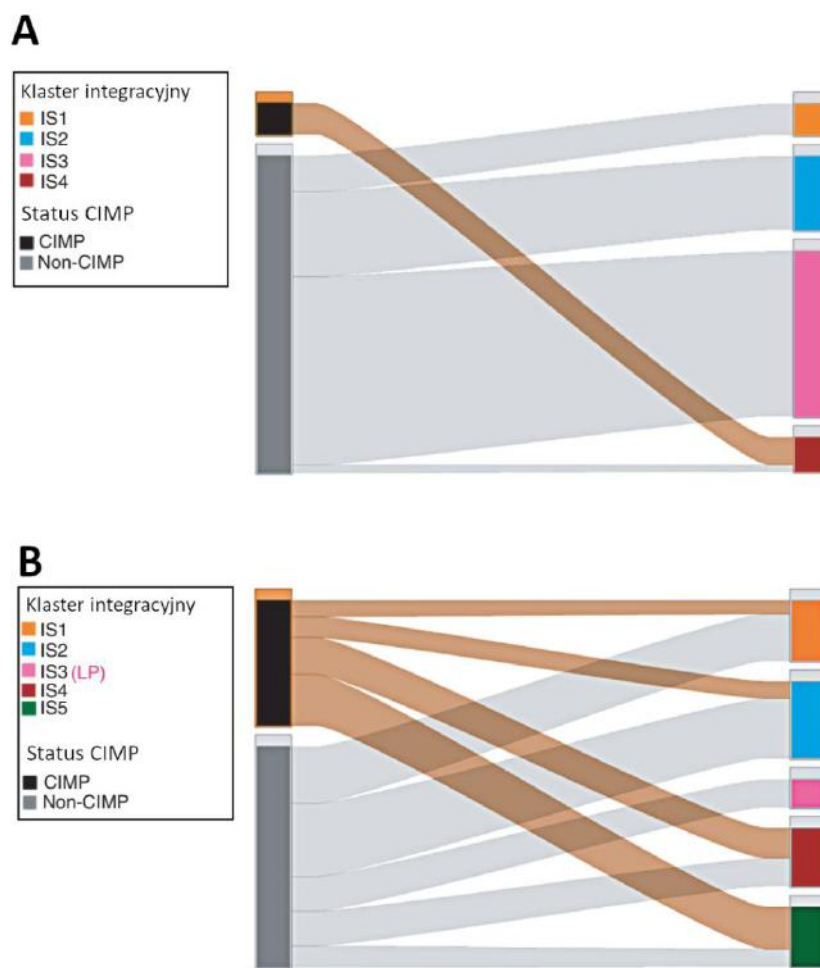
- z 17 typów nowotworów w 10 znaleziono klaster integracyjny, statystycznie znamienne wzbogacony w guzy CIMP (m.in. w raku jelita grubego, raku żołądka, raku brodawkowatym i jasnokomórkowym nerki oraz rakach głowy i szyi). Przeciętnie

rozmiary klastra integracyjnego związanego z CIMP wynosiły od 15% do 20% badanych guzów w danym typie nowotworu.

- z dużym prawdopodobieństwem, fenotyp metylatora jest jednym z czynników napędzających fenotyp guzów należących do wyżej wspomnianych klastrów.
- z 10 klastrów integracyjnych, wzbogaconych w guzy CIMP, tylko jeden - w raku żołądka związany z infekcją wirusem Epsteina-Barr - wykazywał znamienne statystyczne związki z czasem przeżycia po wprowadzeniu korekty na czynniki zakłócające takie jak wiek, płeć i stopień zaawansowania klinicznego choroby.
- pomiędzy 10 klastrami integracyjnymi, wzbogaconymi w guzy CIMP, nie znaleziono cech wspólnych w oparciu o szereg analiz transkryptomicznych (między innymi: różnicową ekspresję (*differential expression*), analizę ścieżek (*pathway analysis*) czy analizę aktywowanych czynników transkrypcyjnych.

Otrzymane wyniki pozwoliły sformułować następujące wnioski:

- CIMP jest fenotypem, który w około 60% typów nowotworów, może mieć cechy napędzające dla pojedynczego, niewielkiego klastra najczęściej stanowiącego 15-20% badanej próby.
- Na poziomie transkryptomu, klaster wzbogacony w guzy CIMP, mają cechy wysoce specyficzne
- CIMP nie jest dobrym markerem całkowitego przeżycia



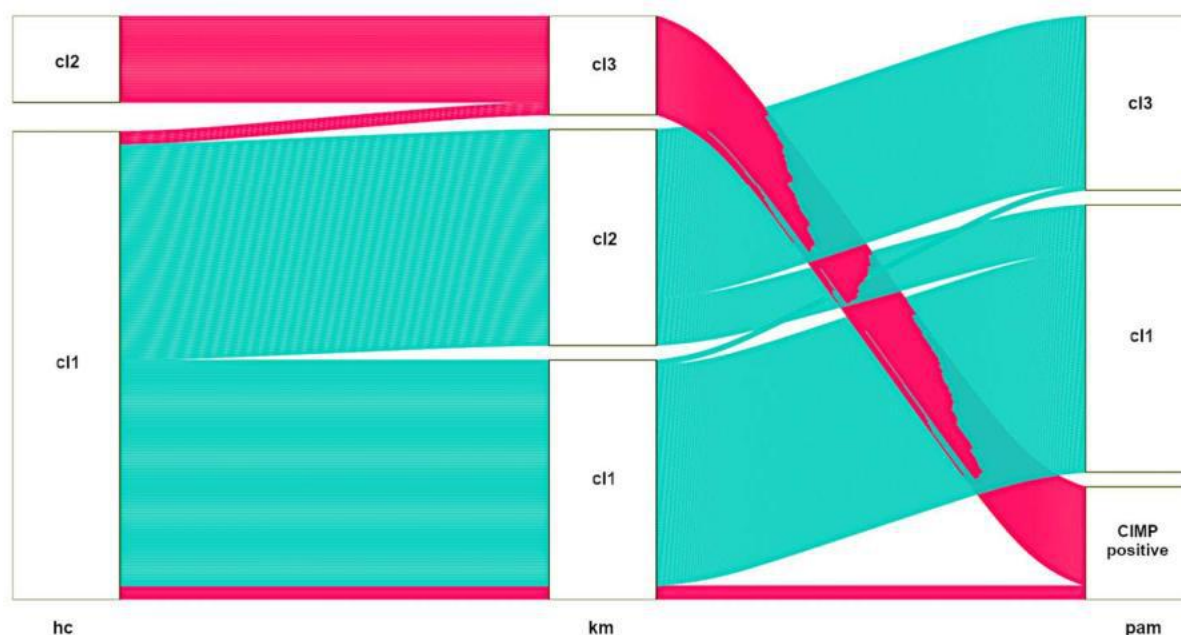
Rycina 6. Wykresy aluwialne przedstawiające rozkład przynależności próbek z fenotypem metylatora (CIMP) [zdefiniowanym na podstawie klasterowania poziomu metylacji wysp CpG] a klastrami intergacyjnymi zdefiniowanymi w tych samych próbkach na podstawie klasterowania 3 rodzajów danych molekularnych (metylacji DNA, ekspresji genów oraz zaburzeń liczby kopii genów) (Karpinski i wsp., 2018; [42]). **A)** Rak jelita grubego, jeden z klastrów integracyjnych jest utworzony prawie wyłącznie przez guzy z CIMP. Jest to przykład CIMP o charakterze napędzającym („driver”) **B)** Rak pęcherza moczowego, guzy CIMP współtworzą wraz guzami non-CIMP prawie wszystkie klastry intergacyjne poza klastrzem IS3, zawierającym próbki niskiej czystości (LP- low purity). Jest to przykład CIMP o charakterze zmiany ubocznej („passenger”).

Ad7. „CpG island methylator phenotype - a hope for the future or a road to nowhere?” jest pracą łączącą przegląd literatury z ostatnich 20 lat na temat fenotypu metylatora (CIMP) oraz obliczenia biostatystyczne służące zbadaniu niektórych tez postawionych w tej pracy. Do badań bibliograficznych dotyczących CIMP wykorzystano techniki automatycznego przeszukiwania bazy PubMed i statystyczną analizę tekstu (*text mining*) zawartych w tytułach publikacji i abstraktach [43]. W latach 1999 do 2020 opublikowano 1164 prac badawczych i przeglądowych dotyczących CIMP. Naszą uwagę zwrócił spadek zainteresowania

tematem CIMP przejawiający się wyraźnym spadkiem publikowanych rocznie artykułów poczynszyszy od 2015 roku. Przyczyn tego stanu rzeczy można upatrywać w kilku czynnikach. Po pierwsze, już od pierwszych lat po odkryciu CIMP (w 1999 roku) kwestionowano jego istnienie raku jelita grubego jako kolejnego, oprócz niestabilności mikrosatelitarnej i niestabilności chromosomowej, fenotypu molekularnego o skali genomowej [44]. Sugerowano, że jest to proces losowej akumulacji zmian epigenetycznych w niektórych guzach związanych ze starzeniem się (tkanki) nie wpływający na fenotyp guza. Nasze badania opisane w pracy nr. 6 dowodzą, że w przypadku około połowy badanych typów nowotworów CIMP może nie mieć związku z podtypem molekularnym guza, co może świadczyć o CIMP jako zmianie typu ubocznego („*passenger*”) w tych typach nowotworów. Po drugie, już od pierwszych lat po odkryciu CIMP zwracał uwagę fakt braku specyficznych markerów (sekwencji lub genów) do detekcji CIMP. Dopiero w roku 2006 i 2007 dwa osobne zespoły zaproponowały zestawy kilku markerów, przetestowanych na setkach próbek, do oceny obecności CIMP w raku jelita grubego [45, 46]. Do dnia dzisiejszego nie pojawiły się doniesienia o opracowaniu podobnych zestawów dla innych typów nowotworów. W związku z tym nie ma obecnie przystępnej, wysoko przepustowej metody, poza rakiem jelita grubego, przeznaczonej do oznaczania i badania CIMP. Co więcej, nawet w obszarze tak jednolitych, w sensie technologicznym, badań metylacji guzów przeprowadzonych w ramach TCGA nie ma konsensusu co do sposobu analizy metylacji DNA. W niniejszej pracy przeanalizowaliśmy w jaki sposób normalizowano dane, selekcjonowano sondy i grupowano dane dla 23 podtypów nowotworów. Każda z 23 analiz przeprowadzonych przez TCGA różniła się w sposób znaczący. W ramach niniejszej pracy przeprowadziliśmy eksperyment, który wyraźnie pokazał, że nawet różnice w doborze liczby sond lub doborze algorytmów klasteryzujących, pozostawiając resztę parametrów niezmiennych, mogą zmienić przyporządkowanie guzów do CIMP (**Rycina 7**) [47]. Podsumowując punkt drugi - nie ma obecnie konsensusu co do doboru markerów i metody do definiowania CIMP w większości typów nowotworów. Po trzecie, jak wskazują nasze własne obliczenia opisane w pracy nr. 6 [41] jak i kwerenda publikacji typu meta-analiza wykonane w ramach niniejszej pracy [47], CIMP nie ma bezpośredniego, statystycznie znaczącego związku z czasem przeżycia pacjentów. Po czwarte, guzy z CIMP, z uwagi na wysoki poziom metylacji wielu genów, były uznawane na początku jak bardzo dobry cel dla terapii epigenetycznych jak chociażby terapii zawierających związki demetylujące (np. inhibitory DNA metyltransferaz) . Terapie przeciwnowotworowe oparte o działanie na mechanizmy epigenetyczne nie odniosły jednak spodziewanego skutku w terapii guzów litych. Przyczyną tego stanu rzeczy jest

niewystarczający wychwyt leku przez guzy lite, który pozostaje głównym problemem chemioterapii. Z uwagi na wysoką, tkankową specyficzność guzów z CIMP nie odnaleziono pozaepigenetycznej zmiany (np. mutacji lub wysokiej ekspresji genu/ów), która mogłaby się stać celem terapii przeciwnowotworowej celującej pośrednio w CIMP. Wyjątkiem od tej reguły są typy nowotworów, w których za fenotyp metylatora odpowiada mutacja w genie *IDH1*, na przykład glejak o niskim stopniu złośliwości, dla którego terapii FDA zatwierdziła dwa inhibitor *IDH-1* vorasidenib i ivosidenib [47].

Podsumowując, CIMP pozostaje nadal słabo poznany, heterogenny, choć bez wątpienia istotnym w wielu typach nowotworów, czynnikiem napędzającym transformację nowotworową. W przypadku znalezienia efektywnej metody transferu terapii epigenetycznych do guzów litych, CIMP może się stać ważnym i pożądanym markerem decydującym o doborze leczenia. Należy spodziewać się więc, że obserwowany spadek zainteresowania zjawiskiem jakim jest CIMP jest tymczasowy.



Rycina 7. Wykres aluwialny przedstawiający przyporządkowanie guzów do CIMP (różowy kolor) w zależności od zastosowanego algorytmu grupującego w raku jelita grubego (Karpinski i wsp., 2022; [47]). Pozostałe parametry analizy normalizacja danych, dobór i liczba sond były niezmiennie. "km" - algorytm k-średnich; hc - grupowanie hierarchiczne; pam - algorytm k-medoidów

Wnioski końcowe

Heterogenność, z którą styka się badacz podczas analizy danych, szczególnie tych otrzymanych za pomocą technik mierzących zjawiska w skali genomowej, jest wielowymiarowym zjawiskiem. Nie jest możliwe podjęcie właściwego wnioskowania o otrzymanych wynikach bez zastosowania technik, które niwelują (zmiennosć techniczną) bądź porządkują heterogenne dane (nienadzorowane klastrowanie czy dekonwolucja).

W podsumowaniu, do najważniejszych osiągnięć kliniczno-poznawczych przedstawionego cyklu należą:

- identyfikacja zaburzeń proporcji komórek NK we krwi pacjentów ze schizofrenią
- jeden z pierwszych opisów heterogenności immunologicznej mikrośrodków molekularnych podtypów raka jelita grubego (CMS) wraz ze wskazaniem potencjalnych genów, które mogą być celem immunoterapii niezależnej od CMS.
- wskazanie na znacząco przyspieszone podziały komórek nowotworowych jako możliwej przyczyny powstawania fenotypu metylatora w większości typów nowotworów
- wskazanie podtypów molekularnych nowotworów, w których fenotyp metylatora może mieć cechy napędzające nowotwór (*driver*)

Bibliografia:

1. Supplitt S, Karpinski P, Sasiadek M, Laczmanska I. Current Achievements and Applications of Transcriptomics in Personalized Cancer Medicine. *Int J Mol Sci*. 2021 Jan 31;22(3).
2. Struck TJ, Mannakee BK, Gutenkunst RN. The impact of genome-wide association studies on biomedical research publications. *Hum Genomics*. 2018 Aug 13;12(1):38.
3. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res*. 2013 Jan;41(Database issue):D991-5.
4. Zhang Z, Hernandez K, Savage J, Li S, Miller D, Agrawal S, et al. Uniform genomic data analysis in the NCI Genomic Data Commons. *Nat Commun*. 2021 Feb 22;12(1):1226.
5. Turajlic S, Sottoriva A, Graham T, Swanton C. Resolving genetic heterogeneity in cancer. *Nat Rev Genet*. 2019 Jul;20(7):404-16.
6. Lauss M, Visne I, Kriegner A, Ringnér M, Jönsson G, Höglund M. Monitoring of technical variation in quantitative high-throughput datasets. *Cancer Inform*. 2013;12:193-201.
7. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell*. 2017 Feb 09;168(4):613-28.
8. Prasetyanti PR, Medema JP. Intra-tumor heterogeneity from a cancer stem cell perspective. *Mol Cancer*. 2017 Feb 16;16(1):41.
9. Abbas AR, Wolslegel K, Seshasayee D, Modrusan Z, Clark HF. Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. *PLoS One*. 2009;4(7):e6098.

10. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015 May;12(5):453-7.
11. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med*. 2015 Nov;21(11):1350-6.
12. Sekula M, Datta S. optCluster: An R Package for Determining the Optimal Clustering Algorithm. *Bioinformatics*. 2017;13(3):101-03.
13. Ayton SG, Pavlicova M, Robles-Espinoza CD, Tamez Peña JG, Treviño V. Multiomics subtyping for clinically prognostic cancer subtypes and personalized therapy: A systematic review and meta-analysis. *Genet Med*. 2022 Jan;24(1):15-25.
14. Sweeney TE, Chen AC, Gevaert O. Combined Mapping of Multiple clUsteriNg ALgorithms (COMMUNAL): A Robust Method for Selection of Cluster Number, K. *Sci Rep*. 2015 Nov;5:16971.
15. Xu T, Le TD, Liu L, Su N, Wang R, Sun B, et al. CancerSubtypes: an R/Bioconductor package for molecular cancer subtype identification, validation and visualization. *Bioinformatics*. 2017 Oct;33(19):3131-33.
16. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007 Jan;8(1):118-27.
17. Guo J, Liu C, Wang Y, Feng B, Zhang X. Role of T helper lymphokines in the immune-inflammatory pathophysiology of schizophrenia: Systematic review and meta-analysis. *Nord J Psychiatry*. 2015 Jul;69(5):364-72.
18. Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B. Meta-analysis of cytokine alterations in schizophrenia: clinical status and antipsychotic effects. *Biol Psychiatry*. 2011 Oct 01;70(7):663-71.
19. Barbosa IG, Rocha NP, Assis F, Vieira É, Soares JC, Bauer ME, et al. Monocyte and lymphocyte activation in bipolar disorder: a new piece in the puzzle of immune dysfunction in mood disorders. *Int J Neuropsychopharmacol*. 2014 Oct 31;18(1).
20. Sweeney TE, Shidham A, Wong HR, Khatri P. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Sci Transl Med*. 2015 May;7(287):287ra71.
21. Darcy CJ, Woodberry T, Davis JS, Piera KA, McNeil YR, Chen Y, et al. Increased plasma arginase activity in human sepsis: association with increased circulating neutrophils. *Clin Chem Lab Med*. 2014 Apr;52(4):573-81.
22. Zhao W, Huang Y, Liu Z, Cao BB, Peng YP, Qiu YH. Dopamine receptors modulate cytotoxicity of natural killer cells via cAMP-PKA-CREB signaling pathway. *PLoS One*. 2013;8(6):e65860.
23. Karpiński P, Frydecka D, Sasiadek MM, Misiak B. Reduced number of peripheral natural killer cells in schizophrenia but not in bipolar disorder. *Brain Behav Immun*. 2016 May;54:194-200.
24. Hannon E, Dempster E, Viana J, Burrage J, Smith AR, Macdonald R, et al. An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. *Genome Biol*. 2016 Aug 30;17(1):176.
25. Teschendorff AE, Breeze CE, Zheng SC, Beck S. A comparison of reference-based algorithms for correcting cell-type heterogeneity in Epigenome-Wide Association Studies. *BMC Bioinformatics*. 2017 Feb;18(1):105.

26. Hess JL, Tylee DS, Barve R, de Jong S, Ophoff RA, Kumarasinghe N, et al. Transcriptome-wide mega-analyses reveal joint dysregulation of immunologic genes and transcription regulators in brain and blood in schizophrenia. *Schizophr Res*. 2016 Oct;176(2-3):114-24.
27. Karpiński P, Samochowiec J, Frydecka D, Sasiadek MM, Misiak B. Further evidence for depletion of peripheral blood natural killer cells in patients with schizophrenia: A computational deconvolution study. *Schizophr Res*. 2018 Nov;201:243-48.
28. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature*. 2000 Aug 17;406(6797):747-52.
29. Karpinski P, Rossowska J, Sasiadek MM. Immunological landscape of consensus clusters in colorectal cancer. *Oncotarget*. 2017 DEC 1 2017;8(62):105299-311.
30. Wang J, Yuan R, Song W, Sun J, Liu D, Li Z. PD-1, PD-L1 (B7-H1) and Tumor-Site Immune Modulation Therapy: The Historical Perspective. *J Hematol Oncol*. 2017 Jan 25;10(1):34.
31. Karpiński P, Łaczmański Ł, Sasiadek MM. Major Histocompatibility Complex Genes as Therapeutic Opportunity for Immune Cold Molecular Cancer Subtypes. *J Immunol Res*. 2020;2020:8758090.
32. Zhu Y, Qiu P, Ji Y. TCGA-assembler: open-source software for retrieving and processing TCGA data. *Nat Methods*. 2014 Jun;11(6):599-600.
33. Peixoto P, Cartron PF, Serandour AA, Hervouet E. From 1957 to Nowadays: A Brief History of Epigenetics. *Int J Mol Sci*. 2020 Oct 14;21(20).
34. Hughes LA, Melotte V, de Schrijver J, de Maat M, Smit VT, Bovée JV, et al. The CpG island methylator phenotype: what's in a name? *Cancer Res*. 2013 Oct;73(19):5858-68.
35. Zheng SC, Widschwendter M, Teschendorff AE. Epigenetic drift, epigenetic clocks and cancer risk. *Epigenomics*. 2016 May;8(5):705-19.
36. Yang Z, Wong A, Kuh D, Paul DS, Rakyan VK, Leslie RD, et al. Correlation of an epigenetic mitotic clock with cancer risk. *Genome Biol*. 2016 Oct;17(1):205.
37. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci U S A*. 1999 Jul;96(15):8681-6.
38. Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell*. 2010 May;17(5):510-22.
39. Gevaert O, Tibshirani R, Plevritis SK. Pancancer analysis of DNA methylation-driven genes using MethylMix. *Genome Biol*. 2015 Jan 29;16:17.
40. Sánchez-Vega F, Gotea V, Margolin G, Elnitski L. Pan-cancer stratification of solid human epithelial tumors and cancer cell lines reveals commonalities and tissue-specific features of the CpG island methylator phenotype. *Epigenetics Chromatin*. 2015;8:14.
41. Karpinski P, Pesz K, Sasiadek MM. Pan-cancer analysis reveals presence of pronounced DNA methylation drift in CpG island methylator phenotype clusters. *Epigenomics*. 2017 Nov;9(11):1341-52.
42. Karpinski P, Patai AV, Hap W, Kielan W, Laczmanska I, Sasiadek MM. Multilevel omic data clustering reveals variable contribution of methylator phenotype to integrative cancer subtypes. *Epigenomics*. 2018 Oct;10(10):1289-99.

43. O'Mara-Eves A, Thomas J, McNaught J, Miwa M, Ananiadou S. Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Syst Rev*. 2015 Jan 14;4(1):5.
44. Yamashita K, Dai T, Dai Y, Yamamoto F, Perucho M. Genetics supersedes epigenetics in colon cancer phenotype. *Cancer Cell*. 2003 Aug;4(2):121-31.
45. Ogino S, Cantor M, Kawasaki T, Brahmandam M, Kirkner GJ, Weisenberger DJ, et al. CpG island methylator phenotype (CIMP) of colorectal cancer is best characterised by quantitative DNA methylation analysis and prospective cohort studies. *Gut*. 2006 Jul;55(7):1000-6.
46. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet*. 2006 Jul;38(7):787-93.
47. Paweł K, Maria Małgorzata S. CpG Island Methylator Phenotype-A Hope for the Future or a Road to Nowhere? *Int J Mol Sci*. 2022 Jan 13;23(2).

5. Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej.

I) Drugim osiągnięciem, które uznaję za istotne i wyróżniające się całościowo w dorobku naukowym, poza zasadniczym, wyodrębnionym cyklem publikacji jest moja współpraca jako analityk danych z prof. Mai P. Hoang (Harvard Medical School, list uwierzytelniający przedstawiono w załącznikach). Pokłosiem tej współpracy jest cykl publikacji dotyczący analizy transkryptomu w rzadkich typach nowotworów: czerniaku błony naczyniowej oka i raku z komórek Merkla.

Donizy Piotr, Krzyżiński Mateusz, Markiewicz Anna, **Karpiński Paweł**, Kotowski Krzysztof, Kowalik Artur, Orłowska-Heitzman Jolanta, Romanowska-Dixon Bożena, Biecek Przemysław, Hoang Mai P.: Machine learning models demonstrate that clinicopathologic variables are comparable to gene expression prognostic signature in predicting survival in uveal melanoma, *European Journal of Cancer*, 2022, vol. 174, s. 251-260

IF: 10

Kowalik Artur, **Karpiński Paweł**, Markiewicz Anna, Orłowska-Heitzman Jolanta, Romanowska-Dixon Bożena, Donizy Piotr, Hoang Mai P.: Molecular profiling of

primary uveal melanoma: results of a Polish cohort, Melanoma Research, 2023, vol. 33, nr 2, s. 104-115

IF: 3,199

Paweł Karpinski, Ivy Rosales, Lukasz Laczmanski, Artur Kowalik, Scott Wenson, Mai P Hoang. Expression of Genes Associated with Epithelial-Mesenchymal Transition in Merkel Cell Polyomavirus-Negative Merkel Cell Carcinoma. Laboratory Investigation. 2023 May 17;100177.doi: 10.1016/j.labinv.2023.100177

IF: 5,511

II) Trzecim osiągnięciem, które uznaję za istotne, poza cyklem publikacyjnym, jest moja praca w Laboratorium Genomiki i Bioinformatyki, Instytut Immunologii i Terapii Doświadczalnej im. Ludwika Hirszfelda Polskiej Akademii Nauk we Wrocławiu (zatrudnienie od 2019). Przedmiotem mojej pracy jest szereg zagadnień z dziedziny biostatystyki i transkryptomiki:

Wagner Marta, Tupikowski Krzysztof, Jasek Monika, Tomkiewicz Anna, Witkiewicz Agata, Ptaszkowski Kuba, **Karpiński Paweł**, Zdrojowy Romuald, Hałoń Agnieszka, Karabon Lidia: SNP-SNP interaction in genes encoding PD-1/PD-L1 axis as a potential risk factor for clear cell renal cell carcinoma, Cancers, 2020, vol. 12, nr 12, art.3521 [15 s.], DOI:10.3390/cancers12123521

IF: 6,639

Buska-Mach Katarzyna, Kędzierska Anna Ewa, Lepczyński Adam, Herosimczyk Agnieszka, Ozgo Małgorzata, **Karpiński Paweł**, Gomułkiewicz Agnieszka, Lorek Daria, Sławek Anna, Dzięgiel Piotr, Chełmońska-Soyta Anna: Differential signals from TNF α -treated and untreated embryos in uterine tissues and splenic CD4⁺ T lymphocytes during preimplantation pregnancy in mice, Frontiers in Veterinary Science, 2021, vol. 8, art.641553 [16 s.], DOI:10.3389/fvets.2021.641553

IF: 3,471

Pawełka Dorota, Łaczmańska Izabela, **Karpiński Paweł**, Supplitt Stanisław, Witkiewicz Wojciech, Knychalski Bartłomiej, Pelak Joanna, Żebrowska Paulina, Łaczmański Łukasz:

Machine-learning-based analysis identifies miRNA expression profile for diagnosis and prediction of colorectal cancer: a preliminary study, *Cancer Genomics & Proteomics*, 2022, vol. 19, nr 4, s. 503-511, DOI:10.21873/cgp.20336

IF: 3,395

Moksud Nafeesa, Wagner Marta, Pawełczyk Konrad Jan, Porębska Irena Maria, Muszczyńska-Bernhard Beata, Kowal Aneta, Wiśniewski Andrzej, Kosacka Monika Maria, Kończak Julia, **Karpiński Paweł**, Frydryk Dominik, Andrzejczak Anna, Karabon Lidia, Kuśnierczyk Piotr, Jasek Monika: Common inherited variants of PDCD1, CD274 and HAVCR2 genes differentially modulate the risk and prognosis of adenocarcinoma and squamous cell carcinoma, *Journal of Cancer Research and Clinical Oncology*, 2023, DOI:10.1007/s00432-023-04602-8

IF: 4,322

Jestem promotorem pomocniczym pracy doktorskiej prowadzonej w Instytucie Immunologii i Terapii Doświadczalnej im. Ludwika Hirszfelda Polskiej Akademii Nauk we Wrocławiu przez mgr Paulinę Żebrowską pt. „The intestinal microbiome and bacterial transcriptome analysis in patients with the autism spectrum disorders”

Jestem wykonawcą następujących grantów realizowanych w Instytucie Immunologii:

1. Modyfikowane egzosomy pochodzenia nowotworowego jako nośniki cząsteczek modulujących środowisko nowotworowe. NCN SONATA BIS 8 UMO-2018/30/E/NZ5/00711 2019-2024
2. Punkty kontroli immunologicznej w niedrobnokomórkowym raku płuc – badanie genetycznych i epigenetycznych mechanizmów regulacji ekspresji mRNA i białka. 2019/33/B/NZ5/03029 2020-2023

6. Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę

A. Jestem zatrudniony na stanowisku adiunkta i prowadzę zajęcia z genetyki laboratoryjnej na Wydziale Lekarskim (kierunek: Lekarski), Wydziale Farmaceutycznym z O. Analityki Medycznej (kierunek: Analityka Medyczna) oraz English Division (Faculty of Medicine) Uniwersytetu Medycznego we Wrocławiu w wymiarze 240 godzin rocznie.

B. Jestem współautorem dwóch rozdziałów w 2 książkach

Karpiński Paweł: Rola zmian epigenetycznych w nowotworach, implikacje terapeutyczne, W: Dlaczego znana aktorka poddaje się zabiegowi mastektomii i adnektomii, czyli onkogenetyka w praktyce klinicznej, (red.) Maria Sasiadek, Agnieszka Stembalska, Wrocław 2014, Wrocławskie Wydawnictwo Naukowe Atla 2, s. 77-83, ISBN 978-83-60732-95-3

Karpiński Paweł, Blin Nikolaus, Sasiadek Maria M.: Genetic and epigenetic alterations in sporadic colorectal cancer: clinical implications, W: Nucleic acids as molecular diagnostics, (red.) Andreas Keller, Eckart Meese, Weinheim 2015, Wiley-VCH Verlag GmbH & Co. KGaA, s. 39-62, ISBN 978-3-527-33556-5

C. Aktywny udział w międzynarodowych i krajowych konferencjach naukowych - jestem współautorem 54 abstraktów konferencyjnych

D. W latach od 2014 do 2018 należałem do Komitetu Naukowego (Sesja Nauk Podstawowych) Międzynarodowej Studenckiej Konferencji Naukowej Młodych Naukowców i Studenckich Kół Naukowych Uczelni Medycznych we Wrocławiu

E. Promotor lub opiekun prac magisterskich:

Joanna Szeluch-Włochal. Badanie polimorfizmu 2R/3R genu kodującego syntazę tymidylanową we krwi oraz guzach pacjentów chorych na raka jelita grubego. 2013 (**Promotor**)

Magdalena Warzecha. Analiza wybranych markerów metylacyjnych w raku jelita grubego. 2014 (**Opiekun**)

Barbara Bosak. Badania metylacji DNA w kontekście uczulenia na jad owadów błonkoskrzydłych. 2016 (**Promotor**)

Aleksandra Teresa Podleśny. Analiza metylacji wybranych genów w raku żołądka. 2019 (**Promotor**)

Paulina Żebrowska. Zaburzenia metylacji DNA w rakach przewodu pokarmowego. 2019 (**Opiekun**)

7. Oprócz kwestii wymienionych w pkt. 1-6, wnioskodawca może podać inne informacje, ważne z jego punktu widzenia, dotyczące jego kariery zawodowej

A. Własna analiza bibliometryczna dorobku

Analizę dla czasopism z *impact factor* przeprowadziłem na podstawie danych pobranych z Web of Science (63 artykuły, Web of Science ResearcherID: HCH-8331-2022, dn. 02.05.2023) oraz Polskiej Platformy Medycznej UMW (<https://ppm.umw.edu.pl>; dn. 02.05.2023).

Mój całkowity dorobek naukowy, obejmujący lata 2006-2023, stanowi 63 publikacje pełnotekstowych, 2 rozdziały w książkach oraz 54 streszczeń ze zjazdów międzynarodowych i krajowych. Z 63 publikacji pełnotekstowych 61 posiada „*impact factor*” (w tym 54 prace oryginalne i 7 prac przeglądowych).

Sumaryczna punktacja za wszystkie publikacje pełnotekstowe wynosi: **IF = 210,977 oraz KBN/MNiSW = 3508 pkt.**

Punktacja cyklu przedłożona jako rozprawa habilitacyjna obejmuje 6 publikacji oryginalnych i 1 pracę przeglądową (**IF = 30,942 KBN/MNiSW = 440 pkt.**).

Aktualna liczba cytowań moich prac (bez autocytowań) wynosi **707, indeks Hirsch’a 16** (wg *Web of Science Core Collection*, z dnia 11.05.2025) - **rycina 1.**

W 26% prac jestem pierwszym autorem. Dokładniejszy rozkład pozycji autorskich przedstawia **rycina 2.**

Rozkład prac publikowanych rocznie przedstawia **rycina 3.**

Rozkład współczynnika *impact factor* dla 63 publikacji przedstawia **rycina 4.**

Results for HCH-8331-2022 ... > Citation Report: HCH-8331-2022 (Author Identifiers)

Citation Report

Q HCH-8331-2022 (Author Identifiers)

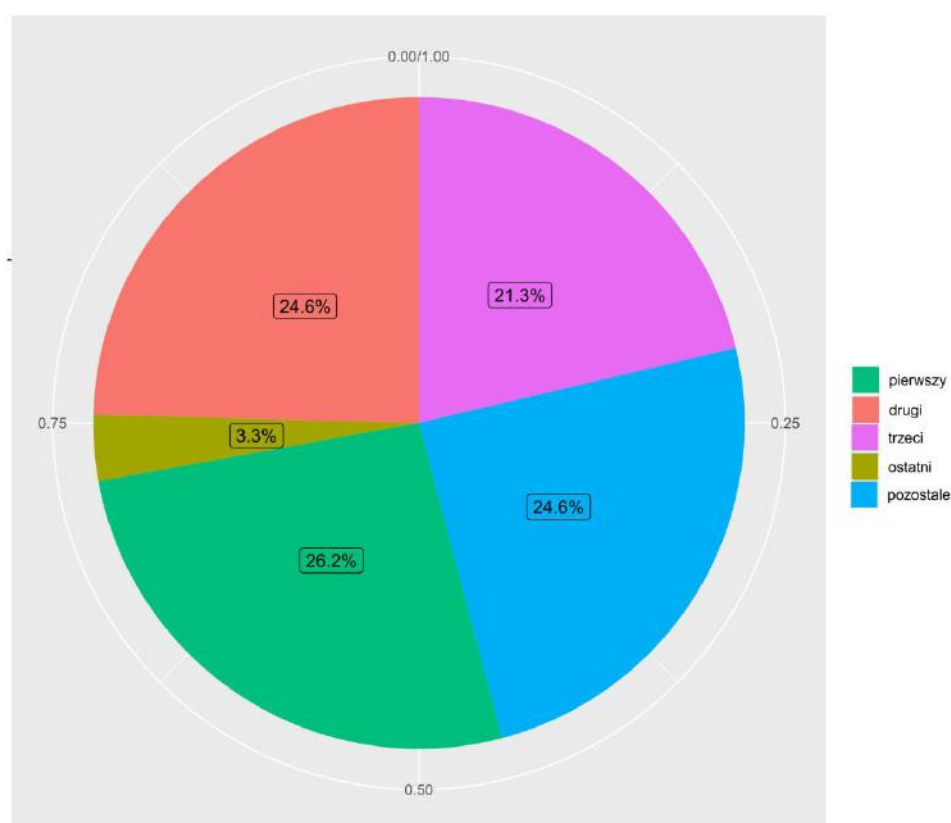
Analyze Results

Create Alert

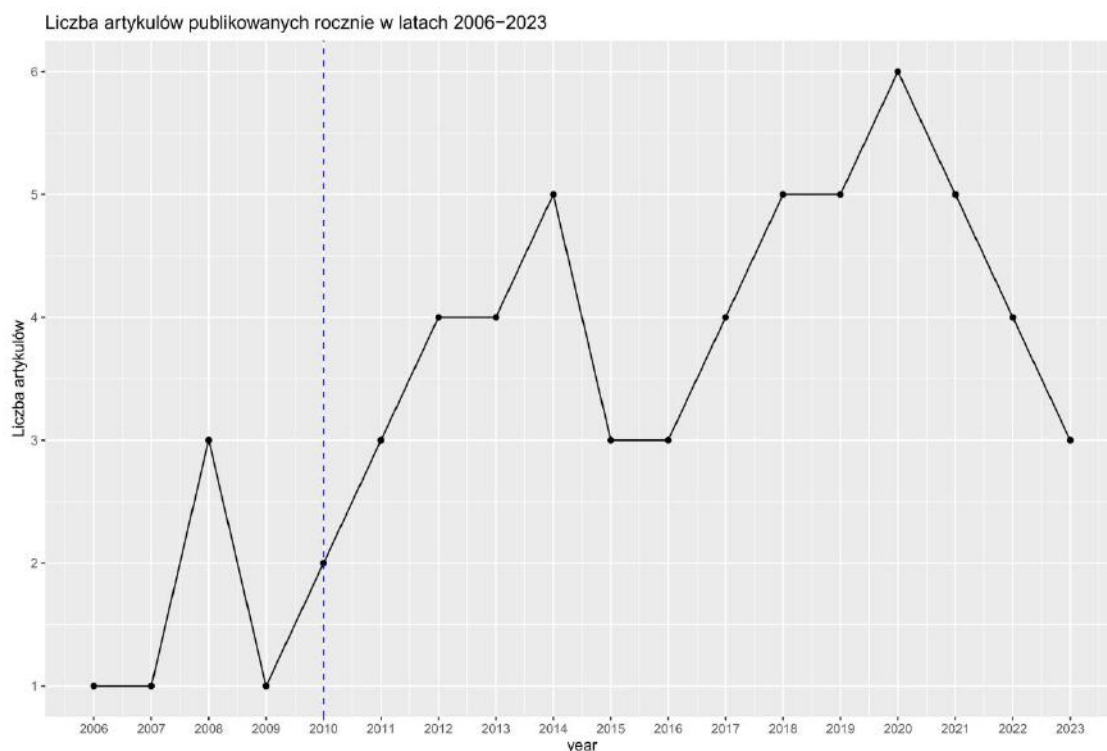
Export Full Report



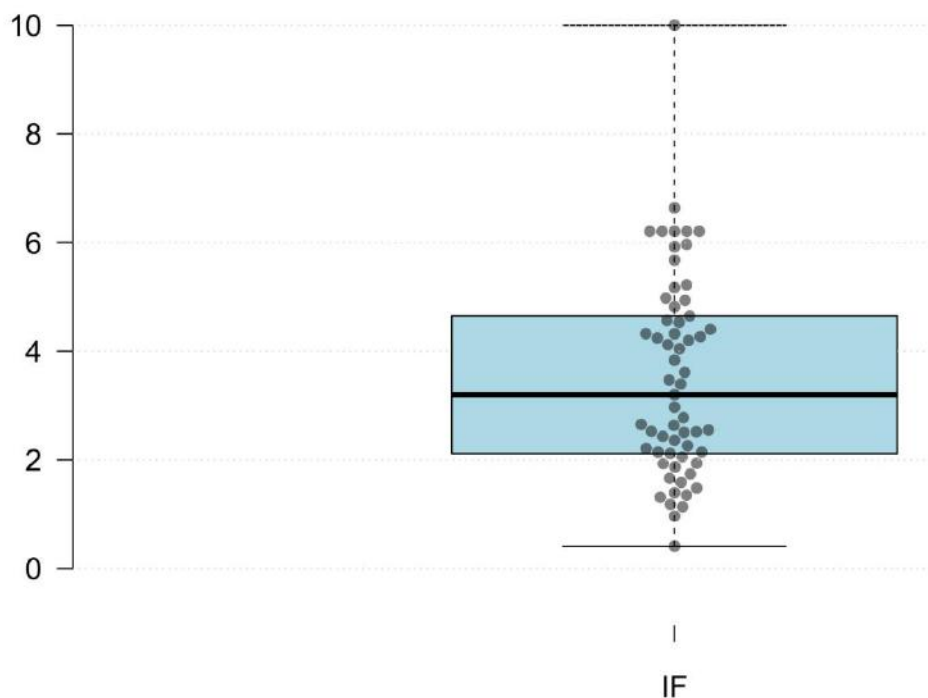
Rycina 1. Raport cytacji z bazy Web of Science (Web of Science ResearcherID: HCH-8331-2022, dn. 02.05.2025) dla 63 artykułów opublikowanych w latach 2006-2023



Rycina 2. Rozkład pozycji autorskich dla 63 artykułów opublikowanych w latach 2006-2023



Rycina 3. Roczny rozkład 63 artykułów opublikowanych w latach 2006-2023. Niebieska, pionowa linia wskazuje datę obrony doktoratu.



Rycina 4. Rozkład współczynnika impact factor dla 63 artykułów opublikowanych w latach 2006-2023.

B. Granty

1. Rak jelita grubego – niestabilność metylacyjna, polimorfizmy w genach kodujących metabolizm ksenobiotyków oraz genów kodujących białka naprawy DNA nr1423/PO1/2007/32, 2007-2010, **(wykonawca)**
2. Badania nad nowymi mechanizmami molekularnymi w etiologii sporadycznego raka jelita grubego – poszukiwanie i charakterystyka dotychczas niezidentyfikowanych obszarów genomu o nieprawidłowej metylacji, grant 736/NCN/2010, 2010 – 2013, **(wykonawca)**
3. Immunoterapia jadem owadów błonkoskrzydłych jako model do badań nad epigenetycznymi mechanizmami immunotolerancji indukowanej w trakcie immunoterapii swoistym alergenem, GR 796/NCN/2013 – 2016, **(wykonawca)**
4. Wykorzystanie innowacyjnej technologii szczypiec optycznych w celu opracowania mało inwazyjnej terapii celowanej chłoniaków. LIDER/016/275/L-5/13/NCBR/2014 2015-2018, **(wykonawca)**
5. Rola stresu w ujęciu koncepcji obciążenia allostatycznego w etiologii schizofrenii, IP2015 052474, 2016 – 2019, **(wykonawca)**
6. Modyfikowane egzosomy pochodzenia nowotworowego jako nośniki cząsteczek modulujących środowisko nowotworowe. NCN SONATA BIS 8 UMO-2018/30/E/NZ5/00711 2019-2024, **(wykonawca)**
7. Punkty kontroli immunologicznej w niedrobnokomórkowym raku płuc – badanie genetycznych i epigenetycznych mechanizmów regulacji ekspresji mRNA i białka. 2019/33/B/NZ5/03029 2020-2023, **(wykonawca)**
8. Analiza transkryptomu w potrójnie ujemnym raku piersi - typowanie molekularne a odpowiedź na terapię adiuwantową, subwencja Uniwersytetu Medycznego we Wrocławiu, SUBK.A290.23.038, 2023-2024, **(kierownik)**

C. Staże zagraniczne

2004 . Półroczny staż w ramach stypendium Socrates-Erasmus: Department of Microbiology, Groningen (Holandia). Kierownik: Prof. Lubbert Dijkhuizen

2010, 2011 - Miesięczne staże w Zakładzie Genetyki Molekularnej, Wydział Antropologii i Genetyki Człowieka, Uniwersytet w Tübingen, Niemcy. Kierownik: Prof. Nicolas Blin

2015 - Trzymiesięczny staż w Instytucie Genetyki Człowieka, Uniwersytet Saarland, Homburg (Niemcy). Kierownik: Prof. Eckart Meese

D. Praca na rzecz czasopism międzynarodowych

- Od 2020 „Guest Editor” w wydaniach specjalnych *International Journal of Molecular Sciences* (MDPI): Epigenetic Mechanisms and Human Pathology, Epigenetic Mechanisms and Human Pathology 2.0, Epigenetic Mechanisms and Human Pathology 3.0, Epigenetic Mechanisms and Human Pathology 4.0
- Od 2021 „Topical Advisory Panel Member” w czasopiśmie *International Journal of Molecular Sciences* (MDPI)
- Od 2022 „Review Editor” w *Cancer Molecular Targets and Therapeutics* (Frontiers)

E. Nagrody

2014 - Nagroda Rektora za ważne osiągnięcia w pracy naukowej za cykl prac dotyczących badań nowych zmian genetycznych i epigenetycznych w etiologii raka jelita grubego opublikowanych w latach 2010-2012

2015 - Nagroda Rektora za cykl prac dotyczących badania zaburzeń metylacji DNA w sporadycznym raku jelita grubego opublikowanych w latach 2012-2013

2015 - Nagroda Zespołowa Rektora za cykl publikacji pt. „Znaczenie zmian genetycznych w powstawaniu, rozwoju oraz diagnostyce i profilaktyce nowotworów”

2015 - "Złota Synapsa" nagroda zespołowa za najbardziej znaczący artykuł o tematyce psychiatrycznej (przyznawana przez Medforum)

2016 - "Złota Synapsa" nagroda zespołowa za najbardziej znaczący artykuł o tematyce psychiatrycznej (przyznawana przez Medforum)

2018 - Nagroda Indywidualna Rektora za publikację w czasopiśmie z bazy JCR

2019 - Nagroda Indywidualna Rektora za publikację w czasopiśmie z bazy JCR

2021 - Nagroda Zespołowa Rektora za publikację w czasopiśmie z bazy JCR

F. Omówienie pozostałych osiągnięć naukowo-badawczych

Brałem udział w realizacji wielu projektów naukowych, które nie weszły w skład osiągnięcia. Główne tematy badań poza tematyką cyklu habilitacyjnego, w które byłem zaangażowany, ogniskowały się na zagadnieniach wypunktowanych poniżej. Zasadniczo można je podzielić na dwie gałęzie:

- analizy oparte o transkryptom z wykorzystaniem uczenia maszynowego i nienadzorowanego grupowania (punkt 1 oraz 2)

- analizy wpływu mononukleotydowych wariantów genomowych na aktywność białek lub ryzyko zachorowania na raka (punkt 3 oraz 4)

1) Badania algorytmów opartych o uczenie maszynowe w detekcji wybranych nowotworów złośliwych

Shah Aftab Ali, Leidinger Petra, Backes Christina, Keller Andreas, **Karpiński Paweł**, Sąsiadek Maria M., Blin Nikolaus, Meese Eckart: A set of specific miRNAs is connected with murine and human gastric cancer, *Genes Chromosomes & Cancer*, 2013, vol. 52, nr 3, s. 237-249

Pawełka Dorota, Łaczmńska Izabela, **Karpiński Paweł**, Supplitt Stanisław, Witkiewicz Wojciech, Knychalski Bartłomiej, Pelak Joanna, Żebrowska Paulina, Łaczmński Łukasz: Machine-learning-based analysis identifies miRNA expression profile for diagnosis and prediction of colorectal cancer: a preliminary study, *Cancer Genomics & Proteomics*, 2022, vol. 19, nr 4, s. 503-511

Donizy Piotr, Krzyżiński Mateusz, Markiewicz Anna, **Karpiński Paweł**, Kotowski Krzysztof, Kowalik Artur, Orłowska-Heitzman Jolanta, Romanowska-Dixon Bożena, Biećek Przemysław, Hoang Mai P.: Machine learning models demonstrate that clinicopathologic variables are comparable to gene expression prognostic signature in predicting survival in uveal melanoma, *European Journal of Cancer*, 2022, vol. 174, s. 251-260

2) Analiza transkryptomu w kontekście alergii na jad owadów błonkoskrzydłych

Karpiński Paweł, Skiba Paweł, Kosińska Magdalena, Rosiek-Biegus Marta, Królewicz Emilia, Blin Nikolaus, Meese Eckart, Panaszek Bernard, Nittner-Marszalska Marita, Sąsiadek Maria Małgorzata: Genome-wide analysis of gene expression after one year of venom immunotherapy, *Immunology Letters*, 2018, vol. 204, s. 23-28

Karpiński Paweł, Kahraman M., Ludwig N., Skiba Paweł, Kosińska Magdalena, Rosiek-Biegus Marta, Królewicz Emilia, Panaszek Bernard, Nittner-Marszalska Marita, Blin Nikolaus, Keller A., Meese E., Sąsiadek Maria M.: Limited long-term impact of insect venom immunotherapy on the micro-RNA landscape in whole blood, *Journal of Investigational Allergology and Clinical Immunology*, 2019, vol. 29, nr 3, s. 206-212

3) Bioinformatyczne modelowanie wpływu patogennych wariantów mononukleotydowych na właściwości i aktywność białek

Ślęzak Ryszard, Łukieńczuk Tadeusz, Noczyńska Anna, **Karpiński Paweł**, Lebioda Arleta, Misiak Błażej, Sąsiadek Maria M.: A novel p.E311K mutation of thyroid receptor beta gene in resistance to thyroid hormone syndrome, inherited in autosomal recessive trait, *Hormone and Metabolic Research*, 2012, vol. 44, nr 9, s. 704-707

Śmigiel Robert, Kostrzewa Grażyna, Kosińska Joanna, Pollak Agnieszka, Stawiński Piotr, Szmida Elżbieta, Błoch Michał, Szymańska Krystyna, **Karpiński Paweł**, Sąsiadek Maria M., Płoski Rafał: Further evidence for GRIN2B mutation as the cause of severe epileptic encephalopathy, *American Journal of Medical Genetics Part A*, 2016, vol. 170, nr 12, s. 3265-3270

Stembalska Agnieszka, Kłapecki Jakub, Pławski Andrzej, **Karpiński Paweł**: In-silico analysis of Thr767Ile pathogenic variant in the MSH6 gene in family with endometrial cancer, *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 2019, vol. 238, s. 54-57

4) Badania polimorficznych wariantów mononukleotydowych (SNP) w kontekście ryzyka zachorowania na choroby nowotworowe

Łaczmańska Izabela, Gil Justyna, **Karpiński Paweł**, Stembalska Agnieszka, Kozłowska Joanna, Busza Halina, Trusewicz Alicja, Pesz Karolina, Ramsey David, Schlade-Bartusiak Kamila, Blin Nikolaus, Sąsiadek Maria Małgorzata: Influence of polymorphisms in xenobiotic-metabolizing genes and DNA-repair genes on diepoxybutane-induced SCE frequency, *Environmental and Molecular Mutagenesis*, 2006, vol. 47, nr 9, s. 666-673

Łaczmańska Izabela, Gil Justyna, **Karpiński Paweł**, Stembalska Agnieszka, Trusewicz Alicja, Pesz Karolina, Ramsey David, Schlade-Bartusiak Kamilla, Blin Nikolaus, Sąsiadek Maria Małgorzata: Polymorphism in nucleotide excision repair gene XPC correlates with bleomycin-induced chromosomal aberrations, *Environmental and Molecular Mutagenesis*, 2007, vol. 48, nr 8, s. 666-671

Karpiński Paweł, Myszkowski Aleksander, Ramsey David, Misiak Błażej, Gil Justyna, Łaczmańska Izabela, Grzebieniak Zygmunt, Sebzda Tadeusz, Śmigiel Robert, Stembalska Agnieszka, Sąsiadek Maria M.: Polymorphisms in methyl-group metabolism genes and risk of sporadic colorectal cancer with relation to the CpG island methylator phenotype, *Cancer Epidemiology*, 2010, vol. 34, nr 3, s. 338-34

Gil Justyna, Ramsey David, Stembalska Agnieszka, **Karpiński Paweł**, Pesz Karolina A., Łaczmańska Izabela, Leszczyński Przemysław, Grzebieniak Zygmunt, Sąsiadek Maria Małgorzata: The C/A polymorphism in intron 11 of the XPC gene plays a crucial role in the

modulation of an individual's susceptibility to sporadic colorectal cancer, *Molecular Biology Reports*, 2012, vol. 39, nr 1, s. 527-534

Pesz Karolina A., Bieniek Andrzej, Gil Justyna, Łaczmańska Izabela, **Karpiński Paweł**, Makowska Izabela, Trusewicz Alicja, Sasiadek Maria M.: Polymorphisms in nucleotide excision repair genes and basal cell carcinoma of the skin, *International Journal of Dermatology*, 2014, vol. 53, nr 12, s. 1474-1477

Łaczmańska Izabela, **Karpiński Paweł**, Gil Justyna, Łaczmański Łukasz, Makowska Izabela, Bębenek M., Ramsey D., Sasiadek Maria M.: The PTPN13 Y2081D (T>G) (rs989902) polymorphism is associated with an increased risk of sporadic colorectal cancer, *Colorectal Disease*, 2017, vol. 19, nr 7, O272-O278

Nguyen-Dumont Tu, Myszka Aleksander, **Karpiński Paweł**, Sasiadek Maria M., Akopyan Hayane, Hammet Fleur, Tsimiklis Helen, Park Daniel J., Pope Bernard J., Ślęzak Ryszard, Kitsera Nataliya, Siekierzyńska Aleksandra, Southey Melissa C.: FANCM and RECQL genetic variants and breast cancer susceptibility: relevance to South Poland and West Ukraine, *BMC Medical Genetics*, 2018, vol. 19, art.12 [7 s.]

Wagner Marta, Tupikowski Krzysztof, Jasek Monika, Tomkiewicz Anna, Witkiewicz Agata, Ptaszkowski Kuba, **Karpiński Paweł**, Zdrojowy Romuald, Hałoń Agnieszka, Karabon Lidia: SNP-SNP interaction in genes encoding PD-1/PD-L1 axis as a potential risk factor for clear cell renal cell carcinoma, *Cancers*, 2020, vol. 12, nr 12, art.3521 [15 s.]

Moksud Nafeesa, Wagner Marta, Pawełczyk Konrad Jan, Porębska Irena Maria, Muszczyńska-Bernhard Beata, Kowal Aneta, Wiśniewski Andrzej, Kosacka Monika Maria, Kończak Julia, **Karpiński Paweł**, Frydryk Dominik, Andrzejczak Anna, Karabon Lidia, Kuśnierczyk Piotr, Jasek Monika: Common inherited variants of PDCD1, CD274 and HAVCR2 genes differentially modulate the risk and prognosis of adenocarcinoma and squamous cell carcinoma, *Journal of Cancer Research and Clinical Oncology*, 2023, DOI:10.1007/s00432-023-04602-8



.....
(podpis wnioskodawcy)