



Uniwersytet im. Adama Mickiewicza w Poznaniu

Wydział Biologii,

Instytut Biologii i Ewolucji Człowieka

Klaudia Staszak

ROZPRAWA DOKTORSKA

**Powiązania pomiędzy retrogenami i nowotworami –
kontekst ewolucyjny**

Promotor: prof. dr hab. Izabela Makałowska

Poznań, 2025



Adam Mickiewicz University in Poznań

Faculty of Biology,

Institute of Human Biology and Evolution

Klaudia Staszak

DOCTORAL DISSERTATION

**Relationships between retrogenes and cancer –
evolutionary context**

Supervisor: prof. dr hab. Izabela Makałowska

Poznań, 2025

Dziękuję mojej Pani Promotor, prof. dr hab. Izabeli Makalowskiej,
za możliwość realizacji pracy doktorskiej, cenne wskazówki
oraz pomoc w przygotowaniu niniejszej rozprawy.

Dziękuję dr Magdalenie Kubiak,
za wsparcie w trakcie realizacji projektu.

Dziękuję Klaudii Samorowskiej, Anecie Łyś, Sylwii Szvec, Alicji Durskiej i
Wenesie Woźniak,
z którymi miałam ogromną przyjemność współpracować. Wasza życzliwość
i niezawodne poczucie humoru rozświeślały nawet najtrudniejsze chwile.

Dziękuję Rodzinie i Przyjaciółom,
jesteście moją największą wartością.

Dziękuję Kacprowi i Fuszemu,
jesteście najwspanialszymi towarzyszami mojej codzienności.

*Pracę dedykuję **Mojej Mamie**,
która osiągnęła dużo więcej niż tytuł naukowy.
Jesteś najlepsza, nigdy o tym nie zapominaj.*

FINANSOWANIE BADAŃ

Badania w ramach niniejszej pracy zostały sfinansowane z następujących źródeł:

1. Grant PRELUDIUM-21 (2022/45/N/NZ2/04105) „*Retrogeny jako istotni gracze w nowotworowych szlakach regulatorowych*” (kierownik: Klaudia Staszak)



N A R O D O W E C E N T R U M N A U K I

2. Minigrant ID-UB „*Analiza funkcjonalna retrogenów związanych z nowotworzeniem w ludzkich liniach komórkowych*” (kierownik: Klaudia Staszak)



INICJATYWA DOSKONAŁOŚCI
UCZELNIA
BADAWCZA

3. Grant Doktorancki Uniwersytetu Jutra (kierownik: Klaudia Staszak)



UNIWERSYTET
JUTRA

SPIS TREŚCI

STRESZCZENIE.....	1
ABSTRACT.....	2
WYKAZ SKRÓTÓW.....	3
1. WSTĘP.....	7
1.1. Retrogeny.....	7
1.2. Ewolucja i funkcje retrogenów.....	8
1.3. Retrogeny, nowotwory i ewolucja.....	11
1.3.1. Mechanizmy ewolucyjne i nowotwory.....	12
1.3.2. Retrogeny i nowotwory.....	16
2. CEL PRACY.....	21
3. MATERIAŁY I METODY.....	22
3.1. Analizy bioinformatyczne.....	22
3.1.1. Dane transkryptomiczne (RNA-Seq).....	22
3.1.2. Adnotacje retrokopii.....	24
3.1.3. Przetwarzanie danych RNA-Seq.....	24
3.1.4. Analizy ekspresji różnicowej.....	26
3.1.5. Analizy korelacji ekspresji.....	27
3.1.6. Potencjalne miejsca wiązania miRNA.....	27
3.1.7. Wizualizacja danych.....	27
3.1.8. Analiza filogenetyczna.....	27
3.1.9. Testy statystyczne.....	28
3.2. Analizy eksperymentalne.....	28
3.2.1. Materiał biologiczny.....	28
3.2.2. Hodowle komórkowe.....	28
3.2.3. Izolacja RNA i odwrotna transkrypcja.....	29
3.2.4. PCR.....	29
3.2.5. Wyciszanie ekspresji.....	31
3.2.6. qPCR.....	31
3.2.7. Test żywotności.....	31
4. WYNIKI.....	32
4.1. Identyfikacja retrokopii o zmienionej ekspresji w nowotworach człowieka, psa i kury.....	32
4.1.1. Szacowanie poziomu ekspresji retrokopii.....	32
4.1.2. Identyfikacja retrogenów o zmienionym poziomie ekspresji w nowotworach człowieka.....	36
4.1.3. Retrogeny o zmienionej ekspresji w różnych typach i stadiach nowotworów człowieka.....	38
4.1.4. Identyfikacja retrogenów o zmienionym poziomie ekspresji w nowotworach psa.....	39
4.1.5. Identyfikacja retrogenów o zmienionym poziomie ekspresji w nowotworze kury.....	40
4.2. Analiza ewolucyjna retrogenów o zmienionej ekspresji w nowotworach.....	41
4.2.1. Retrogeny związane z nowotworami wspólne dla człowieka, psa i kury.....	41
4.2.2. Retrogeny specyficzne dla badanych gatunków kręgowców.....	45
4.3. Analiza eksperymentalna retrogenów o zmienionej ekspresji w nowotworach.....	46
4.3.1. Ekspresja w nowotworach i zdrowych tkankach/narządach <i>RHOB</i> , <i>HSPA2</i> i <i>CALML5</i>	46
4.3.2. Ekspresja retrogenów specyficznych gatunkowo w nowotworach oraz zdrowych tkankach i narządach.....	51
4.4. Profil ekspresji retrogenów a różnicowanie typów nowotworów.....	56
4.5. Wielkoskalowe analizy retrokopii.....	57
4.5.1. Ekspresja retrokopii w danych TCGA.....	57
4.5.2. Retrogeny o zróżnicowanej ekspresji w nowotworach człowieka – dane TCGA.....	59
4.5.3. Retrogeny związane z ewolucją komórek w obrębie guza nowotworowego oraz z odpowiedzią na terapię.....	66
4.5.4. Retrogeny o zróżnicowanej ekspresji pochodzące od genów powiązanych z nowotworami.....	69
4.6. Eksperymentalne analizy wpływu retrogenów na ekspresję genów rodzicielskich i żywotność komórek.....	76
4.6.1. Wpływ retrogenów na poziom ekspresji genów rodzicielskich.....	76
4.6.2. Wpływ wyciszenia ekspresji retrogenów na żywotność komórek.....	81

4.7. Identyfikacja retrogenów związanych zarówno z nowotworami jak i rozwojem łożyska człowieka.....	82
5. DYSKUSJA.....	85
6. WNIOSKI.....	96
SPIS RYCIN.....	97
SPIS TABEL.....	99
LITERATURA.....	101
ANEKS.....	112

STRESZCZENIE

Proces retropozycji, w którym dojrzały mRNA podlega odwrotnej transkrypcji do cDNA i ponownej integracji z genomem, skutkuje powstaniem dodatkowej kopii istniejącego genu (genu rodzicielskiego). Kopie takie, nazywane retrokopiami lub retrogenami, często osadzają się w intronach innych genów, zwanych genami gospodarzy. Retrokopie, dotychczas opisywane jako нефunkcjonalne elementy, są obecnie uznawane za istotne w kontekście ewolucji, a ich rola w kształtowaniu genomów, transkryptomów i proteomów staje się coraz bardziej widoczna. Rosnąca liczba danych z wysokoprzepustowych eksperymentów umożliwia identyfikację nowych funkcjonalnych retrogenów. Bardzo często regulują one poziom ekspresji swoich genów rodzicielskich działając jako gąbki miRNA czy antysensowne transkrypty oraz mogą brać udział w regulacji epigenetycznej, modyfikacji alternatywnego splicingu czy w formacji transkryptów fuzyjnych. Zjawisko retropozycji było szczególnie wzmożone podczas ewolucji naczelnych. Związek między wyjątkowo dużą liczbą retrokopii człowieka, w porównaniu do innych gatunków zwierząt, a transformacją nowotworową u naszego gatunku, nie jest dobrze poznany. Dlatego głównym celem pracy było sprawdzenie na ile intensywny proces retropozycji u naczelnych może być powiązany ze wzmożeniami, w porównaniu z innymi gatunkami, procesami nowotworzenia u człowieka.

Materiał do badań stanowiły dane transkryptomiczne człowieka, psa oraz kury. Ekspresję retrogenów związanych z nowotworami zweryfikowano także eksperymentalnie w ludzkich liniach komórkowych, w panelu cDNA pochodzącego z różnych narządów badanych gatunków oraz w świeżo mrożonych próbach nowotworowych i tkankach okalających pobranych od psów i szczurów. Ekspresję wybranych retrogenów wyciszono w ludzkich liniach komórkowych, by zbadać wpływ na ekspresję genu rodzicielskiego lub genu gospodarza.

W oparciu o wyniki analizy ekspresji różnicowej zidentyfikowano 135 retrogenów o zmienionej ekspresji w nowotworach człowieka. Spośród nich ponad trzykrotnie więcej należy do retrogenów specyficznych dla naczelnych. Retrogeny *HSPA2*, *RHOB* i *CALML5* wskazano jako retrogeny powiązane z nowotworami odległych ewolucyjnie gatunków kręgowców. Ponadto, wskazano retrogeny powiązane z nowotworami, które są specyficzne dla badanych gatunków kręgowców a także retrogeny, które mogą działać jako ważne niekodujące RNA regulujące geny znane jako pełniące role w komórkach nowotworowych. Wyniki uzyskane w ramach niniejszej pracy umożliwiły po raz pierwszy w tak dużej skali spojrzeć na profil ekspresji retrogenów w nowotworach w kontekście ewolucyjnym.

ABSTRACT

The process of retroposition, in which mature mRNA is reverse transcribed into cDNA and reintegrated into the genome, results in the creation of an additional copy of an existing gene (parental gene). Such copies, called retrocopies or retrogenes, are often embedded in introns of other genes, called host genes. Retrocopies, previously described as non-functional elements, are now recognized as important in the context of evolution, and their role in shaping genomes, transcriptomes and proteomes is becoming increasingly evident. An increasing amount of data from high-throughput experiments is allowing the identification of new functional retrogenes. Very often, they regulate the expression level of their parental genes by acting as miRNA sponges or antisense transcripts, and may participate in epigenetic regulation, modification of alternative splicing or in the formation of fusion transcripts. The phenomenon of retroposition has been particularly intensified during the evolution of primates. The relationship between the exceptionally high number of human retrocopies, compared to other animal species, and neoplastic transformation in our species is not well understood. Therefore, the main aim of the study was to investigate to what extent the intensive process of retroposition in primates can be linked to increased, compared to other species, processes of human carcinogenesis.

The study material consisted of transcriptomic data from human, dog and chicken. The expression of retrogenes associated with tumors was also experimentally verified in human cell lines, in a panel of cDNA from different organs of the studied species and in fresh-frozen tumor samples and surrounding tissues collected from dogs and rats. The expression of selected retrogenes was silenced in human cell lines to examine the effect on the expression of the parental gene or the host gene.

Based on the results of the differential expression analysis, 135 retrogenes with altered expression in human tumors were identified. Of these, more than three times as many belong to primate-specific retrogenes. The retrogenes *HSPA2*, *RHOB* and *CALML5* were identified as retrogenes associated with cancers of evolutionarily distant vertebrate species. In addition, we identified cancer-associated retrogenes that are specific to the vertebrate species studied, as well as retrogenes that may act as important non-coding RNAs regulating genes known to play a role in cancer cells. The results of this work allow, for the first time on such a large scale, to consider the expression profile of retrogenes in cancers in an evolutionary context.

WYKAZ SKRÓTÓW

ActB – z ang. *actin beta*)
APTX – z ang. *aprataxin*
BRCA – z ang. *breast carcinoma*
CALM3 – z ang. *calmodulin 3*
CALML5 – z ang. *calmodulin like 5*
cDNA – z ang. *complementary DNA*
cis-NAT – z ang. *cis natural antisense transcript*
CMC – z ang. *canine mammary carcinoma*
COAD – z ang. *colon adenocarcinoma*
CSCs – z ang. *cancer stem cells*
CSMD3 – z ang. *CUB and Sushi multiple domains 3*
CTNNA1P1 – z ang. *catenin alpha 1 pseudogene 1*
DCAF12 – z ang. *DDB1 and CUL4 associated factor12*
DEG – z ang. *differential expressed gene*
EEF1A1 – z ang. *eukaryotic translation elongation factor 1 alpha 1*
EEF1A1P13 – z ang. *eukaryotic translation elongation factor 1 alpha 1 pseudogene 13*
ENO2 – z ang. *enolase 2*
ER+ – z ang. *estrogen receptor positive*
fgf4 – z ang. *fibroblast growth factor 4*
FMO5 – z ang. *flavin containing dimethylaniline monooxygenase 5*
FTH1P3 – z ang. *ferritin heavy chain 1 pseudogene 3*
GBM – z ang. *glioblastoma multiforme*
H3P6 – z ang. *H3 histone pseudogene 6*
HCC – z ang. *hepatocellular carcinoma*
HMGAI – z ang. *high mobility group AT-hook 1*
HMGB1 – z ang. *high mobility group box 1*
HMGB1P5 – z ang. *high mobility group box 1 pseudogene 5*
HMGB1P6 – z ang. *high mobility group box 1 pseudogene 6*
HMM-HA – z ang. *high-molecular-mass hialuronian*
HSPA2 – z ang. *heat shock protein family A (Hsp70) member 2*
HSPA8 – z ang. *heat shock protein family A (Hsp70) member 8*
INTS6 – z ang. *integrator complex subunit 6*
INTS6P1 – z ang. *integrator complex subunit 6 pseudogene 1*
ISF – z ang. *inward stranded forward*

ISR – z ang. *inward stranded reverse*
 IU – z ang. *inward unstranded*
 KIRC – z ang. *kidney renal clear cell carcinoma*
 LEP – z ang. *leptin*
 LIHC – z ang. *liver hepatocellular carcinoma*
 lncRNA – z ang. *long non-coding RNA*
 log2FC – z ang. *log2 of Fold Change*
 LUSC – z ang. *lung squamous cell carcinoma*
 MALAT1 – z ang. *metastasis associated lung adenocarcinoma transcript 1*
 MCF-7 – linia komórkowa ludzkiego estrogenozależnego gruczolaka piersi
 MCF-12F – linia komórkowa ludzkiego gruczołu mlekowego
 miRNA – z ang. *micro RNA*
 MRC-5 – linia komórkowa ludzkich fibroblastów płuca
 mRNA – z ang. *messenger RNA*
 MSL3P1 – z ang. *MSL complex subunit 3B*
 NANOGP8 – z ang. *Nanog homeobox retrogene P8*
 NAT – z ang. *natural antisense transcript*
 NCBI SRA – z ang. *National Center for Biotechnology Information Sequence Read Archive*
 NCI-H1703 – linia komórkowa niedrobnokomórkowego raka płuc
 NKAPL – z ang. *NFkB activating protein like*
 NOPCHAP1 – ang. *NOP protein chaperon 1*
 OV – z ang. *ovary cancer*
 PCBP1 – z ang. *poly(rC) binding protein 1*
 PCNA – z ang. *proliferating cell nuclear antigen*
 PCR – z ang. *polymerase chain reaction*
 PDIA3P1 – z ang. *protein disulfide isomerase family A member 3 pseudogene 1*
 PLEKHA7 – z ang. *pleckstrin homology domain containing A7*
 POU5F1P4/P5 – z ang. *POU class 5 homeobox 1 pseudogene 4/pseudogene 5*
 PPIAP43 – z ang. *peptidylprolyl isomerase A pseudogene 43*
 PPP1R14B – z ang. *protein phosphatase 1 regulatory inhibitor subunit 14B*
 PPP1R14BP3 – z ang. *protein phosphatase 1 regulatory inhibitor subunit 14B pseudogene 3*
 PRAD – z ang. *prostate adenocarcinoma*
 PTEN – z ang. *phosphatase and tensin homolog*
 PTENP1 – z ang. *phosphatase and tensin homolog pseudogene 1*
 pz - par zasad

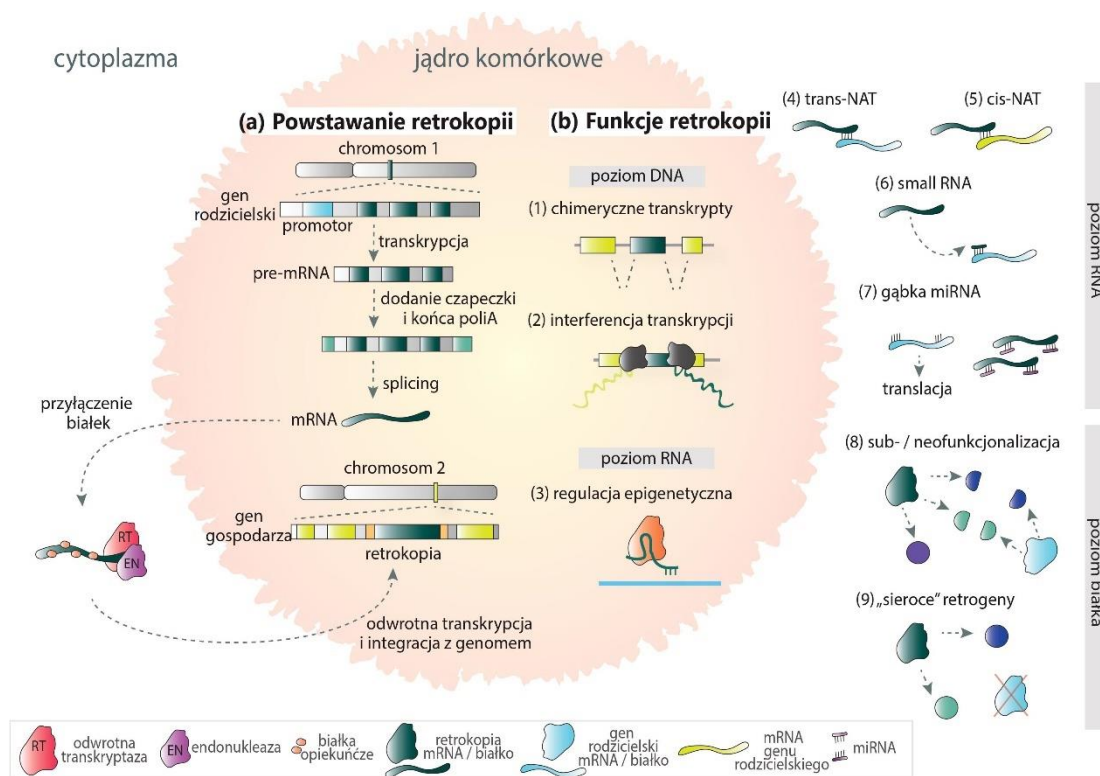
qPCR – z ang. *quantitative polymerase chain reaction*
 READ – z ang. *rectum adenocarcinoma*
 RHOA – z ang. *ras homolog family member A*
 RHOB – z ang. *Ras homolog family member B*
 RHOG – z ang. *ras homolog family member G*
 RNA-Seq – z ang. *RNA sequencing*
 RNF113B – z ang. *ring finger protein 113B*
 RPL10P9 – z ang. *ribosomal protein L10 pseudogene 9*
 RPL36A – z ang. *ribosomal protein L36a*
 RPL9 – z ang. *ribosomal protein L9*
 RPS18 – z ang. *ribosomal protein S18*
 Rps23r1 – z ang. *ribosomal protein S23, retrogene 1*
 RPS28P7 – z ang. *ribosomal protein S28 pseudogene 7*
 rRNA – z ang. *ribosomal RNA*
 RT-PCR – ang. *reverse transcribed polymerase chain reaction*
 SOCS2 – z ang. *suppressor of cytokine signaling 2*
 SF – z ang. *stranded forward*
 SR – z ang. *stranded reverse*
 TCGA – z ang. *The Cancer Genome Atlas*
 THSD4 – z ang. *thrombospondin type 1 domain containing 4*
 TNBC – z ang. *triple-negative breast cancer*
 TP53 – z ang. *tumor protein p53*
 TPM – z ang. *transcripts per million*
 trans-NAT – z ang. *trans natural antisense transcript*
 TSEEN – z ang. *tumor-specifically expressed, evolutionarily novel*
 UCEC – z ang. *uterine corpus endometrial carcinoma*
 UCP1 – z ang. *uncoupling protein 1*
 UTP14C – z ang. *UTP14C small subunit processome component*
 UTR – z ang. *untranslated regions*
 YWHAQ – z ang. *tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein theta*

1. WSTĘP

1.1. Retrogeny

Zmieniające się warunki środowiska determinują pojawienie się nowych cech adaptacyjnych, umożliwiających przetrwanie gatunków. Występowanie tych charakterystycznych cech fenotypowych może być, między innymi, efektem działania nowopowstałych genów. Duplikacja genów, dostarczając "materiał" dla działania doboru naturalnego, jest głównym mechanizmem związanym z powstawaniem nowych genów (Kaessmann i in., 2009; Pegueroles i in., 2013). Przez długi czas uważano, że funkcjonalne kopie genów powstają głównie poprzez duplikację regionów DNA. Kolejne badania wykazały, że duplikacja oparta na RNA (proces retropozycji) może być jeszcze bardziej znaczącym źródłem nowych genów, pełniących istotne funkcje w komórce (Kaessmann i in., 2009; Kubiak i in., 2020; Kubiak & Makałowska, 2017). Kopie te odgrywają ważną rolę w ewolucji molekularnej genomów i zostały nazwane „zarodkami ewolucji” (ang. *seeds of evolution*) (Brosius, 1991).

Proces powstawania retrokopii rozpoczyna się od transkrypcji genu rodzicielskiego (ang. *parental gene*) (Rycina 1a). Dojrzały mRNA przechodzi do cytoplazmy, gdzie przyłączają się do niego białka kodowane przez retrotranspozon L1, odwrotna transkryptaza i endonukleaza kodowane przez ORF2 oraz białka opiekuńcze kodowane przez ORF1. Powstały kompleks jest następnie przenoszony z powrotem do jądra komórkowego, gdzie mRNA genu rodzicielskiego łączy się z uszkodzonym końcem DNA. Zachodzi proces odwrotnej transkrypcji, a powstałe cDNA jest włączane ponownie do genomu jako tzw. retrokopia. Retrokopie są pozbawione intronów i elementów regulatorowych, posiadają ogon poli-A oraz powtórzenia flankujące w miejscu insercji. Kopie te, mimo początkowego braku elementów regulatorowych mogą ulegać transkrypcji. Aktywność tę uzyskują poprzez wykorzystanie m.in. promotorów sąsiadujących genów, odległych sekwencji CpG, a czasami nawet części swoich własnych sekwencji. Ponadto, gdy retrokopia wstawiona zostaje do intronu innego genu może wykorzystywać elementy regulatorowe tego tzw. genu gospodarza (ang. *host gene*) (Kaessmann i in., 2009; Vinckenbosch i in., 2006). Retrokopie aktywne transkrypcyjnie nazywane są retrogenami. Ich mRNA przez długi czas uważano za szum transkrypcyjny, z uwagi na ich znaczące podobieństwo do genów rodzicielskich oraz powszechne traktowano jako pseudogeny.



Rycina 1. Mechanizm powstawania retrokopii genów kodujących białka **(a)** oraz możliwe funkcje pełnione w komórce **(b)** (Staszak & Makałowska, 2021).

1.2. Ewolucja i funkcje retrogenów

Gdy powstaje dodatkowa kopia genu kodującego białko, podlega ona mniejszym ograniczeniom ewolucyjnym niż gen, z którego się wywodzi (Magadum i in., 2013). Duplikacja pozwala na swobodną selekcję oczyszczającą, więc kopie mają tendencję do szybszego gromadzenia mutacji, co sprawia, że ewoluują szybciej (Tutar, 2012). Spośród dwóch rodzajów duplikacji, to duplikaty powstałe na drodze retropozycji częściej nabywają nowy wzór ekspresji i funkcji. Duplikacja na poziomie DNA prowadzi do powstania potomnych kopii o takiej samej organizacji genomowej jak gen rodzicielski, z tym, że nie zawsze zduplikowana jest cała sekwencja genu. Duplikaty zawierające całość sekwencji genu, włączając w to rejon regulatorowy, zazwyczaj odzwierciedlają funkcję oraz wzór ekspresji ich przodka (Kaessmann i in., 2009; Vinckenbosch i in., 2006). Ponadto, sugeruje się, że ewolucja przebiega wolniej w obszarach ekson-intron oraz w genach zawierających wiele intronów (Parmley i in., 2007). Co ciekawe, zauważono, że tempo ewolucji w retrokopiach jest największe w obszarach

poprzednich miejsc splicingowych w genie rodzicielskim (Kaessmann i in., 2009; Parmley i in., 2007).

Insercja retrokopii może być źródłem cech gatunkowo specyficznych. Dobrym przykładem jest retrogen *fgf4* (ang. *fibroblast growth factor 4*), który występuje u ras psów o krótkich nogach i jest odpowiedzialny za chondrodysplazję (Parker i in., 2009). Z kolei retrokopia genu cyklofiliny A związana jest z odpornością na HIV u małp ponocnicowatych (Richardson i in., 2014). Zatem zrozumienie różnic międzygatunkowych ma kluczowe znaczenie w badaniach medycznych, ponieważ mogą one być powodem, dla którego wyników badań na zwierzętach nie można przenieść na ludzi. Przykładowo funkcjonalny retrogen *Rps23r1* (ang. *ribosomal protein S23, retrogene 1*), którego rola związana jest z redukcją poziomu beta-amyloidu oraz fosforylacji białka tau w przypadku choroby Alzheimera, jest specyficzny dla myszy i nie występuje u człowieka (Y. Zhang i in., 2009). Liczne analizy wskazują, że proces retropozycji był szczególnie wzmożony wśród ewolucji naczelnych (ang. *burst of retroposition*), co związane jest z występowaniem dużej ilości retrokopii w genomach tego rzędu ssaków (Marques i in., 2005; Navarro & Galante, 2015). Pojawiły się także doniesienia, w których stosując inne metody szacowania ilości zduplikowanych genów pokazano, że u myszy proces retropozycji był równie intensywny (Pan & Zhang, 2007). Te rozbieżności mogą wynikać z różnic w ilości i jakości danych dostępnych dla różnych gatunków.

Retrokopie, początkowo jednoegzonowe, mogą z czasem zyskiwać introny lub dodatkowe eksony. Szcześniak i in. opisali dwa retrogeny, *RNF113B* (ang. *ring finger protein 113B*) i *DCAF12* (ang. *DDB1 and CUL4 associated factor 12*), w których introny powstały w wyniku mutacji i pojawienia się nowych miejsc splicingowych (Szcześniak i in., 2011). Z kolei zespoły Fablet i in. oraz Carelli i in. wskazały przykłady retrogenów, które nabyły nowe eksony (Carelli i in., 2016; Fablet i in., 2009). Gdy retrokopia zostanie włączona w rejon innego genu, może zostać aktywowana podczas transkrypcji genu gospodarza, tworząc chimeryczne transkrypty. Taka integracja retrokopii przyczynia się do powstania nowych form splicingowych, często ze zmodyfikowanym białkiem i specyficznym wzorem ekspresji. Dobrym przykładem może być retrokopia *retro_hsap_4001*¹, zlokalizowana w locus genu *CSMD3* (ang. *CUB and Sushi multiple domains 3*), która jest związana z powstaniem dwóch pierwszych eksonów krótszego wariantu *CSMD3* (*CSMD3-203*). Kompleksowa analiza

¹ Nazwy rozpoczynające się od „retro” pochodzą z bazy danych RetrogeneDB, retrokopie te nie mają swojego odpowiednika lub nazwy w innych bazach.

retrokopii człowieka wskazała 71 transkryptów, które zawierały sekwencje 56 retrokopii. (Kubiak i in., 2020).

Dotychczas opisano wiele przykładów retrogenów, które posiadają potencjał do kodowania białek (Kaessmann i in., 2009; S. Chen i in., 2010; Kubiak i in., 2020) (Rycina 1b). Zidentyfikowano aż 106 znanych genów kodujących białka, które powstały na drodze retropozycji (Rosikiewicz i in., 2017). W toku ewolucji retrogeny te mogą podlegać subfunkcjonalizacji, dzieląc funkcje ze swoim genem rodzicielskim lub nawet zastąpić swoje geny rodzicielskie, mówimy wówczas o retrogenach „sierocych” (ang. *orphan retrogenes*) (Ciomborowska i in., 2013). Jak wspomniano wcześniej, sekwencje retrogenów często ewoluują w szybkim tempie, wskutek czego mogą zyskiwać nowe funkcje, a więc podlegać neofunkcjonalizacji. Funkcje te wynikają z uzyskania nowych wzorców ekspresji, narzuconych przez zawartość sekwencji genomowej otaczającej retrokopię w miejscu wstawienia oraz szybko kumulujące się mutacje (Poliseno i in., 2010).

Gen zagnieżdżony w intronie innego genu może powodować przedwczesne zakończenie transkrypcji genu gospodarza (Kaer i in., 2011). Kolejnym przykładem funkcji retrogenów może być zatem ich udział w interferencji transkrypcji. W zjawisku tym transkrypcja retrogenu znajdującego się na przeciwnej nici DNA wpływa hamująco na transkrypcję innego genu. Przykładem może być zagnieżdżona w intronie genu *ERLIN2* (ang. *ER lipid raft associated 2*) retrokopia *retro_hsap_4044*. Sugeruje się, że ekspresja tej retrokopii może być związana z wcześniejszą terminacją transkrypcji genu *ERLIN2*, a co za tym idzie z występowaniem krótszych form splicingowych tego genu (*ERLIN2-002*, *ERLIN2-006*) (Kubiak i in., 2020).

Na etapie potranskrypcyjnym, RNA pochodzące z retrogeny może wpływać na regulację ekspresji innych genów na poziomie epigenetycznym. Dobry przykład stanowi transkrypt antysensowny do retrokopii *Oct4-pseudogen 5* (*asOct4-pg5*, ang. *antisense to Oct4-pseudogen 5*), który jest częścią kompleksu modulującego na poziomie epigenetycznym aktywność transkrypcyjną genu rodzicielskiego *Oct4* (Hawkins & Morris, 2010).

Retrogeny mogą pełnić także funkcje regulatorowych RNA, działając jako *cis*-NAT (ang. *cis natural antisense transcript*) oraz *trans*-NAT (ang. *trans natural antisense transcript*). Ze względu na wysoką komplementarność sekwencji retrogenów do genów rodzicielskich, retrogeny mogą tworzyć dupleksy RNA:RNA (Kubiak i in., 2020), które odgrywają ważną rolę w modulacji splicingu pre-mRNA, edycji RNA, kontroli stabilności mRNA, a także znoszeniu represji indukowanej miRNA (Rosikiewicz & Makałowska, 2016). W badaniu Kubiak i in.,

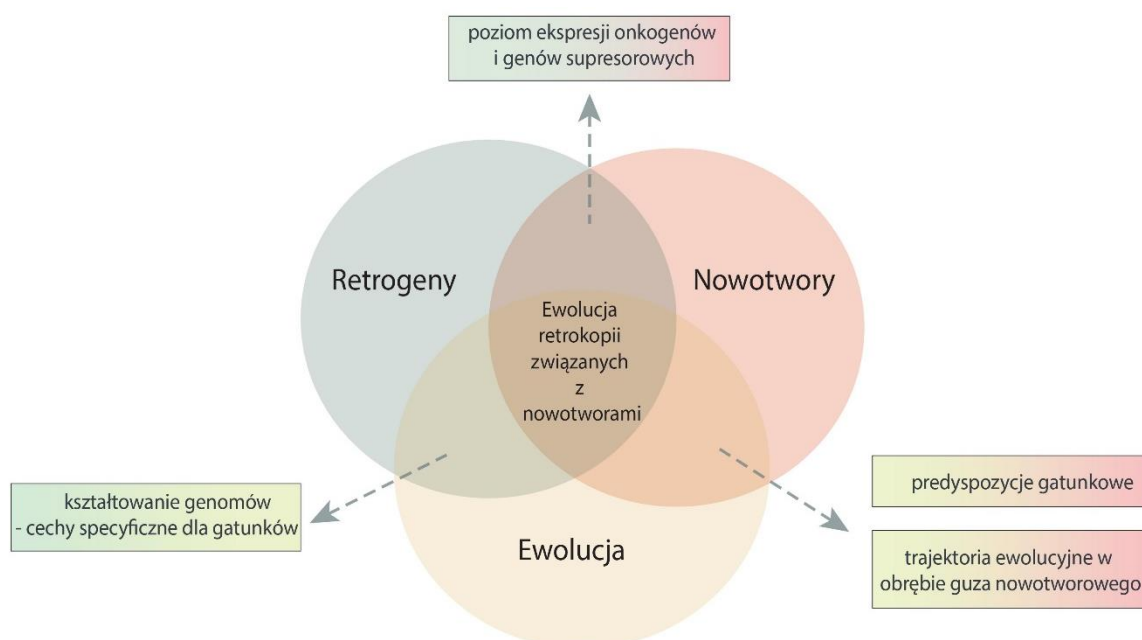
wskazano 78 retrokopii potencjalnie regulujących ekspresję swoich genów rodzicielskich działając jako *trans*-NATs (Kubiak i in., 2020).

Ponad 40% retrokopii jest zagnieżdżonych w intronach innych genów (2019 spośród 4611 retrogenów, na podstawie bazy danych RetrogeneDB2). Kiedy retrokopie występują na przeciwnej nici DNA mogą regulować ekspresję gospodarza działając jako *cis*-NAT (Groen i in., 2014; Kubiak i in., 2020). Wyjaśnienie działania tego mechanizmu nie jest do końca znane. W wielkoskalowej analizie zidentyfikowano 656 retrokopii, wykazujących istotną statystycznie korelację ekspresji ze swoim genem gospodarzem – co może sugerować ich funkcję jako *cis*-NAT (Kubiak i in., 2020). Przykładem retrogeny potencjalnie funkcjonującego jako *cis*-NAT może być *PDIA3P1* (ang. *protein disulfide isomerase family A member 3 pseudogene 1, retro_hsap_217*) zagnieżdżony w onkogenie *FMO5* (ang. *flavin containing dimethylaniline monooxygenase 5*). *FMO5* wykazuje zwiększoną ekspresję w raku jelita grubego (T. Zhang i in., 2018). Jego ekspresja jest dodatnio skorelowana z ekspresją retrogeny co może sugerować pewną regulacyjną rolę zagnieżdżonej retrokopii (Kubiak i in., 2020).

Retrokopie aktywne transkrypcyjnie mogą także pełnić funkcje gąbek miRNA. Takim przykładem są gen *HMGAI* (ang. *high mobility group AT-hook 1*) oraz jego dwie retrokopie. *HMGAI*P6 i *HMGAI*P7, które działają jak „wabiki” dla cząsteczek miRNA i w ten sposób regulują ekspresję rodzicielskiego genu *HMGAI* (Esposito i in., 2014). Innym przykładem jest długi niekodujący RNA *INTS6P1* (ang. *integrator complex subunit 6 pseudogene 1*) wraz ze swoim genem rodzicielskim *INTS6*, który jest genem supresorowym w raku wątrobowokomórkowym. Mechanizm regulacji ekspresji także powiązano z konkurencją o onkogenny *miR-17-5p* (Peng i in., 2015).

1.3. Retrogeny, nowotwory i ewolucja

Retrokopie jako ważne elementy biorące udział w ewolucji gatunków, powiązano także z występowaniem nowotworów. W literaturze opisano retrokopie, które wykazują właściwości zarówno onkogenne jak i supresorowe (Esposito i in., 2014; Poliseno i in., 2015; Rohozinski i in., 2012; Staszak & Makałowska, 2021; Yu i in., 2014). Natomiast szybkie tempo ewolucji, które charakteryzuje geny powstałe w wyniku retropozycji, należy także do cech typowych dla guzów nowotworowych. W rozdziale tym opisane są zatem powiązania i podobieństwa pomiędzy retrogenami i nowotworami z uwzględnieniem aspektów ewolucyjnych (Rycina 2).



Rycina 2. Schematyczne ujęcie związków pomiędzy retrogenami, procesem nowotworzenia a ewolucją (na podstawie Staszak & Makałowska, 2021).

1.3.1. Mechanizmy ewolucyjne i nowotwory

Kancerogeneza jest zjawiskiem złożonym, wieloczynnikowym i wciąż owianym wieloma niewiadomymi. Wysokie ryzyko transformacji nowotworowej u człowieka może być spowodowane na przykład stylem życia czy czynnikami środowiskowymi. Kluczowym aspektem są jednak czynniki genetyczne, predysponujące do rozwoju nowotworu. Spoglądając na drzewo ewolucji zwierząt, zjawisko to wydaje się dość powszechne (Kitsoulis i in., 2020; Vincze i in., 2022).

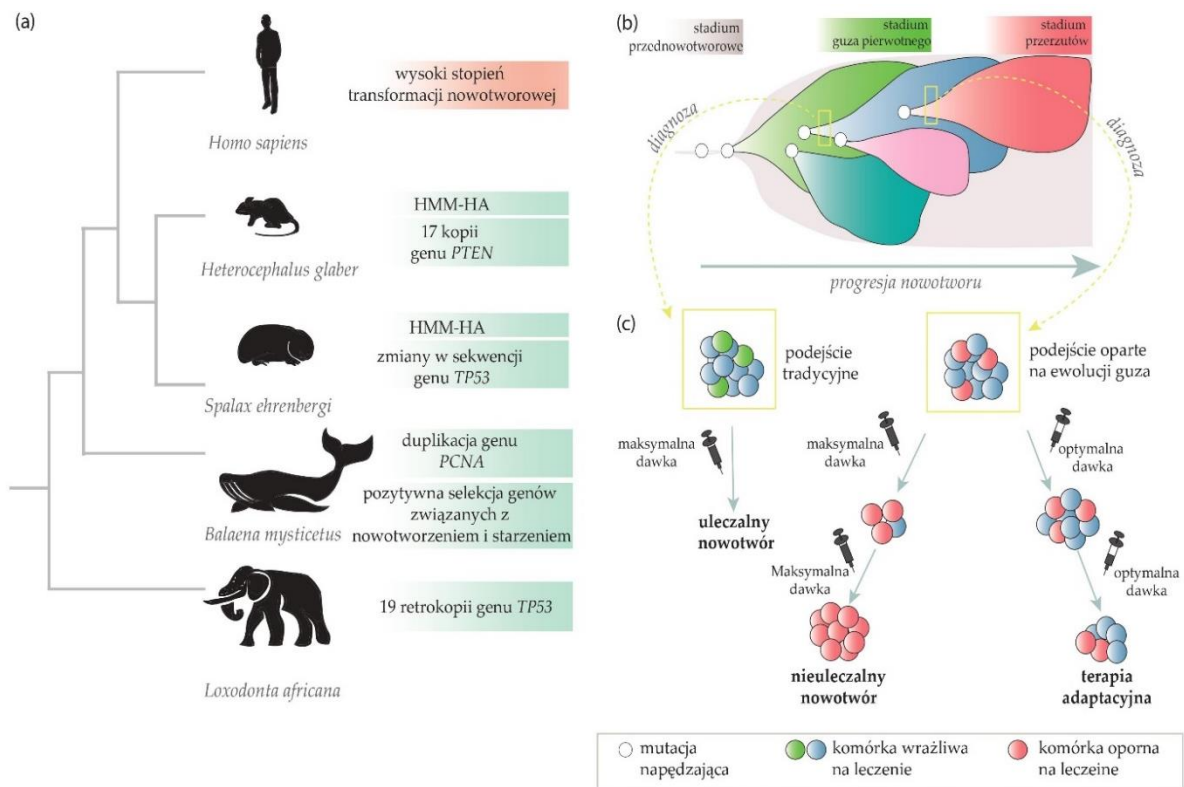
U kilku gatunków zwierząt zidentyfikowano geny, które powiązano ze zmniejszoną podatnością na transformację nowotworową (Rycina 3a). Mechanizm „odporności” na procesy nowotworzenia zaobserwowano w populacji długowiecznych wielorybów grenlandzkich (żyją ok 200 lat). Zjawisko to powiązano m.in. z wysokim tempem ewolucji genów zaangażowanych zarówno w powstawanie nowotworów, jak i procesy starzenia tj. *SOC2* (ang. *suppressor of cytokine signaling 2*), *APTX* (ang. *apoptosis*), *LEP* (ang. *leptin*) oraz specyficznymi mutacjami w obrębie genów związanych z termoregulacją (gen *UCP1*, ang. *uncoupling protein 1*). Dodatkowo, zaobserwowano duplikację genu *PCNA* (ang. *proliferating cell nuclear antigen*), który jest kluczowy dla procesów naprawy uszkodzeń DNA. Podejrzewa się, że geny te są związane ze zmniejszoną częstotliwością mutacji, co z kolei przyczynia się do zapobiegania powstawaniu nowotworów (Keane i in., 2015; Seim i in., 2014).

U słonia mechanizm odporności na nowotwory powiązано ze skuteczniejszą reakcją na uszkodzenia DNA dzięki dodatkowym kopiom genu *TP53* (ang. *tumor protein p53*). W przeciwieństwie do ludzkich komórek, gdzie zazwyczaj występuje jedna kopia genu *TP53*, słonie afrykańskie mają aż 20 kopii tego genu, z czego 19 powstało w wyniku retropozycji (Abegglen i in., 2015). Gen *TP53* koduje białko p53, które jest nazywane „strażnikiem genomu” (Efeyan & Serrano, 2007). Białko to jest kluczowym supresorem nowotworów, a mutacje w genie *TP53* występują często w różnych typach nowotworów u ludzi (Giacomelli i in., 2018; Olivier i in., 2010). Zaburzenie funkcji białka p53 prowadzi do wystąpienia cech charakterystycznych dla komórek nowotworowych. Dodatkowe kopie genu *TP53* u słoni afrykańskich sprzyjają efektywnej reakcji na uszkodzenia DNA poprzez nadaktywny szlak z udziałem *TP53* (Abegglen i in., 2015).

Specyficzny mechanizm "odporności na nowotwory" możemy znaleźć także wśród przedstawicieli gryzoni. Ślepy kretoszczur (łac. *Spalax ehrenbergi*) potrafi dożyć 21 lat, wyjątkowo długo w porównaniu z innymi przedstawicielami tego rzędu ssaków (mysz żyje ok. 4 lata). Ponadto, należy do gatunków o wyjątkowo niskim wskaźniku transformacji nowotworowej – nigdy nie zaobserwowano u jego przedstawicieli występowania nowotworów. Żyjąc pod ziemią, gatunek ten charakteryzuje się silną tolerancją na ekstremalne warunki niedotlenienia. Białko p53 kontroluje odpowiedź komórkową na różne warunki stresowe, w tym uszkodzenie DNA i niedotlenienie, prowadząc do zatrzymania podziałów komórkowych i apoptozy. Inaktywacja p53, występująca w 40–50% ludzkich nowotworów, zapewnia selektywną przewagę w niedotlenionym mikrośrodku podczas progresji nowotworu. Poddano zatem badaniom białko p53 u ślepego kretoszczura i zidentyfikowano substytucję argininy na lizynę w domenie wiążącej DNA (miejsce 174). Białko to jest zdolne do indukowania zatrzymania cyklu komórkowego, ale nie inicjuje apoptozy. Późniejsze badania eksperymentalne wykazały, że dominującym mechanizmem przeciwnowotworowym w komórkach tego gatunku nie jest apoptoza, lecz nekroza. Jedną z korzyści martwicy może być usunięcie wszystkich komórek wokół zmiany przednowotworowej, całego potencjalnie reaktywnego mikrośrodku (Gorbunova i in., 2012).

Innym przykładem gryzonia, u którego wyewoluowała strategia przeciwdziałania procesowi nowotworzenia, jest nagi kretoszczur (łac. *Heterocephalus glaber*) – został opisany jako gatunek niemal „wolny od raka” (Tollis i in., 2020). Mechanizm tej niezwyklej odporności polega na ograniczeniu proliferacji komórek poprzez ekspresję hialuronianu o dużej masie cząsteczkowej (HMM-HA, ang. *high-molecular-mass hyaluronan*). Dłuższa forma HMM-HA

pełni rolę cytostatyczną, hamując podziały komórkowe (Tian i in., 2013). Ponadto, w genomie tego gatunku odnotowano 17 dodatkowych kopii genu *PTEN* (ang. *phosphatase and tensin homolog*), który jest ważnym supresorem nowotworowym - co może dodatkowo przyczyniać się do silnej odporności na nowotwory u tego gatunku (J. Tang i in., 2016).



Rycina 3. Mechanizmy ewolucyjne a nowotwory.

Geny związane z „odpornością” na występowanie nowotworów u niektórych gatunków ssaków **(a)**, kolejne subpopulacje komórek podczas progresji guza nowotworowego **(b)**, ewolucyjne podejście do leczenia onkologicznego **(c)** (na podstawie Staszak & Makiłowska, 2021).

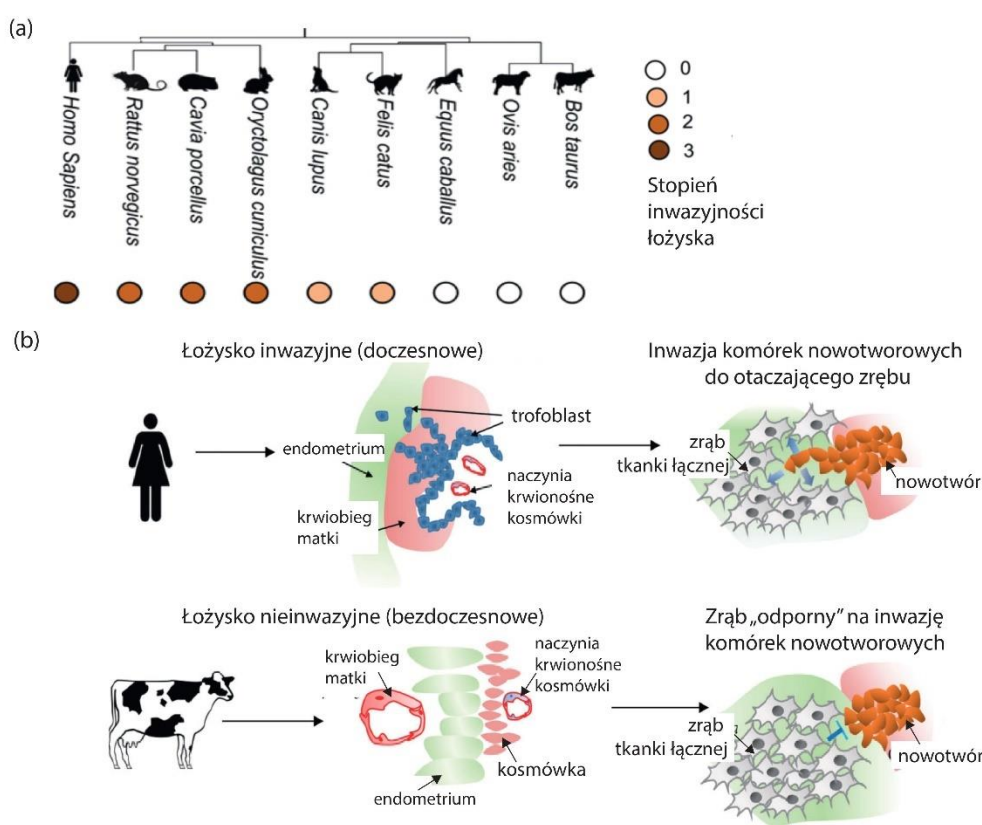
Ewolucyjny charakter kancerogenezy znajduje również odzwierciedlenie w mikroewolucji komórek w obrębie samego nowotworu (Venkatesan i in., 2017; Zhu i in., 2021). Guz nowotworowy składa się z wielu kopii pojedynczej komórki. Podczas rozwoju nowotworu mutacje prowadzą do powstania heterogenicznych grup komórek, które stanowią punkt wyjścia dla działania presji ewolucyjnej. W trakcie kolejnych niekontrolowanych podziałów komórkowych dochodzi do powstania różnych subklonów genetycznych (McPherson i in., 2018) (Rycina 3b). W tym momencie selekcja ewolucyjna zaczyna działać i niektóre grupy komórek nowotworowych zaczynają szybciej ewoluować. Takie sytuacje mogą mieć miejsce, gdy pewne komórki nabiorą cech, które dają im przewagę w określonym środowisku nowotworu, co skutkuje większą liczbą kolejnych komórek potomnych. Ponadto wysoki

stopień różnorodności nowotworów i niestabilność genomowa zwiększają ryzyko wystąpienia mutacji korzystnych dla komórek nowotworowych, co z kolei przyspiesza postęp choroby (Fortunato i in., 2017).

Kolejnym argumentem przemawiającym za istotnym działaniem mechanizmów ewolucyjnych w procesie nowotworzenia jest to, że zastosowanie terapii przeciwnowotworowej może wywoływać presję selekcyjną – działającą analogicznie do presji selekcyjnej podczas zmieniających się warunków środowiska naturalnego (Rycina 3c). Ta sama terapia może wywołać odmienny skutek u różnych pacjentów, u których zdiagnozowano ten sam typ nowotworu. Zakładając, że u pierwszego pacjenta zdiagnozowany nowotwór składa się wyłącznie z komórek wrażliwych na terapię, wówczas rokowanie pacjenta jest dobre i najprawdopodobniej pacjent zostanie wyleczony. Jednakże u niektórych pacjentów po zastosowaniu określonych terapii obserwuje się początkową odpowiedź, po czym progresja nowotworu niespodziewanie przyspiesza. To może sugerować obecność subpopulacji komórek opornych na terapię we wstępnie leczonym guzie. Presja selekcyjna wynikająca z terapii może prowadzić do wzrostu populacji opornych na leczenie, co z kolei może prowadzić do rozwoju "oporności nabytej". Z tego powodu zaproponowano terapię adaptacyjną, która polega na utrzymaniu populacji komórek wrażliwych na lek, co ogranicza wzrost populacji opornych na leczenie (Enriquez-Navas i in., 2015; McPherson i in., 2018; Russo & Bardelli, 2017; Venkatesan i in., 2017).

Rola mechanizmów ewolucyjnych w powstawaniu podatności na nowotwory może mieć odzwierciedlenie w pojawieniu się cech adaptacyjnych u poszczególnych gromad kręgowców. Jedną z nich jest występowanie łożyska u ssaków łożyskowych. Pomiedzy procesem inwazji łożyska a inwazją komórek nowotworowych występują duże podobieństwa (Kshitiz i in., 2019). Inwazję komórkową do sąsiednich tkanek obserwuje się w czasie ciąży, gdy łożysko płodu nacieka endometrium matki oraz w przypadku raka, gdy rozsiew komórek nowotworowych do otaczającego zrębu komórek jest etapem poprzedzającym wystąpienie przerzutów. Stopień inwazji łożyska do endometrium różni się znacznie u różnych ssaków łożyskowych (Rycina 4a) (Suhail i in., 2022). Sugeruje się, że gatunki charakteryzujące się bardziej inwazyjnym typem łożyska (np. człowiek) są bardziej podatne na nowotwory złośliwe. Przeprowadzono badania, gdzie porównano proces inwazji trofoblastu do fibroblastów z użyciem komórek ludzkich (gatunek o inwazyjnym łożysku) oraz bydłowych (gatunek o nieinwazyjnym łożysku) (Rycina 4b). Zidentyfikowano geny o wysokiej ekspresji w ludzkich fibroblastach, które nie ulegały ekspresji w fibroblastach bydłowych. Następnie wyciszono ich

ekspresję w ludzkich fibroblastach i zauważono, że były one bardziej odporne na inwazję komórek nowotworowych niż fibroblasty stanowiące kontrolę. Wywnioskowano, że geny związane ze zdolnością wnikania łożyska do endometrium macicy, mogą być reaktywowane w komórkach nowotworowych, co może prowadzić do przerzutów (Kshitiz i in., 2019). Ponadto, obecność hipermetylacji genów supresorowych zarówno w łożysku człowieka, jak i pawiana, ale nie w łożysku myszy, sugeruje, że geny te przyczyniają się do specyficznej dla naczelných inwazji łożyska w sposób porównywalny z rolą tych genów w inwazyjnych nowotworach (Macaulay i in., 2017; Suhail i in., 2022).



Rycina 4. Inwazja łożyska a proces inwazji komórek nowotworowych. Stopień inwazji łożyska u wybranych gatunków ssaków łożyskowych (na podstawie Suhail i in., 2022) (a). Schemat dwóch rodzajów łożyska oraz ich związek z inwazją komórek nowotworowych (na podstawie Kshitiz i in., 2019) (b).

1.3.2. Retrogeny i nowotwory

Ewolucję specyficznych dla gatunku mechanizmów oporności na raka oraz występowanie określonych typów nowotworów i odpowiedzi pacjentów na zastosowane leczenie powiązano w niektórych przypadkach z ekspresją retrogenów (R. Wang i in., 2018; S.

Wang & Yu, 2019). Jednym z najlepiej opisanych retrogenów związanych z procesem nowotworzenia jest *RHOB* (ang. *Ras homolog family member B*), znany głównie jako supresor nowotworu, który powstał w wyniku retropozycji na wczesnych etapach ewolucji kręgowców (J. A. Ju & Gilkes, 2018; Mazieres i in., 2004). Kolejnym przykładem retrogenu o funkcji supresorowej jest *NKAPL* (ang. *NFκB activating protein like*). Zmniejszenie poziomu ekspresji tej retrokopii powiązано z krótszym okresem przeżycia pacjentów z rakiem brodawkowatokomórkowym nerki, gruczolakorakiem trzustki czy rakiem migdałkowo-torbielowatym (M. Li i in., 2017). Zmniejszony poziom ekspresji w nowotworach wykazały także inne retrogeny jak np. *PTENP1* (Yu i in., 2014), *CTNNA1P1* (X. Chen i in., 2016), *TUSC2P1* (Rutnam i in., 2014) i *INTS6P1* (Peng i in., 2015).

Dużą grupę stanowią retrogeny o podwyższonej ekspresji w nowotworach. W wielu przypadkach zwiększona ekspresja tych retrokopii sprzyja rozwojowi nowotworu. Jednym z przykładów takiego retrogenu jest *UTP14C* (ang. *UTP14C small subunit processome component*). Predyspozycja do raka jajnika jest powiązana z podniesioną ekspresją *UTP14C*, retrokopii genu *UTP14A*. Mechanizm regulatorowy najprawdopodobniej polega na tym, że *UTP14C* kieruje białko p53 na drogę degradacji. W komórkach o wysokiej ekspresji *UTP14C* białko p53 nie może spełniać swojej funkcji, czyli blokować cyklu komórkowego i kierować komórek na drogę apoptozy (Rohozinski i in., 2012). Z kolei, zwiększona ekspresja *MSL3P1* (ang. *MSL complex subunit 3B*) została skorelowana z występowaniem raka nerkowokomórkowego (B. Chen i in., 2018). Do grupy retrogenów o zwiększonym poziomie ekspresji w nowotworach należą również m.in. *KRASP1* (Poliseno i in., 2010), *UBE2CP3* (Lin i in., 2018), *ANXA2P2* (Q.-S. Wang i in., 2019), *RACGAP1P* (M.-Y. Wang i in., 2019) oraz *LGMNP1* (H. Xu i in., 2019).

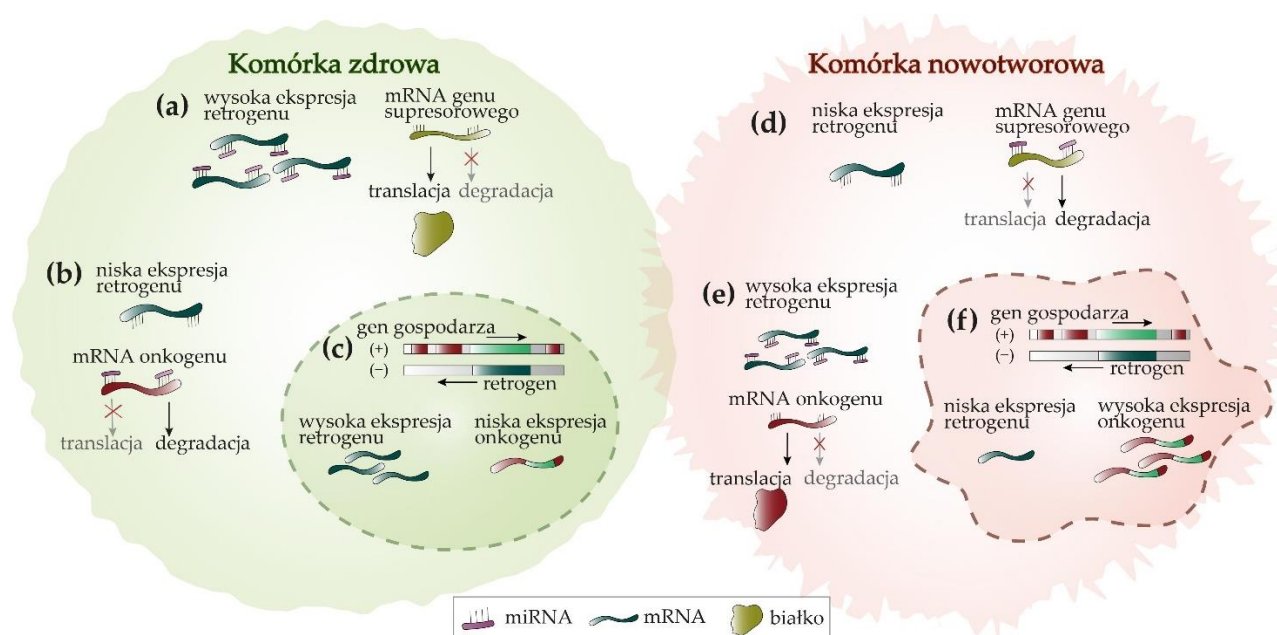
Jednym z retrogenów, które powiązано z występowaniem przerzutów jest wspomniany wcześniej *CTNNA1P1* (ang. *catenin alpha 1 pseudogene 1*). U pacjentów z rakiem jelita grubego wykazano istotną korelację ekspresji tego retrogenu z inwazją do węzłów chłonnych (X. Chen i in., 2016). Zwiększoną ekspresję retrogenów *NANOGP8* (ang. *Nanog homeobox retrogene P8*) i *POU5F1P4/P5* (ang. *POU class 5 homeobox 1 pseudogene 4/pseudogene 5*) powiązано natomiast z występowaniem fenotypu macierzystych komórek nowotworowych (CSCs, ang. *cancer stem cells*). Występowanie tej subpopulacji komórek o wysokiej zdolności do inwazji, zwiastuje intensywną ekspansję nowotworu (Fairbanks i in., 2012).

W literaturze znajdujemy także przykłady retrogenów istotnych w odpowiedzi na zastosowane leczenie. Retrokopię *PPIAP43* (ang. *peptidylprolyl isomerase A pseudogene 43*)

powiązano z wrażliwością na radioterapię wśród pacjentów z niedrobnokomórkowym rakiem płuc (S. Wang & Yu, 2019). Interesujący przykład stanowi również retrogen *FTHIP3* (ang. *ferritin heavy chain 1 pseudogene 3*), którego poziom ekspresji powiązano z opornością na leczenie paclitaxelem (Wang i wsp., 2019).

Dokładne mechanizmy działania retrokopii w komórkach nowotworowych w większości przypadków nie są dobrze poznane. Retrokopie mogą pełnić różnorodne funkcje, ale wydaje się, że w większości przypadków działają jako niekodujące gąbki miRNA (Ciomborowska-Basheer i in., 2021). Ze względu na mechanizm powstawania, retrokopie charakteryzują się wysokim stopniem podobieństwa sekwencji do swoich przodków (genów rodzicielskich). Mimo, że potencjał kodowania białek nie zawsze jest zachowany, miejsca docelowe miRNA często pozostają w sekwencji retrokopii niezmienione. Przykłady funkcji retrokopii w komórkach zdrowej i nowotworowej ilustruje Rycina 5. Sytuację, w której retrogen reguluje ekspresję genu rodzicielskiego będącego genem supresorowym przedstawia Rycina 5a. W zdrowej komórce poziom ekspresji tego retrogenu jest zrównoważony. Wystarczająca ekspresja retrokopii reguluje ekspresję genów supresorowych poprzez konkutowanie o wspólne miRNA. Zapobiega to degradacji transkryptu genu supresorowego i umożliwia proces translacji. Analogicznie, obniżony poziom ekspresji tego typu retrogenów przyczynia się do zwiększonego wiązania miRNA z transkryptem genu supresorowego, co kieruje go na ścieżkę degradacji i sprzyja transformacji nowotworowej (Rycina 5d). Dobrym przykładem jest niski poziom ekspresji retrokopii *PTENP1* (ang. *phosphatase and tensin homolog pseudogene 1*), który powiązano z występowaniem nowotworów (C.-L. Chen i in., 2015; Yu i in., 2014). W normalnych warunkach sensowny transkrypt *PTENP1* chroni gen rodzicielski przed wiązaniem miRNA, a tym samym przed degradacją i obniżonym poziomem translacji. W komórkach nowotworowych obniżenie ekspresji *PTENP1* prowadzi do degradacji *PTEN* pod wpływem wiązania miRNA (Poliseno i in., 2015). Innym przykładem jest długi niekodujący RNA *INTS6P1* wraz z jego genem rodzicielskim *INTS6* (ang. *integrator complex subunit 6*), supresorem raka wątrobowokomórkowego. Mechanizm ich współdziałania powiązano z konkurencją o onkogenny *miR-17-5p* (Peng i in., 2015). Odwrotnie jest w przypadku retrokopii, które pochodzą od genów rodzicielskich należących do onkogenów. W zdrowej komórce niska ekspresja retrogenu, który ma wspólne miejsca wiązania miRNA z danym onkogenem, powoduje, że miRNA wiąże się głównie z mRNA onkogenu i kieruje go na ścieżkę degradacji (Rycina 5b). W komórce nowotworowej podwyższona ekspresja retrokopii powoduje wiązanie większej ilości miRNA, co z kolei zapobiega degradacji transkryptu onkogenu (Rycina 5e),

prowadząc do progresji nowotworu. Dobrym przykładem jest gen *HMGA1* i jego retrokopie *HMGA1P6* i *HMGA1P7* działające jako gąbki dla wspólnych miRNA, regulujące w ten sposób ekspresję onkogenu *HMGA1* (Esposito i in., 2014).



Rycina 5. Retrogeny związane z procesem nowotworzenia jako gąbki miRNA i *cis*-NATs w komórce zdrowej i nowotworowej.

Wiązanie miRNA przez retrogen o wysokiej ekspresji i translacja genu supresorowego (a). Degradacja mRNA onkogenu poprzez wiązanie miRNA z powodu niskiej ekspresji retrogeny (b). Wysoka ekspresja retrogeny i spadek ekspresji onkogennego genu gospodarza (c). Degradacja mRNA genu supresorowego poprzez wiązanie miRNA z powodu niskiego poziomu retrogeny (d). Wiązanie miRNA z retrogenem o wysokiej ekspresji i translacja onkogenu (e). Niska ekspresja retrogeny i podwyższony poziom ekspresji gospodarza działającego jako onkogen (f) (na podstawie Staszak & Makałowska, 2021).

Istnieje kilka możliwych scenariuszy, w których retrogeny mogą regulować funkcje swoich gospodarzy zarówno w komórkach zdrowych, jak i nowotworowych. Retrogen działający jako *cis*-NAT może, w zależności od mechanizmu działania, zarówno promować, jak i hamować transkrypcję gospodarza (Rosikiewicz & Makałowska, 2016). Rycina 5c przedstawia przypadek, w którym retrogen jest zagnieżdżony w onkogenie a ich ekspresja jest ze sobą negatywnie skorelowana. W związku z czym w zdrowej komórce wysoka ekspresja retrogeny może utrzymać ekspresję onkogenu na niskim poziomie poprzez np. zjawisko interferencji transkrypcji. Z kolei w komórce nowotworowej niska ekspresja retrogeny nie zatrzymuje transkrypcji onkogennego genu gospodarza, co skutkuje jego wysoką ekspresją (Rycina 5f).

Coraz więcej badań podkreśla funkcjonalność retrogenów oraz ich związek z występowaniem chorób, w tym nowotworów. Mimo to, retrokopie są nadal często pomijane jako geny kandydujące w badaniach nad nowotworami. Może to wynikać z tego, że analiza tych duplikatów stanowi duże wyzwanie, głównie ze względu na ich niską ekspresję w porównaniu z genami kodującymi białka oraz wysoki poziom podobieństwa sekwencji z ich genami rodzicielskimi (Staszak & Makałowska, 2021). Ustalenie profilu ekspresji i funkcji retrogenów może stanowić istotny kierunek badań nad korelacją wydarzeń ewolucyjnych z podatnością na proces nowotworzenia wśród różnych gatunków, w tym człowieka. Wyselekcjonowane retrogeny-biomarkery danych typów nowotworów, mogłyby znaleźć zastosowanie w standardowej diagnostyce onkologicznej, wpływając na zwiększenie skuteczności identyfikowania poszczególnych rodzajów nowotworów.

2. CEL PRACY

Retropozycja dostarcza nowego materiału genetycznego, który przynajmniej początkowo nie podlega presji ewolucyjnej i dość szybko ewoluuje. Proces retropozycji genów kodujących białka był szczególnie intensywny w czasie ewolucji naczelnych. Wiele z tych nowo powstałych duplikatów genów ulega transkrypcji. Część koduje białka, inne funkcjonują jako regulatorowe RNA. Związek pomiędzy tymi nowymi funkcjonalnymi cząsteczkami a nowotworami nie był dotąd badany. A zatem, **głównym celem pracy było sprawdzenie na ile intensywny proces retropozycji u naczelnych może być powiązany ze wzmożonymi, w porównaniu z innymi gatunkami, procesami nowotworzenia u człowieka.** Postawiono trzy hipotezy badawcze:

1. Specyficzne dla naczelnych lub człowieka retrogeny częściej wykazują związane z nowotworami zmiany ekspresji w porównaniu z retrogenami ewolucyjnie starszymi.
2. Wśród retrogenów wykazujących zmienioną ekspresję w nowotworach, przeważają retrogeny powstałe z genów kodujących istotne dla procesów nowotworowych białka lub ulokowane w intronach takich genów.
3. Retrogeny, które powstały ze znanych onkogenów i genów supresorowych lub są zagnieżdżone w takich genach, powiązane są z procesami nowotworowymi poprzez regulację ekspresji tychże genów.

Bardzo duże podobieństwa pomiędzy procesem inwazji łożyska a inwazją komórek nowotworowych sugerują, że gatunki charakteryzujące się bardziej inwazyjnym typem łożyska (np. człowiek) są bardziej podatne na nowotwory złośliwe. Dlatego postawiono kolejną, związaną z tym zagadnieniem hipotezę:

4. Retrogeny, które wykazują zmienioną poziom ekspresji w nowotworach, w podobny sposób zmieniają ekspresję w trakcie rozwoju łożyska.

3. MATERIAŁY I METODY

3.1. Analizy bioinformatyczne

3.1.1. Dane transkryptomiczne (RNA-Seq)

Materiał do badań stanowiły dane transkryptomiczne (RNA-Seq, ang. *RNA Sequencing*) człowieka (łac. *Homo sapiens*), psa (łac. *Canis lupus familiaris*) oraz kury (łac. *Gallus gallus*) zdeponowane w bazie danych NCBI SRA (*The National Center for Biotechnology Information Sequence Read Archive*) (Leinonen i in., 2011) (Tabela 1, Tabela 2). Selekcja zestawów danych opierała się głównie na liczbie prób oraz w miarę możliwości wspólnym pochodzeniu (próby z tkanek nowotworowych i zdrowych od tych samych osobników). Wymienione gatunki zwierząt zostały wybrane jako grupa porównawcza ze względu na dystans ewolucyjny i ilość dostępnych danych nowotworowych. Dodatkowo pies posiada stosunkowo dobrze poznany genom i najbliższy człowiekowi poziom opieki medycznej, co pozwalało na zebranie materiału badawczego.

Tabela 1. Zestawy danych nowotworowych z repozytorium NCBI SRA użyte do analiz bioinformatycznych. ER+- rak piersi z pozytywnym statusem receptorów estrogenowych, TNBC- potrójnie ujemny rak piersi, C2, C1- podtypy molekularne wątrobiaka zarodkowego.

Zestaw danych (numer ID)	Próby - tkanka nowotworowa	Próby - tkanka okalająca (zdrowa)
Człowiek (łac. <i>Homo sapiens</i>)		
Rak piersi (PRJNA251383) (Varley i in., 2014)	42 ER+ 42 TNBC	30 okalające guz ER+, 21 okalające guz TNBC, 5 z mammoplastyki
Gruzołakorak jelita grubego (PRJNA218851) (Kim i in., 2014)	18 pierwotny rak jelita grubego, 18 przerzuty do wątroby	18
Gruzołakorak płuc (PRJEB2784) (Ju i in., 2012)	85	77
Rak wątrobowokomórkowy (PRJNA310012) (Liu i in., 2016)	50	50
Wątrobiak zarodkowy (PRJNA413799) (Hooks i in., 2018)	18 subtyp C1 7 subtyp C2	25
Pies (łac. <i>Canis lupus familiaris</i>)		
Rak gruczołu mlekowego psa (PRJNA489087) (Kim i in., 2020)	138 (9 ras)*	61
Rak gruczołu mlekowego psa (PRJNA561580) (Graim i in., 2020)	22 (<i>carcinoma</i>) 41 (<i>adenoma</i>)	26 próby zdrowe z widocznymi atypiami
Kura (łac. <i>Gallus gallus</i>)		
LMH Linia komórkowa raka wątrobowokomórkowego kury (PRJNA239321) (Sun i in., 2015)	6	-
Próbki zdrowej wątroby kury (PRJNA362606) (Tang i in., 2017)	-	6

*przeanalizowano tylko te rasy, dla których dostępne były zarówno próby nowotworowe, jak i zdrowe, czyli wskazane 138 prób nowotworowych i 61 prób okalających

Tabela 2. Zestaw danych z eksperymentu inwazji łożyska z repozytorium NCBI SRA użyty do analiz bioinformatycznych.

Zestaw danych (numer ID)	Fibroblasty zrębu endometrium po inwazji trofoblastem	Fibroblasty zrębu endometrium
Człowiek (łac. <i>Homo sapiens</i>)		
Inwazja łożyska (PRJNA562073) (Kshitiz i in., 2019)	3	3

W celu zwiększenia skali analiz ekspresji retrokopii w nowotworach człowieka, przeanalizowano dane dotyczące ekspresji 10 rodzajów nowotworów znajdujące się w repozytorium TCGA (*The Cancer Genome Atlas*) (*The Cancer Genome Atlas Program (TCGA)* - NCI, 2022; Weinstein i in., 2013) (Tabela 3). Dane z glejaka wielopostaciowego (TCGA-GBM) oraz raka jajnika (TCGA-OV) z uwagi na małą ilość lub brak prób zdrowych zostały użyte tylko do analizy ogólnej ekspresji retrogenów, nie uwzględniono ich w analizach ekspresji różnicowej, gdzie porównywano próby nowotworowe i próby zdrowe.

Tabela 3. Zestawy danych z repozytorium TCGA użyte do analiz bioinformatycznych.
*oznaczono pakiety danych, których nie uwzględniono w analizach ekspresji różnicowej, gdzie porównywano próby nowotworowe i próby zdrowe z uwagi na zbyt małą liczbę prób zdrowych

Typ nowotworu	Próby nowotworowe	Próby zdrowe	Próby nowotworowe			
			Stadium IV	Stadium I	Efekt terapii – postępujący nowotwór	Efekt terapii – całkowita remisja
Rak piersi (TCGA-BRCA, ang. <i>breast carcinoma</i>)	1113	113	-	-	-	-
Gruźlakorak jelita grubego (TCGA-COAD, <i>colon adenocarcinoma</i>)	473	41	66	78	24	134
Glejak wielopostaciowy (TCGA-GBM, ang. <i>glioblastoma multiforme</i>)*	169	5	-	-	-	-
Rak jasnokomórkowy nerki (TCGA-KIRC, ang. <i>kidney renal clear cell carcinoma</i>)	538	72	81	272	4	48
Rak wątrobowokomórkowy (TCGA-LIHC, ang. <i>liver hepatocellular carcinoma</i>)	374	50	5	177	-	-
Rak płaskonabłonkowy płuc (TCGA-LUSC, ang. <i>lung squamous cell carcinoma</i>)	501	51	-	-	13	274
Rak jajnika (TCGA-OV, ang. <i>ovary cancer</i>)*	429	0	-	-	32	243
Gruźlakorak prostaty (TCGA-PRAD, ang. <i>prostate adenocarcinoma</i>)	502	52	-	-	16	313

Gruczolakorak odbytnicy (TCGA-READ, ang. <i>rectum adenocarcinoma</i>)	167	10	25	33	6	42
Rak endometrium trzonu macicy (TCGA-UCEC, ang. <i>uterine corpus endometrial carcinoma</i>)	550	35	-	-	17	388

3.1.2. Adnotacje retrokopii

Jako źródło adnotacji retrokopii wykorzystano bazę danych RetrogeneDB2 (Rosikiewicz i in. 2017). W związku z tym, że RetrogeneDB2 opiera się na poprzedniej wersji ludzkiego genomu (GRCh37), Ensembl ID zostały przekonwertowane do nowszej wersji ludzkiego genomu, w oparciu o koordynaty genomowe przy użyciu Assembly Converter (Yates i in., 2020). Ostateczna lista obejmowała zaktualizowane do wersji GRCh38 retrokopie znajdujące się zarówno w Ensembl jak i RetrogeneDB2, retrokopie adnotowane jedynie w RetrogeneDB2 (tzw. *novel*) oraz dodatkowo retrokopie związane z nowotworami pochodzące z danych literaturowych (niezaadnotowane w RetrogeneDB2). W przypadku pozostałych gatunków nie była dostępna nowsza wersja genomu i w związku z tym nie było potrzeby aktualizacji identyfikatorów i nazw retrokopii. W sytuacji braku nazwy genu w Ensembl pozostawiono ID z RetrogeneDB2.

3.1.3. Przetwarzanie danych RNA-Seq

Bezpośrednio pobrane dane RNA-Seq z NCBI SRA zostały poddane kontroli jakości z użyciem programu FastQC (Andrews, 2010) (Rycina 6a). Wykonane następnie wstępne przetwarzanie danych z użyciem narzędzia BBduk2 (qtrim=w trimq=20 maq=10 k=23 mink=11 hdist=1 tbo tpe minlength=2/3length; <https://jgi.doe.gov>) obejmowało przycinanie i filtrowanie odczytów oraz usuwanie sekwencji adapterów, stanowiących potencjalne zanieczyszczenia prób (Tabela 4). Zgodnie z informacją zawartą w repozytoriach, z których pobrano dane, podczas przygotowywania bibliotek RNA-Seq usunięte zostały odczyty pochodzące z rybosomalnego RNA (rRNA). W celu usunięcia ewentualnych pozostałości rRNA odczyty zostały zmapowane do referencyjnego rRNA (Ensembl, NCBI) z użyciem Bowtie2 (Langmead & Salzberg, 2012). Przetworzone dane zostały poddane ponownej kontroli jakości (Rycina 6b). Przed podjęciem kolejnych kroków określony został rodzaj każdej z bibliotek, gdyż jest to istotnym parametrem na etapie mapowania. W tym celu posłużono się narzędziem infer_experiment.py (<http://rseqc.sourceforge.net/>). Poziom ekspresji został oszacowany z użyciem programu Salmon przy użyciu parametrów domyślnych oraz flag —seqBias i —gcBias, w celu skorygowania potencjalnych odchyleń w danych wejściowych

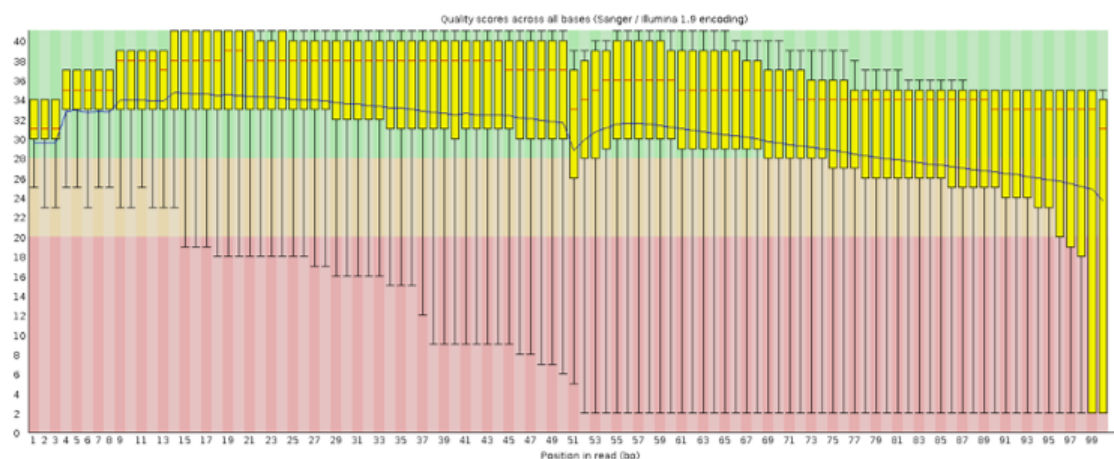
(Patro i in., 2017). Do mapowania użyto następujących wersji transkryptomów: człowiek i pies (Ensembl 101) oraz kury (Ensembl 103), które wzbogacono o sekwencje retrokopii zaadnotowanych jedynie w bazie RetrogeneDB2 (*novel*) (Rosikiewicz i in., 2017). Wartości ekspresji z poziomów transkryptów przekonwertowano na poziom genów (dodano wartości ekspresji wszystkich transkryptów danego genu) a następnie dane zaimportowano do środowiska R (wersja 2.1.6) przy pomocy pakietu tximport (<https://bioconductor.org/packages/release/bioc/html/tximport.html>).

Tabela 4. Przykładowe podsumowanie przetwarzania danych RNA-Seq z repozytorium NCBI SRA (*dotyczy pierwszej próby z każdego z pakietów danych).

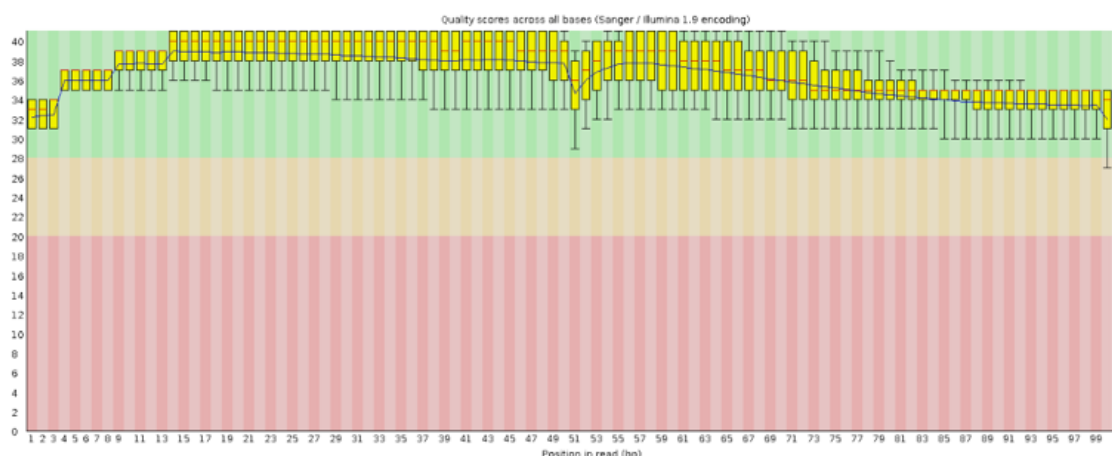
Możliwe rodzaje bibliotek dla odczytów sparowanych (ang. *paired-end*): IU-niespecyficzna niciowo, ISF- pierwszy odczyt pochodzi z nici wiodącej, ISR- pierwszy odczyt pochodzi z nici przeciwnej; dla odczytów niesparowanych (ang. *single-end*): SF- pierwszy odczyt pochodzi z nici wiodącej; SR- pierwszy odczyt pochodzi z nici przeciwnej.

Pakiet danych RNA-Seq	Rodzaj odczytów	Długość odczytów (pz)	Rodzaj biblioteki	Wejściowa liczba odczytów*	Liczba odczytów po filtrowaniu*	Współczynnik mapowania*
Nowotwór piersi (PRJNA251383) (Varley i in., 2014)	Paired-end	50	IU	131 191 930	105 582 550	65%
Rak jelita grubego (PRJNA218851) (S.-K. Kim i in., 2014)	Paired-end	101	IU	78 493 548	42 122 512	93%
Gruźlakorak płuc (PRJEB2784)(Y. S. Ju i in., 2012b)	Paired-end	101	IU	45 688 760	26 456 920	86%
Rak wątrobowokomórkowy (PRJNA310012) (G. Liu i in., 2016)	Paired-end	100	IU	93 946 356	60 199 452	89%
Wątrobiak zarodkowy (PRJNA413799) (Hooks i in., 2018)	Paired-end	125	ISR	169 073 930	144 450 926	85%
Inwazja łożyska (PRJNA562073) (Kshitiz i in., 2019)	Paired-end	50	ISR	120 980 520	101 485 886	94%
Rak gruczołu mlekowego psa (PRJNA489087) (T.-M. Kim i in., 2020)	Paired-end	101	ISR	74 696 512	68 308 274	67%
Rak gruczołu mlekowego psa (PRJNA561580) (Graim i in., 2020)	Single-end	141	SF	131 221 568	66 584 592	64%
LMH Linia komórkowa raka wątrobowokomórkowego kury (PRJNA239321) (Sun i in., 2015)	Single-end	50	SR	22 491 917	19 094 035	72%
Próbki zdrowej wątroby kury (PRJNA362606) (Q. Tang i in., 2017, s.)	Paired-end	125	ISR	48 426 516	38 814 304	80%

a) ❌ Per base sequence quality



b) ✅ Per base sequence quality



Rycina 6. Kontrola jakości odczytów pochodzących z próby SRR3129838 (rak wątrobowokomórkowy) przed (a) i po oczyszczeniu z adapterów, odczytów niskiej jakości oraz rRNA (b).

Z repozytorium TCGA dla każdej próby pobrano dane dotyczące liczby surowych odczytów oraz wartości TPM (ang. *Transcript Per Milion*) (Goldman i in., 2020; Grossman i in., 2016) z użyciem pakietu TCGAbiolinks (Colaprico i in., 2016). Dane te były następnie wykorzystane do analizy ekspresji różnicowej.

3.1.4. Analizy ekspresji różnicowej

Geny o zmienionym poziomie ekspresji (DEGs ang. *Differential Expressed Genes*) zidentyfikowano przy użyciu pakietu DESeq2 (Love i in., 2014), porównując próby nowotworowe i próby zdrowe. Z analiz usunięto wartości odstające tylko tam, gdzie

znajdowały się one powyżej wartości 100 na wykresie głównych składowych (PCA ang. *Principal Component Analysis*), a jednocześnie opis kliniczny nie wyjaśniał tej zmienności.

Ustalono próg istotności $p_{adj} < 0,05$ i $|\log_2FC| > 0,58$ (p_{adj} , ang. *adjusted p-value*; $\log_2FC$, ang. *log2 of fold change*). Następnie przefiltrowano wyniki by uzyskać retrokopie o istotnej statystycznie zmianie ekspresji. Dodatkowo filtrowano wyniki na podstawie różnic w średnich wartościach TPM między nowotworem a zdrową tkanką ($|\text{różnica TPM}| > 10$), tak by wyniki odzwierciedlały jedynie biologicznie istotne zmiany w poziomie ekspresji. Miało to na celu odfiltrowanie przypadków, w których wartości ekspresji były bardzo małe i choć zmiany były istotne statystycznie, biologicznie zmiany te były nieistotne. Dla przykładu, retrogen *MSL3P1* istotnie zwiększył swoją ekspresję ($\log_2FC=2,34$; $p_{adj}=6.41 \text{ e-}55$), przy czym różnica ta wynosiła 3,3 TPM – w tkance zdrowej jego średnia ekspresja wynosiła 0,86 TPM a w tkance nowotworowej 4,16 TPM co oznacza, że poziom retrogeny w obu przypadkach był bardzo mały.

3.1.5. Analizy korelacji ekspresji

Aby obliczyć współczynniki korelacji pomiędzy poziomem ekspresji retrokopii i genami rodzicielskimi/genami gospodarzy, wykorzystano funkcję *corr.test* z testem Spearmana z pakietu *psych* (R, wersja 2.1.6). Aby korelacja została uznana za istotną wymagane było: $p < 0,001$ i $|\rho| > 0,25$.

3.1.6. Potencjalne miejsca wiązania miRNA

Miejsca wiązania miRNA ustalono z użyciem bazy danych miRDB (Chen & Wang, 2020; Liu & Wang, 2019).

3.1.7. Wizualizacja danych

Wykresy pudełkowe przygotowano za pomocą pakietu *ggplot2* (Wickham, 2009) oraz *ggsignif* z użyciem testu t-studenta (Ahlmann-Eltze & Patil, 2021). Wykresy typu mapa cieplna (ang. *heatmap*) wykonano z użyciem narzędzia *ComplexHeatmap* (Gu i in., 2016, Gu, 2022). Wykresy wyników z eksperymentów qPCR zostały wykonane w GraphPad Prism 8 (GraphPad, USA).

3.1.8. Analiza filogenetyczna

Do określenia czasu powstania retrogenów wykorzystano bazę danych GenTree (<http://gentree.ioz.ac.cn>) (Shao i in., 2019). Aby potwierdzić czas powstania retrokopii

specyficznych dla człowieka, sprawdzono również ortologi w bazach RetrogeneDB2 (Rosikiewicz i in., 2017), NCBI Orthologs (O’Leary i in., 2016) i Ensembl Comparative (Herrero i in., 2016). Przy użyciu programu BLASTN (Camacho i in., 2009) sprawdzono także, czy istnieją ludzkie ortologi retrogenów związanych z nowotworami specyficznych dla psa i kury oraz obliczono podobieństwo sekwencji pomiędzy retrogenami różnych gatunków zwierząt (człowiek, szympan, mysz, pies, kura). Programu BLASTP (Johnson i in., 2008) użyto do określenia podobieństwa białek kodowanych przez wybrane retrogeny.

3.1.9. Testy statystyczne

W celu końcowego zweryfikowania postawionych hipotez przeprowadzono testy statystyczne z wykorzystaniem funkcji *chisq.test* z pakietu *stats* (R wersja 4.2.2) (R: The R Project for Statistical Computing, 2025).

3.2. Analizy eksperymentalne

3.2.1. Materiał biologiczny

Ekspresję retrogenów związanych z nowotworami zweryfikowano w ludzkich liniach komórkowych: MCF-7 (linia komórkowa ludzkiego estrogenozależnego gruczolakoraka piersi), MCF-12F (linia komórkowa ludzkiego gruczołu mlekowego), NCI-H1703 (linia komórkowa niedrobnokomórkowego raka płuc) i MRC-5 (ludzkie fibroblasty płuca). Ekspresję retrokopii sprawdzono także w panelu ludzkiego cDNA pochodzącego z 8 narządów (Human MTC™ Panel I, Takara Bio), panelu cDNA pochodzącego z 6 narządów psa (Dog Tissue CDNA, Zyagen) i cDNA pochodzącego z 6 narządów kury (Chicken Tissue CDNA, Zyagen) oraz w świeżo mrożonych próbach nowotworowych i próbach tkanek okalających pobranych od psów i szczurów. Próby do analiz pozyskano z kliniki weterynaryjnej „Klinika Małych Zwierząt Wąsiatycz” w Poznaniu. Uzyskano zgodę Inspekcji Weterynaryjnej Powiatowego Lekarza Weterynarii w Poznaniu (decyzja nr 821/2021).

3.2.2. Hodowle komórkowe

Linie komórkowe utrzymywano w odpowiednich dla danej linii pożywkach (MCF-7 i NCI-H1703 w pożywce RPMI; MCF-12F w pożywce DMEM F12; MRC-5 w pożywce DMEM) w temperaturze 37°C, 5% CO₂, w inkubatorze z systemem nawilżającym powietrze (zgodnie z wymaganiami ATCC).

3.2.3. Izolacja RNA i odwrotna transkrypcja

Całkowity RNA ekstrahowano przy użyciu odczynnika TRI Reagent (Molecular Research Center, USA) i poddano odwrotnej transkrypcji za pomocą cDNA RevertAid KIT (Qiagen, Niemcy) zgodnie z protokołami producenta.

3.2.4. PCR

Startery specyficzne dla poszczególnych genów zaprojektowano w PrimerBlast (Ye i in., 2012) (Tabela 5,

Tabela 6) i zamówiono w Merck Life Science Sp.z.o.o. Amplifikacja PCR została przeprowadzona przy użyciu EconoTaq PLUS2X Master Mix (Lucigen, USA) w objętości 10 µl. 5 µl produktów PCR zostało użyte do elektroforezy żelowej w 1,5% żelu agarozowym zawierającym GelRed i 1x bufor TAE. Zdjęcia żelu zostały wykonane za pomocą G:Box EF2 (Syngene, UK) i oprogramowania analitycznego GeneSys (Syngene, UK). Pozostałe 5 µl produktów PCR zostało poddane oczyszczeniu enzymatycznemu odczynnikiem ExoSAP-IT™ PCR Product Cleanup Reagent (ThermoScientific, USA) i sekwencjonowane przy użyciu Big Dye V3.1 Terminator Kit w analizatorze ABI Prism 3130XL Analyzer (Applied Biosystems, UK). Do analizy zsekwencjonowanych produktów PCR wykorzystane zostało oprogramowanie BioEdit (Hall, 1999).

Tabela 5. Startery PCR stosowane w analizach ekspresji genów człowieka.

Gen	Nazwa startera	Sekwencja startera 5' – 3'
<i>CALM3</i>	CALM3_F1_127	AGGGAAAGTAGTCCGGCGA
	CALM3_R1_127	TGCAATCTGCTCCTCAGTCA
<i>CALML5</i>	CALML5_221_F	GAAGTGTCTGGGACAGACGC
	CALML5_221_R	GGCATTGATGGTGCCGTTTC
<i>HMGB1</i>	HMGB1_F1_193	ACCGATAGGAAACGAGGCTG
	HMGB1_R1_193	TGACATTTTGCCTCTCGGCT
<i>HMGB1P6</i>	HMGB1P6_F_159	TTTTTGTGCAAACGTCGGGA
	HMGB1P6_R_159	TAATGGGCCTTGTCCACCTT
<i>HSPA2</i>	HSPA2_F1_131	ATCATCGCCAACGACCAGG
	HSPA2_R1_131	TTGGCGTCGAAGATGGTGTT
<i>PLEKHA7</i>	PLEKHA7_F1_134	CGGGACACTTTACCTGAGCA
	PLEKHA7_R1_134	ATCATGTGGCCCGAGTTGAC
<i>retro_hsap_852</i>	hsap_852_F2_168	TGGAGAACAGATCAGCGGAAG

	hsap_852_R2_168	GCGGAAAGAAGCCAGGTTTT
<i>retro_hsap_1605</i>	retro_hsap_1605_UNI_F_182	AGCTCTTGATGTTATCTCTGCACT
	retro_hsap_1605_UNI_R_182	CACGATAACTGTGCGTCCCT
<i>RHOB</i>	RHOB_UNI_F2_145	CAAGAAGCTGGTGGTGGTGG
	RHOB_UNI_R2_145	CACCTGCTTGCCGTCCA
<i>RPL36A</i>	RPL36A_F1_154	ATAAGAAGAGAAAGGGCCAAGTG
	RPL36A_R1_154	TCCCCACAAGCGGAATTGTATT
<i>RPL9</i>	RPL9_F3_193	CCCCTCTCGGAACCTTGTC
	RPL9_R3_193	GAATTACGAGGCCAGCAGT
<i>THSD4</i>	THSD4_F1_190	GTACAGTGAGGGAGAGCCGA
	THSD4_R1_190	GCTGTTTGAACGTGGAGCAA

Tabela 6. Startery PCR stosowane w analizach ekspresji genów psa, kury i szczura.

Gen	Nazwa startera	Sekwencja startera 5' – 3'
<i>CALML5 pies</i>	CALML5_dog_F_191	GACGGGCAAGTGAACCTACGA
	CALML5_dog_R_191	AAGAGAGGCCATGCTTTCCC
<i>CALML5 szczur</i>	CALML5_rat_131_F1	TGTGTGGGAACTACCTCATGG
	CALML5_rat_131_R1	GACCAACCCAAAGAAACCGTC
<i>HSPA2 pies</i>	HSPA2_dog_F_149	AGGACGACATTGACCGGATG
	HSPA2_dog_R_149	CTTGCCCCTGAGTTTCTCGT
<i>HSPA2 kura</i>	HSPA2_chicken_F_158	ACGTATTCTTGCGTGGGTGT
	HSPA2_chicken_R_158	GTGTTGGTGGGGTTCATTGC
<i>HSPA2 szczur</i>	HSPA2_rat_F_110	CCTAACGTTGCTTTGCCTGT
	HSPA2_rat_R_110	CTTGCCATGTTGGAAGACCC
<i>RHOB pies</i>	RHOB_UNI_F2_145	CAAGAAGCTGGTGGTGGTGG
	RHOB_UNI_R2_145	CACCTGCTTGCCGTCCA
<i>RHOB kura</i>	RHOB_chicken_F_154	TCTCTCCTACCCAGACACGG
	RHOB_chicken_R_154	GCGCAGGTCTTTCTTGTTGG
<i>retro_cfam_1036</i>	cfam_1036_F1_130	GGATGACCTCCTTTGCACGA
	cfam_1036_R1_130	CTGCTTGCGTGTGGCAAT
<i>retro_cfam_1061</i>	cfam_1061_F1_148	CTGAAGAATCCGCACCCTGA
	cfam_1061_R1_148	CCGAATATGCCACGGACAGT
<i>retro_ggal_15</i>	ggal_15_F2_156	ATGTCAACCAGTTACGGAGGG
	ggal_15_R2_156	ATCATTTGCATGCGCCATCTG

3.2.5. Wyciszanie ekspresji

Ekspresję retrogenów wyciszono w ludzkich liniach komórkowych, by zbadać jego wpływ na ekspresję genu rodzicielskiego lub genu gospodarza. Mieszanina transfekcyjna antysensownych LNA gapmeRs (Qiagen, Niemcy) (Tabela 7) została przygotowana w Opti-MEM (Gibco, USA) i Lipofectamine 3000 (Invitrogen™, USA). Mieszanina transfekcyjna została dodana do hodowanych komórek. Komórki zostały pobrane do analizy po 72 godzinach od momentu transfekcji. Stężenie poszczególnych gapmeRów wynosiło *CALML5*=50nM, *retro_hsap_852*=30nM, *HMGB1P6*=50nM. Wyciszenie ekspresji *MALAT1* (ang. *metastasis associated lung adenocarcinoma transcript 1*) stanowiło kontrolę pozytywną gapmeRa. Kontrole negatywne obejmowały sprawdzenie ewentualnego wpływu lipofektaminy (LIPO) na ekspresję genów oraz wpływu samego gapmeRa – kontrola producenta (NEG).

Tabela 7. Sekwencje gapmeRów użyte do wyciszenia ekspresji poszczególnych retrogenów.

Gen	Nazwa gapmeRa	Sekwencja 5' – 3'
<i>CALML5</i>	CALML5	ACTCGGCAGTCCTTTA
<i>HMGB1P6</i>	HMGB1P6_flank_200	GGTCGGGAATGGTATG
<i>Retro_hsap_852</i>	Human_852_5'_200nt_2	CGGAAAGAAGCCAGGT

3.2.6. qPCR

RNA został wyizolowany i przepisany na cDNA, jak opisano powyżej. qPCR został przeprowadzony na platformie QuantStudio™ 7 Flex Real-Time PCR System (Thermo Fisher Scientific, Waltham, USA) przy użyciu PowerUp™ SYBR Green Master Mix (Applied Biosystems, UK) zgodnie ze standardowym protokołem. Ekspresja *ActB* (ang. *actin beta*) stanowiła kontrolę jako gen o konstytutywnej ekspresji. Eksperymenty zostały przeprowadzone w trzech powtórzeniach technicznych dla trzech powtórzeń biologicznych. Względne ilościowe oznaczenie ekspresji genów zostało przeprowadzone metodą $2^{-\Delta\Delta Ct}$ (Livak & Schmittgen, 2001).

3.2.7. Test żywotności

Wpływ wyciszenia retrogeny na żywotność komórek był analizowany przy użyciu TrypanBlue (Merck) zgodnie z protokołem producenta. Do 0,1 ml roztworu komórek dodano 0,1 ml TrypanBlue. Procentową zawartość żywych komórek policzono z użyciem The Countess™ 3 Automated Cell Counter.

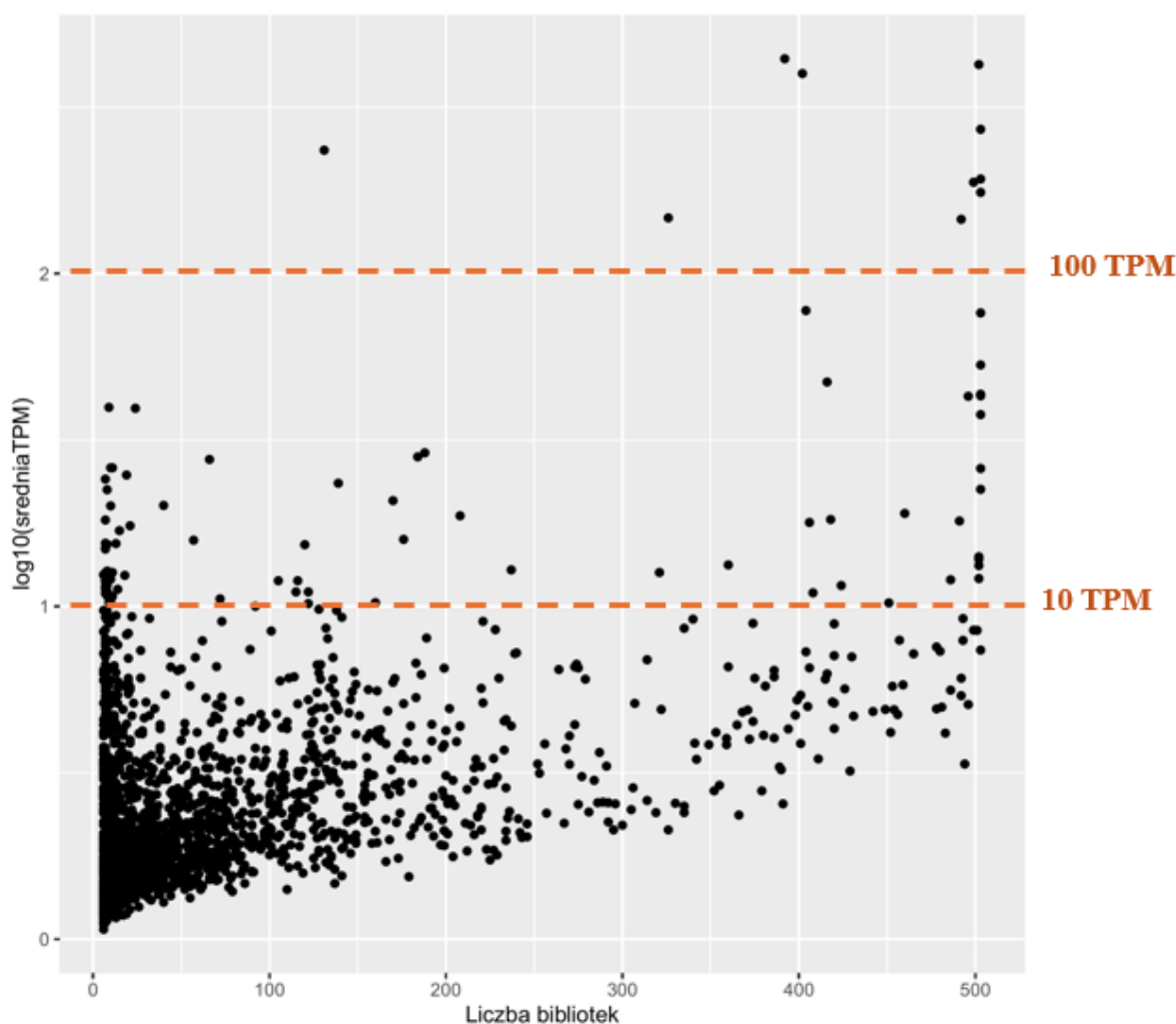
WYNIKI

3.3. Identyfikacja retrokopii o zmienionej ekspresji w nowotworach człowieka, psa i kury

3.3.1. Szacowanie poziomu ekspresji retrokopii

W analizach danych transkryptomicznych człowieka wzięto pod uwagę 62 319 kodujących białka i niekodujących genów, wliczając retrokopie. Uwzględnionych było 4580 retrokopii, w tym: 4486 retrokopii adnotowanych zarówno w RetrogeneDB2 jak i Ensembl, 83 retrokopie adnotowane jedynie w bazie RetrogeneDB2 (tzw. *novel*) oraz 11 retrokopii związanych z procesem nowotworowym wybranych na podstawie doniesień w literaturze (nieadnotowane w RetrogeneDB2).

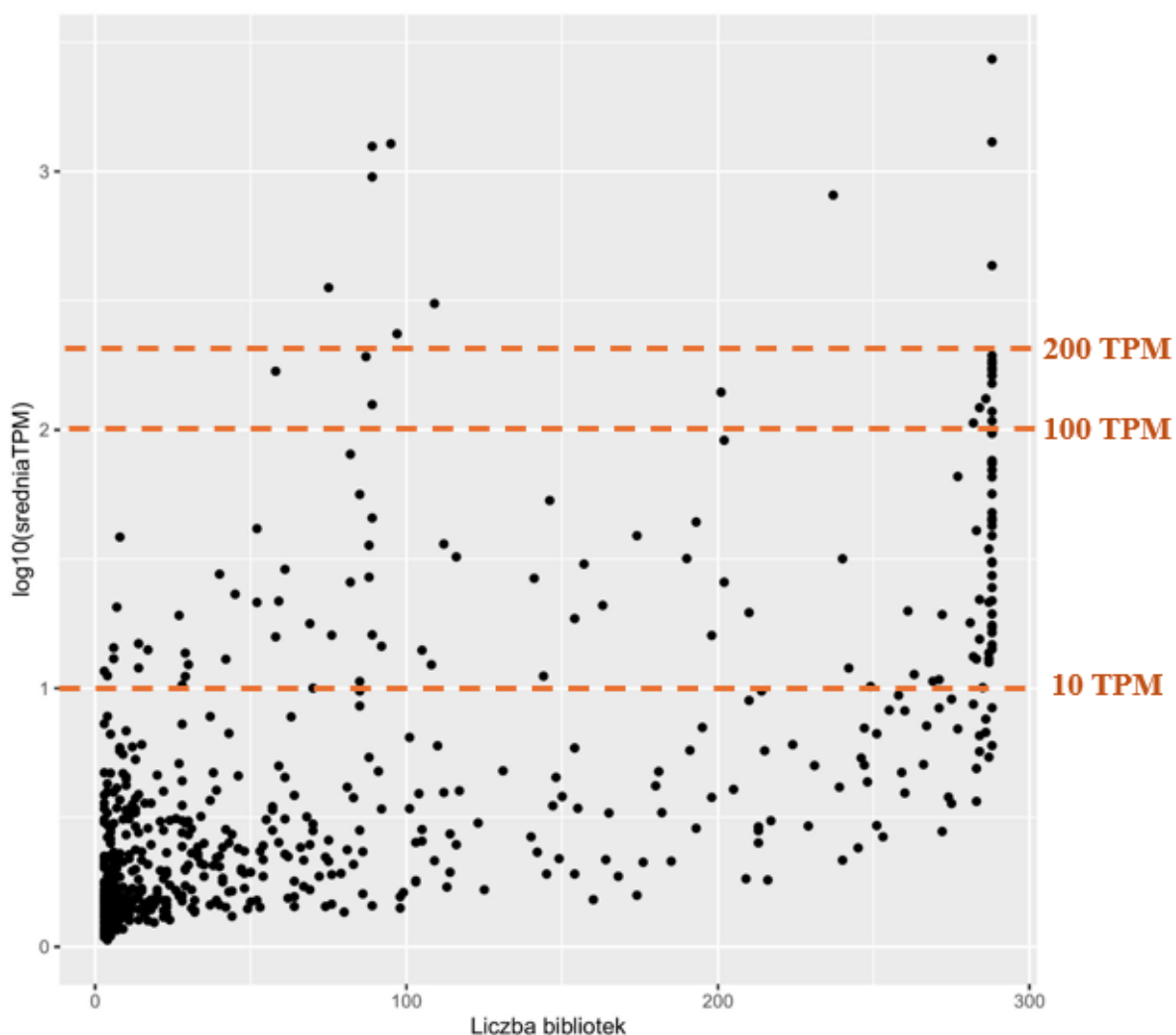
Ekspresję retrokopii w tkankach człowieka sprawdzono w 506 bibliotekach RNA-Seq. W celu wyeliminowania artefaktów wzięto pod uwagę tylko retrokopie, które wykazywały ekspresję minimum 1 TPM w co najmniej 1% bibliotek. W przypadku człowieka było to 6 bibliotek. Średnia wartość TPM dla retrokopii uznanych za ulegające ekspresji została wyliczona tylko dla prób, w których ekspresja danego retrogeny wynosiła minimum 1 TPM. Analiza wskazała 1851 retrogenów wykazujących ekspresję oraz 2729 retrokopii niespełniających przyjętych warunków. 143 retrokopie wykazywały ekspresję w ponad połowie (powyżej 253) prób, a 89 retrokopii wykazywało średnią ekspresję minimum 10 TPM (Rycina 7). Do retrogenów o najwyższej średniej wartości ekspresji, przekraczającej 100 TPM (od 146 do 442 TPM) należało 10 retrogenów: *RPS2P5* (ENSG00000240342), *H3P6* (ENSG00000235655), *RPL36AL* (ENSG00000165502), *RPS28P7* (ENSG00000227097), *PCBPI* (ENSG00000169564), *CALML5* (ENSG00000178372), *RHOB* (ENSG00000143878), *RPL23AP4* (ENSG00000234851), *retro_hsap_852* (ENSG00000244398) oraz *RPL21P1* (ENSG00000220842).



Rycina 7. Ekspresja retrokopii człowieka w analizowanych próbach (NCBI SRA). Liczba bibliotek, w których retrokopia wykazywała ekspresję minimum 1 TPM (oś x) oraz średnia wartość ekspresji (oś y).

W przypadku analiz danych transkryptomycznych pochodzących z tkanek psa wzięto pod uwagę 32 270 genów wliczając 2091 retrokopii. Wśród retrokopii 611 adnotowanych było w obu bazach, Ensembl i RetrogeneDB, a 1480 retrokopii adnotowanych było jedynie w bazie RetrogeneDB (tzw. *novel*). Podobnie jak w przypadku analiz tkanek człowieka, warunkiem uznania, że dana retrokopia ulega ekspresji była wartość minimum 1 TPM w co najmniej 1% bibliotek (w tym przypadku w 3 bibliotekach). Średnia wartość TPM liczona była jedynie dla retrokopii spełniających wcześniejsze kryterium i liczona została tylko dla prób, w których ekspresja wynosiła minimum 1 TPM. W analizach bibliotek RNA-Seq psa (288 prób), 591 retrokopii wykazywało ekspresję (Rycina 8). 142 retrokopie ulegały ekspresji w ponad połowie prób. W przypadku 27 retrogenów średnia wartość TPM przekraczała 100 i wynosiła pomiędzy 106 a 2722 TPM. Do retrogenów o najwyższej ekspresji (średnia wartość TPM > 200) należało

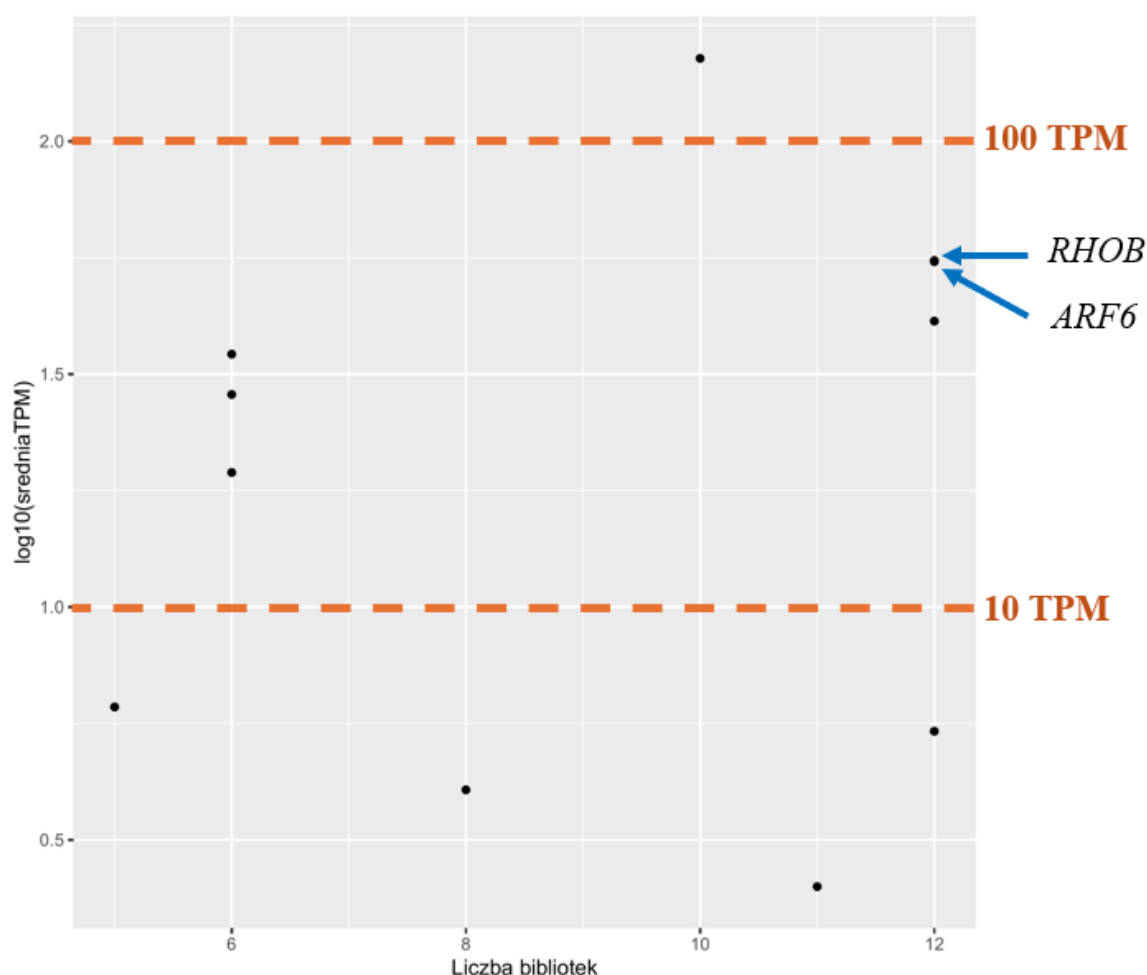
10 retrogenów: *retro_cfam_203* (ENSCAFG00000005494), *retro_cfam_1011* (ENSCAFG000000004857), *CALML5* (ENSCAFG00000005235), *retro_cfam_113* (ENSCAFG000000024837), *retro_cfam_36* (ENSCAFG000000016065), *retro_cfam_145* (ENSCAFG000000010850), *retro_cfam_1671* (ENSCAFG000000018652), *retro_cfam_1642* (ENSCAFG000000024420), *retro_cfam_2167* (ENSCAFG000000017231), *retro_cfam_150* (ENSCAFG000000030465).



Rycina 8. Ekspresja retrokopii psa w analizowanych próbach (NCBI SRA). Liczba bibliotek, w których retrokopia wykazywała ekspresję minimum 1 TPM (oś x) oraz średnia wartość ekspresji (oś y).

W analizach danych kury wzięto pod uwagę 24 194 genów wliczając retrokopie. Ostateczna liczba retrokopii uwzględnionych w analizie wynosiła 25, w tym 10 retrokopii adnotowanych zarówno w bazie Ensembl jak i RetrogeneDB oraz 15 retrokopii adnotowanych jedynie w bazie RetrogenDB (tzw. *novel*). W przypadku analiz 12 bibliotek RNA-Seq kury

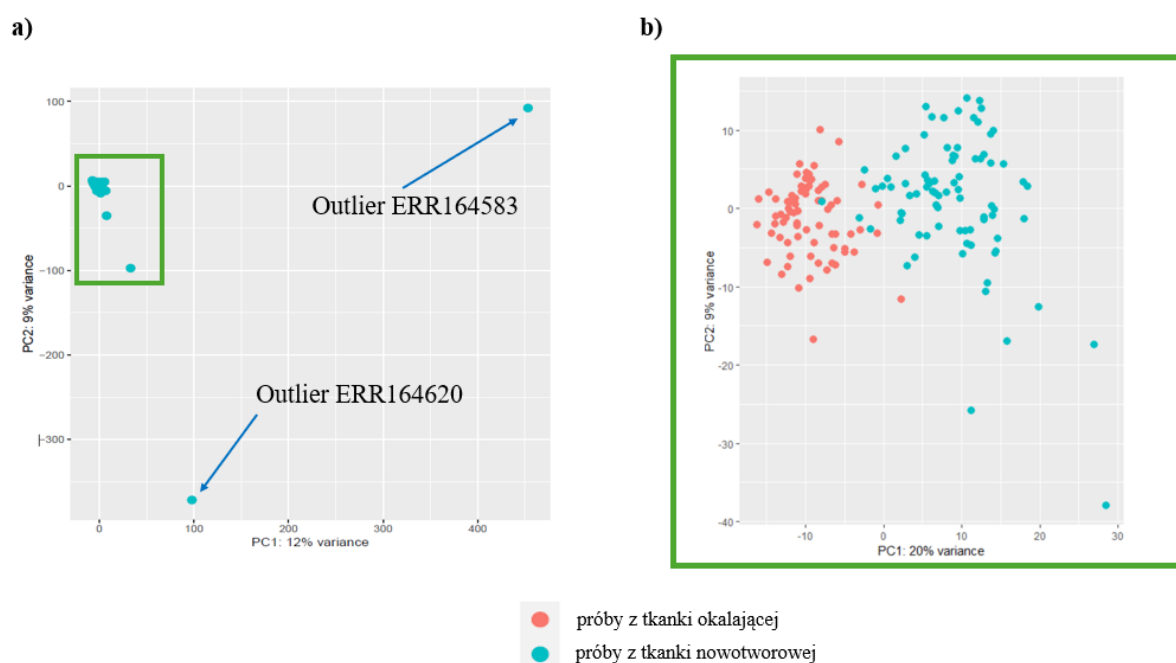
zastosowano inne kryterium ze względu na bardzo małą liczbę bibliotek – retrogen uznano za ulegający ekspresji, gdy wartość TPM wynosiła minimum 1 TPM, w co najmniej dwóch próbach (16% bibliotek). Średnia wartość TPM dla danej retrokopii, podobnie jak w przypadku człowieka i psa, została wyliczona tylko dla retrokopii określonych jako ulegające ekspresji i jedynie na podstawie prób, w których ekspresja przekraczała 1 TPM. Zgodnie z tymi kryteriami, w analizowanych próbach ekspresji ulegało 11 retrokopii (Rycina 9). Retrogeny *RHOB* (ENSGALG00000016485) i *ARF6* (ENSGALG00000026944) pokrywają się na Rycinie 9 – obydwie wykazywały średnią ekspresję na podobnym poziomie. W przypadku 7 retrokopii średni poziom ekspresji był większy niż 10 TPM. Do retrogenów o najwyższej ekspresji, w zakresie średniej wartości, przekraczającej 100 TPM, należał retrogen *retro_ggal_21*.



Rycina 9. Ekspresja 11 retrokopii kury w 12 analizowanych bibliotekach (NCBI SRA). Liczba bibliotek, w których retrokopia wykazywała ekspresję minimum 1 TPM (oś x) oraz średnia wartość ekspresji (oś y).

3.3.2. Identyfikacja retrogenów o zmienionym poziomie ekspresji w nowotworach człowieka

Przed rozpoczęciem szczegółowych analiz wykonano analizę składowych głównych w celu identyfikacji prób, które odstają od pozostałych i mogłyby zaburzać późniejsze analizy. Przed analizą ekspresji różnicowej usunięto następujące próby „odstające” (ang. *outliers*): PRJNA251383 – SRR1313203, SRR1313175; PRJEB2784 – ERR164583, ERR164620; PRJNA561580 – SRR10012983; PRJNA489087 – SRR10012983. Przykład dwóch odstających prób przedstawiono na rycinie 10.



Rycina 10. Przykład identyfikacji prób odstających na podstawie analizy głównych składowych dla gruczolaka płuca PRJEB2784. Wykres PCA przed usunięciem (a) oraz po usunięciu prób ERR164583 i ERR164620 (b).

Na podstawie wyników analiz bioinformatycznych, udało się zidentyfikować 33 retrogeny o istotnie zróżnicowanym poziomie ekspresji w co najmniej jednym z pięciu ludzkich nowotworów (Tabela 8). W przypadku raka piersi (BRCA) oraz wątrobiaka zarodkowego (HB) wszystkie dwa typy nowotworu zostały ujęte wspólnie jako próby nowotworowe. Szczegółowe informacje dotyczące retrogenów o zmienionej ekspresji (log2foldchange, wartości p oraz zmiany w wartościach TPM) znajdują się w tabelach na końcu niniejszej pracy (Aneks, Tabele 21 do 25).

Tabela 8. Retrogeny o zmienionym poziomie ekspresji w 5 ludzkich nowotworach (dane NCBI SRA). Czerwoną strzałką oznaczono zwiększony poziom ekspresji a zieloną zmniejszony poziom ekspresji. BRCA- rak piersi, COAD- gruczolakorak jelita grubego, HB- wątrobiak zarodkowy, HCC- rak wątrobowokomórkowy, LUAD- gruczolakorak płuc.

Retrogen	Gen rodzicielski	BRCA	COAD	HCC	LUAD	HB
<i>ALDH1B1</i>	<i>ALDH2</i>	NA	↑	↓	NA	↓
<i>BTF3L4P2</i>	<i>BTF3L4</i>	NA	NA	NA	↑	NA
<i>CALML5</i>	<i>CALM3</i>	↑	NA	NA	NA	NA
<i>CCT8P1</i>	<i>CCT8</i>	NA	NA	NA	NA	↓
<i>CHMP1B</i>	<i>CHMP1A</i>	NA	↓	NA	NA	NA
<i>FOSL1P1</i>	<i>FOSL1</i>	↑	NA	NA	NA	NA
<i>FTLP14</i>	<i>FTL</i>	↑	NA	NA	NA	NA
<i>H3P6</i>	<i>H3-3B</i>	↑	NA	NA	NA	NA
<i>HMGB1P3</i>	<i>HMGB1</i>	↑	NA	NA	NA	NA
<i>HMGB1P5</i>	<i>HMGB1</i>	↑	NA	NA	NA	NA
<i>HMGB1P6</i>	<i>HMGB1</i>	↑	↑	NA	NA	NA
<i>HMG4</i>	<i>HMG2</i>	NA	NA	↑	NA	NA
<i>LRRFIP1P1</i>	<i>LRRFIP1</i>	↑	NA	NA	NA	NA
<i>MEIS3P1</i>	<i>MEIS3</i>	↓	NA	NA	NA	NA
<i>NXT1</i>	<i>NXT2</i>	NA	NA	↑	NA	NA
<i>PA2G4P4</i>	<i>PA2G4</i>	↑	NA	NA	NA	NA
<i>PPP1R14BP3</i>	<i>PPP1R14B</i>	↑	NA	NA	NA	NA
<i>retro_hsap_154</i>	<i>TMA7</i>	↓	NA	↓	NA	NA
<i>retro_hsap_1605</i>	<i>RPL9</i>	↓	NA	↑	↓	NA
<i>retro_hsap_2297</i>	<i>PPIA</i>	↑	NA	NA	NA	NA
<i>retro_hsap_3179</i>	<i>FAM133B</i>	↑	NA	NA	NA	NA
<i>retro_hsap_3325</i>	<i>RPL31</i>	↑	NA	NA	NA	NA
<i>retro_hsap_794</i>	<i>RPL31</i>	↑	NA	NA	NA	NA
<i>retro_hsap_852</i>	<i>RPL36A</i>	NA	↑	↑	↑	NA
<i>RHOB</i>	<i>RHOA</i>	↓	NA	↓	↓	↓
<i>RHOG</i>	<i>RAC1</i>	NA	NA	↑	NA	NA
<i>RPL23AP42</i>	<i>RPL23A</i>	NA	NA	↑	NA	NA
<i>RPL21P16</i>	<i>RPL21</i>	↓	NA	NA	NA	NA
<i>RPS20P33</i>	<i>RPS20</i>	↑	NA	NA	NA	NA
<i>RPS2P5</i>	<i>RPS2</i>	NA	NA	↑	NA	NA
<i>RRAGA</i>	<i>RRAGB</i>	↓	NA	NA	NA	NA
<i>SFN</i>	<i>YWHAQ</i>	NA	NA	↑	↑	NA
<i>TMEM30B</i>	<i>TMEM30A</i>	↓	↓	NA	NA	NA

Najwięcej retrogenów o zmienionej ekspresji zostało zidentyfikowanych w nowotworze piersi (22) a najmniej w wątrobiaku zarodkowym (3) (Tabela 8). 8 spośród wskazanych retrokopii zmieniło swoją ekspresję w co najmniej dwóch nowotworach człowieka. Retrokopie

TMEM30B, *RHOB* oraz *retro_hsap_154* to grupa retrokopii, które posiadają obniżony poziom ekspresji w co najmniej dwóch nowotworach. Natomiast *SFN*, *retro_hsap_852* a także *HMGB1P6* to przykłady retrogenów, których ekspresja wzrasta w co najmniej dwóch analizowanych zestawach danych. Większość retrokopii (25 z 33) reprezentuje geny, które zmieniają poziom ekspresji tylko w jednym nowotworze.

3.3.3. Retrogeny o zmienionej ekspresji w różnych typach i stadiach nowotworów człowieka

W celu wyselekcjonowania retrogenów o zmienionej ekspresji w różnych typach nowotworów człowieka przeprowadzono analizy ekspresji różnicowej porównując ekspresję w dwóch typach nowotworu piersi – ER+ (ang. *estrogen receptor positive*) oraz TNBC (ang. *triple-negative breast cancer*) oraz w dwóch subtypach molekularnych wątrobiaka zarodkowego – C1 oraz C2. Subtyp C2 to bardziej zaawansowane stadium nowotworu wątrobiaka zarodkowego z gorszym rokowaniem, w porównaniu z subtypem C1 (Hooks i in., 2018). Wyniki analiz wskazały *RHOB* i *RPS2P5* jako retrogeny o zmienionej ekspresji w bardziej inwazyjnym subtypie C2 (Tabela 9). Z kolei, w porównaniu TNBC i ER+ zidentyfikowano trzy retrogeny o zwiększonej ekspresji w bardziej złośliwym typie nowotworu piersi (TNBC): *CALML5*, *retro_hsap_3325* i *retro_hsap_852*. Wszystkie spośród nich zostały zidentyfikowane w opisaney wcześniej analizie jako retrokopie o zmienionej w nowotworach ekspresji. Ponadto, 7 retrogenów (*CALML3*, *FOSL1P1*, *HSPA2*, *RHOB*, *RPS20P33*, *TMEM30B*, *TOMM40P4*) wykazało zmniejszony poziom ekspresji w TNBC a wśród nich trzy: *CALML3*, *HSPA2* oraz *TOMM40P4* to niezidentyfikowane we wcześniejszej analizie przypadki.

W celu identyfikacji retrogenów o zmienionej ekspresji podczas rozwoju przerzutów gruczolakoraka jelita grubego porównano dane pochodzące z guzów przerzutujących z próbami z guzów pierwotnych. Zidentyfikowano jeden retrogen, który zmienił swoją ekspresję w tej inwazyjnej formie nowotworu (Tabela 9). Retrogenem tym jest *HSPA2* (ang. *heat shock protein family A (Hsp70) member 2*), którego zmianę ekspresji zaobserwowano dopiero w późniejszych stadiach progresji nowotworu jelita grubego, podobnie jak w bardziej inwazyjnym typie TNBC w przypadku raka piersi. Sugeruje to, że obniżenie poziomu ekspresji retrogeny *HSPA2* może być wskaźnikiem stopnia zaawansowania nowotworów w przypadku raka piersi i gruczolakoraka jelita grubego.

Tabela 9. Retrogeny o zmienionym poziomie ekspresji w różnych typach lub stadiach 3 ludzkich nowotworów (dane NCBI SRA). Czerwoną strzałką oznaczono zwiększony poziom ekspresji a zieloną zmniejszony poziom ekspresji. BRCA- rak piersi, ER+- rak piersi z pozytywnym statusem receptorów estrogenowych, TNBC- potrójnie ujemny rak piersi, COAD- gruczolakorak jelita grubego, HB- wątrobiak zarodkowy, C2, C1- subtypy molekularne wątrobiaka zarodkowego.

Retrogen	Gen rodzicielski	BRCA (TNBC vs ER+)	COAD (przerzuty vs guz pierwotny)	HB (C2 vs C1)
CALML3	CALM3	↓	NA	NA
CALML5	CALM3	↑	NA	NA
FOSL1P1	FOSL1	↓	NA	NA
HSPA2	HSPA8	↓	↓	NA
retro_hsap_3325	RPL31	↑	NA	NA
retro_hsap_852	RPL36A	↑	NA	NA
RHOB	RHOA	↓	NA	↓
RPS20P33	RPS20	↓	NA	NA
RPS2P5	RPS2	NA	NA	↑
TMEM30B	TMEM30A	↓	NA	NA
TOMM40P4	TOMM40	↓	NA	NA

3.3.4. Identyfikacja retrogenów o zmienionym poziomie ekspresji w nowotworach psa

Analogiczne analizy przeprowadzono dla dwóch innych gatunków kręgowców, psa i kury, używając dostępnych danych transkryptomicznych. W przypadku psa analizowano dwa zestawy danych z gruczołu mlekowego. Zestawy danych pochodzą z różnych laboratoriów więc analizowano je osobno ze względu na tzw. efekt wsadowy (ang. *batch effect*). Analiza ekspresji różnicowej danych RNA-Seq pierwszego z nich (PRJNA489087) pozwoliła na zidentyfikowanie 5 retrogenów o obniżonym poziomie ekspresji (*CALML5*, *retro_cfam_942*, *HSPA2*, *retro_cfam_179*, *RHOB*) oraz 3 retrogenów o zwiększonej ekspresji (*retro_cfam_1061*, *retro_cfam_1036*, *retro_cfam_320*) (Tabela 10). Trzy spośród nich – *RHOB*, *HSPA2* i *CALML5*, zidentyfikowano także we wcześniejszych analizach dotyczących człowieka. Dokładne informacje dotyczące retrogenów o zmienionej ekspresji ($\log_2\text{foldchange}$, wartości p oraz zmiany w wartościach TPM) znajdują się w aneksie do niniejszej pracy (Aneks, Tabele 19-20).

Tabela 10. Retrogeny o zmienionym poziomie ekspresji w nowotworach psa. Czerwoną strzałką oznaczono zwiększony poziom ekspresji a zieloną zmniejszony poziom ekspresji.

Retrogen	Gen rodzicielski	Rak gruczołu mlekowego (PRJNA489087)	Rak gruczołu mlekowego (PRJNA561580)
<i>CALML5</i>	<i>CALM3</i>	↓	NA
<i>retro_cfam_942</i>	<i>CAPNS1</i>	↓	NA
<i>HSPA2</i>	<i>HSPA8</i>	↓	NA
<i>retro_cfam_179</i>	<i>RPS3A</i>	↓	NA
<i>RHOB</i>	<i>RHOA</i>	↓	NA
<i>retro_cfam_1061</i>	<i>LDHA</i>	↑	NA
<i>retro_cfam_1036</i>	<i>ENSCAFG000000007199</i>	↑	NA
<i>retro_cfam_320</i>	<i>ENO2</i>	↑	↑

W przypadku drugiego zestawu danych psa (PRJNA561580) wyselekcjonowano tylko jeden retrogen o istotnie zmienionej ekspresji – *retro_cfam_320*, który wykazał zwiększony poziom ekspresji w tkance nowotworowej w porównaniu z tkanką zdrową (Tabela 10). Z opisu klinicznego wynika, że w próbach zdrowych widoczne były pierwsze zmiany predysponujące do kancerogenezy tj. atypowa hiperplazja przewodowa, co może tłumaczyć identyfikację tylko jednego retrogeny o zmienionej ekspresji. Niemniej jednak podkreśla to potencjalną rolę *retro_cfam_320* jako retrogeny związanego z nowotworem gruczołu mlekowego u tego gatunku.

3.3.5. Identyfikacja retrogenów o zmienionym poziomie ekspresji w nowotworze kury

Analizie poddano także dane RNA-Seq z linii komórkowej raka wątrobowokomórkowego kury (PRJNA239321, PRJNA362606). Ze względu na wcześniej już wspomniany efekt wsadowy bezpośrednie porównywanie danych z różnych laboratoriów w jednej analizie nie jest wskazane. Niestety były to jedyne dane, które były dostępne w repozytorium NCBI SRA dla bardziej odległego ewolucyjnie gatunku kręgowca. Zidentyfikowano 5 retrokopii o zmienionym w próbach nowotworowych poziomie ekspresji, w tym *RHOB*, *HSPA2*, *retro_ggal_14*, *retro_ggal_15* oraz *retro_ggal_18* (Tabela 11). Szczegółowe informacje dotyczące retrogenów o zmienionej ekspresji ($\log_2\text{foldchange}$, wartości p oraz zmiany w wartościach TPM) znajdują się w tabeli na końcu niniejszej pracy (Aneks, Tabela 31).

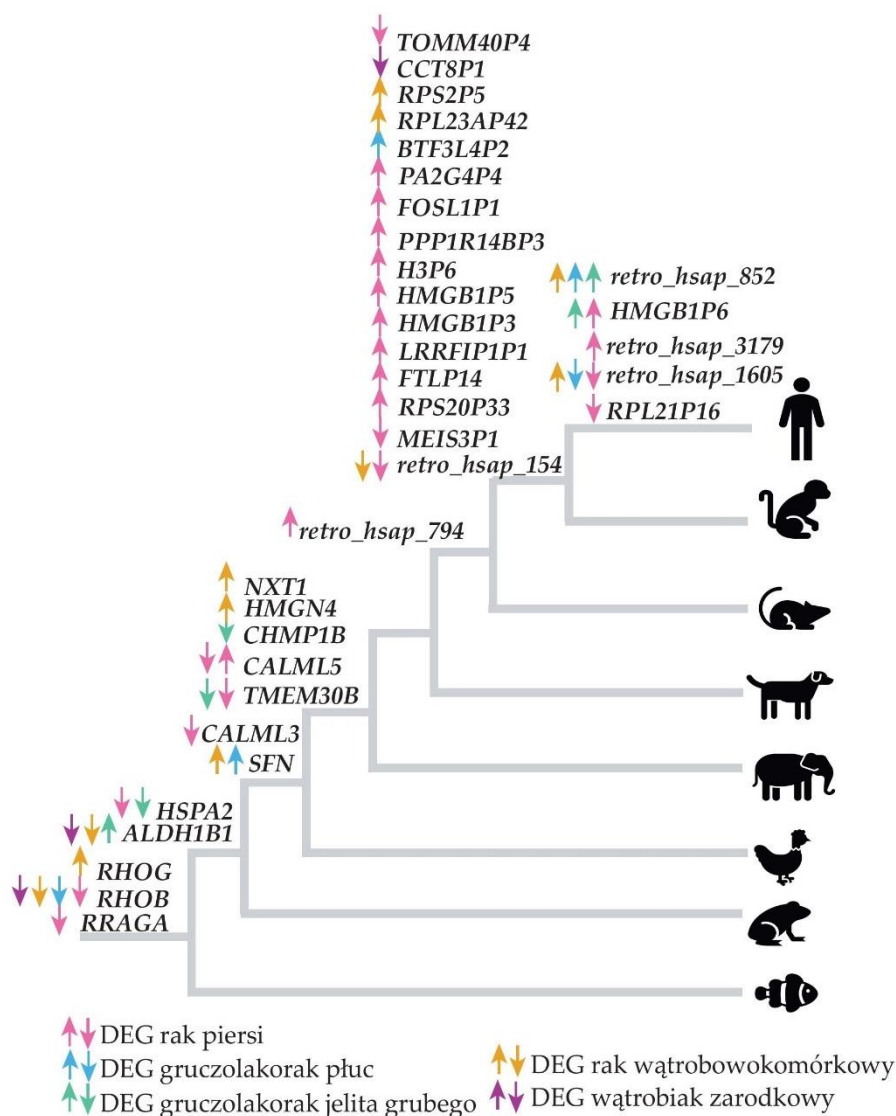
Tabela 11. Retrogeny o zmienionym poziomie ekspresji w nowotworze kury. Czerwoną strzałką oznaczono zwiększony poziom ekspresji a zieloną zmniejszony poziom ekspresji.

Retrogen	Gen rodzicielski	Rak wątrobowokomórkowy
<i>RHOB</i>	<i>RHOA</i>	
<i>HSPA2</i>	<i>HSPA8</i>	
<i>retro_ggal_15</i>	<i>C12orf45</i>	
<i>retro_ggal_18</i>	<i>Pou5f3</i>	
<i>retro_ggal_14</i>	<i>SCO1</i>	

3.4. Analiza ewolucyjna retrogenów o zmienionej ekspresji w nowotworach

3.4.1. Retrogeny związane z nowotworami wspólne dla człowieka, psa i kury

Retrokopie o zmienionym poziomie ekspresji w nowotworach człowieka poddano analizie ewolucyjnej. Baza danych GenTree pozwoliła na określenie czasu powstania 34 spośród 36 zidentyfikowanych retrokopii. Nie ustalono czasu powstania *retro_hsap_2297* oraz *retro_hsap_3325* ze względu na brak adnotacji w bazie danych GenTree (Shao i in., 2019). Większość spośród zidentyfikowanych retrokopii (21 spośród 34) o zmienionym w nowotworach człowieka poziomie ekspresji powstała podczas ewolucji naczelnych i należy do ewolucyjnie młodych genów (Rycina 11). Spośród nich 16 jest wspólnych dla naczelnych, a 5 specyficznych dla człowieka (*retro_hsap_852*, *HMGB1P6*, *retro_hsap_3179*, *retro_hsap_1605*, *RPL21P16*). Pozostałe 13 retrokopii powstało podczas wcześniejszej ewolucji kręgowców – wśród nich 6 należy do retrogenów specyficznych dla ssaków (*NXT1*, *HMGN4*, *RPL13AP17*, *CHMP1B*, *CALML5*, *TMEM30B*), a tylko trzy występują u wszystkich kręgowców (*RHOG*, *RHOB*, *RRAGA*).

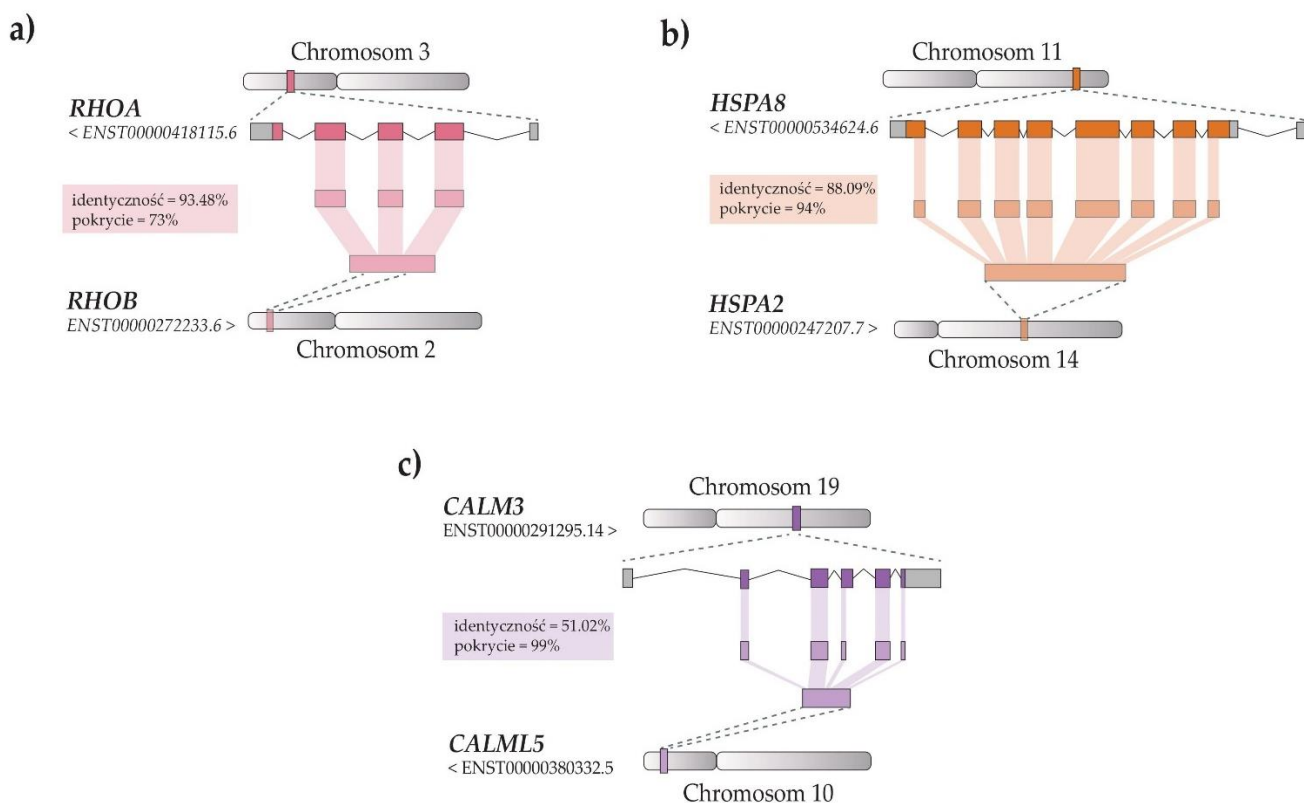


Rycina 11. Czas powstania retrogenów o zmienionej ekspresji w nowotworach człowieka (na podstawie GenTree <http://gentree.ioz.ac.cn>, brak *retro_hsap_2297* oraz *retro_hsap_3325* ze względu na brak adnotacji w bazie danych GenTree).

Strzałka w górę symbolizuje retrogen o zwiększonej ekspresji w nowotworze, a strzałka w dół retrogen o zmniejszonej ekspresji.

Przeprowadzone analizy pozwoliły na zidentyfikowanie trzech ewolucyjnie starych retrokopii: *RHOB* i *HSPA2*, które wykazały zmianę ekspresji w danych nowotworowych wszystkich trzech gatunków oraz retrogen specyficzny dla ssaków *CALML5* (ang. *calmodulin like 5*), o zmienionej ekspresji w nowotworach człowieka i psa. *CALML5* wykazał spadek ekspresji w łagodniejszym typie nowotworu piersi ER+, natomiast wzrost w bardziej agresywnym typie TNBC. W przypadku psa ekspresja tego retrogenu spadła w próbach nowotworowych. Z kolei, *HSPA2* i *RHOB* zostały zidentyfikowane jako retrogeny o

zmienionym poziomie ekspresji w nowotworach człowieka, psa i kury. U wszystkich gatunków oba retrogeny wykazały spójnie, zmniejszoną ekspresję w próbach nowotworowych. *RHOB* należy do znanych genów związanych z nowotworami. Według danych literaturowych ekspresja *RHOB* w większości nowotworów spada (Mazieres i in., 2004). Podobnie w 4 z analizowanych nowotworów człowieka retrogen ten również posiadał obniżoną ekspresję (Tabela 8). W przypadku raka piersi i wątrobiaka zarodkowego *RHOB* wykazał niższą ekspresję w bardziej zaawansowanych typach nowotworów – odpowiednio TNBC oraz w podklasie molekularnej wątrobiaka zarodkowego C2 (Tabela 9). Z kolei, *HSPA2* wykazywał znaczącą zmianę ekspresji tylko w bardziej agresywnym typie nowotworu piersi – TNBC (Tabela 9). W próbach nowotworu jelita grubego ekspresja *HSPA2* zmieniła się istotnie statystycznie pomiędzy różnymi postaciami raka, z większym spadkiem w próbach stanowiących guzy przerzutowe ($\log_2FC = -1,93$). W celu potwierdzenia powstania na drodze retropozycji trzech ewolucyjnie starych retrogenów, wspólnych dla nowotworów człowieka, psa i kury, porównano u człowieka ich sekwencje i strukturę genomową z ich genami rodzicielskimi (Rycina 12). Gen rodzicielski *RHOA* (ang. *ras homolog family member A*) i retrokopia *RHOB* znajdują się odpowiednio na chromosomach 3 i 2. W jednoeksonowej retrokopii *RHOB* zachowana jest większość eksonów kodujących, co zostało zobrazowane poprzez odpowiadające sobie fragmenty zaznaczone w postaci bloków (Rycina 12a). Identyczność sekwencji kodujących tych genów wynosi 93,48%, przy pokryciu równym 73%, co znaczy, że część sekwencji genu rodzicielskiego nie została w retrogenie zachowana. Pokrycie dotyczy sekwencji kodującej, bez sekwencji UTR (ang. *untranslated regions*). W przypadku pary *HSPA8* (ang. *heat shock protein family A (Hsp70) member 8*) i *HSPA2* również widoczne jest wysokie podobieństwo sekwencji, z identycznością na poziomie 88,09% i pokryciem wynoszącym 94% (Rycina 12b). Z kolei retrogen *CALML5* wykazuje znacznie mniejszy stopień podobieństwa do swojego genu rodzicielskiego *CALM3* (ang. *calmodulin 3*), w porównaniu do *RHOB* lub *HSPA2*, co znajduje odzwierciedlenie w niższym stopniu identyczności sekwencji (51,02%). Natomiast 99% pokrycie oznacza, że praktycznie cała sekwencja kodująca genu *CALM3* została zachowana w retrogenie (Rycina 12c).



Rycina 12. Porównanie struktury i lokalizacji retrokopii *RHOB* (a) *HSPA2* (b) i *CALML5* (c) z ich genami rodzicielskimi.

Wyniki w Tabeli 12. przedstawiają procentowe podobieństwo sekwencji aminokwasowych *RHOB*, *HSPA2*, *CALML5* w odniesieniu do wybranych gatunków kręgowców. Wysokie podobieństwo sekwencji pomiędzy retrogenami człowieka a retrogenami szympansa odzwierciedla to, że jest on najbliższym ewolucyjnym krewnym człowieka. *RHOB* należy do retrogenów o wysoce konserwatywnej sekwencji (Tabela 12). Najprawdopodobniej, retrogen ze względu na pełnioną funkcję jest pod większą presją selekcyjną. Podobnie w przypadku *HSPA2*, który także wykazuje duże podobieństwo sekwencji w badanych gatunkach. Może to sugerować, że zarówno białko *RHOB*, jak i *HSPA2* pełnią istotną rolę w podstawowych procesach biologicznych. Interesującym przypadkiem jest *CALML5*. Retrogen ten jest najmłodszym z analizowanej trójki, a jednocześnie wykazuje największe zróżnicowanie sekwencji. Niskie podobieństwo sekwencji tego retrogeny w porównaniu między człowiekiem a myszą i psem (odpowiednio 53,04% i 57,05%) może wskazywać na znacznie mniejszą presję ewolucyjną i specyficzne funkcje adaptacyjne tego białka u odległych ewolucyjnie gatunków.

Tabela 12. Porównanie podobieństwa sekwencji aminokwasowej RHOB, HSPA2 oraz CALML5 (BLASTP) między białkami człowieka i innych gatunków kręgowców.

id- identyczność, cov- pokrycie

Retrogeny człowieka	Szympan	Mysz	Pies	Kura
RHOB	id= 100% cov=100%	id=100% cov=100%	id=100 % cov=100%	id=97.45% cov=100%
HSPA2	id= 99.38% cov=100%	id=98.44% cov=100%	id=99.22% cov=100%	id=92.18 % cov=100%
CALML5	id=95.89% cov=100%	id=53.04 % cov=96%	id=57.05% cov=100%	-

4.2.2. Retrogeny specyficzne dla badanych gatunków kręgowców

Na podstawie bazy danych GenTree (Shao i in., 2019) zidentyfikowano 5 retrogenów, które należą do grupy retrogenów specyficznych dla naszego gatunku. Należą do nich: *retro_hsap_852*, *retro_hsap_1605*, *retro_hsap_3179*, *HMGB1P6* oraz *RPL21P16*. Dane zweryfikowano także w bazie RetrogeneDB2. W tym przypadku analizy potwierdziły, że retrokopie te nie posiadają ortologów wśród innych gatunków kręgowców (Rosikiewicz i in., 2017).

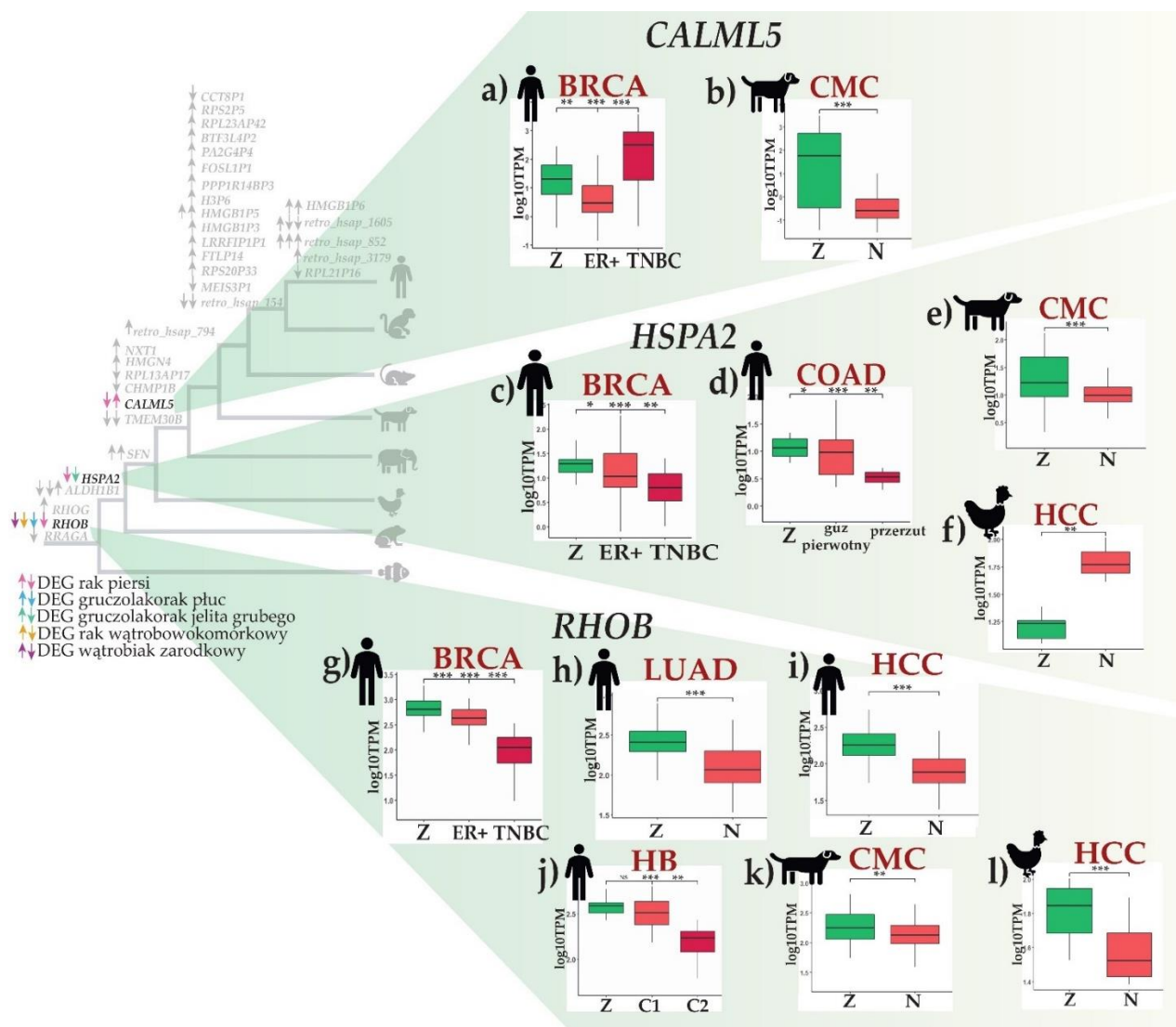
Jednym ze zidentyfikowanych retrogenów, specyficznych dla człowieka, jest *HMGB1P6* (ang. *high mobility group box 1 pseudogene 6*). Analizy wykazały, że ekspresja tego retrogenu jest zwiększona w próbach nowotworowych piersi i jelita grubego. Nie stwierdzono natomiast istotnych zmian w ekspresji *HMGB1P6* pomiędzy poszczególnymi typami nowotworu piersi czy stadiami nowotworu jelita grubego. Druga retrokopia – *retro_hsap_1605* – wykazała obniżony poziom ekspresji w nowotworze piersi i nowotworze płuc. W próbach raka wątrobowokomórkowego *retro_hsap_1605* wykazywał zwiększoną ekspresję. Natomiast retrogen *retro_hsap_852* wykazał jednolity kierunek zmian – wzrost ekspresji w trzech ludzkich nowotworach.

Jako retrogeny o zwiększonej ekspresji w próbach nowotworu gruczołu mlekowego psa zidentyfikowano *retro_cfam_320*, *retro_cfam_1061* i *retro_cfam_1036*. Retrokopie te należą do retrokopii specyficznych dla tego gatunku. Zgodnie z bazą RetrogeneDB2 nie posiadają ortologów w genomach innych kręgowców (Rosikiewicz i in., 2017). W przypadku analizy retrokopii kury zidentyfikowano jeden retrogen specyficznie występujący w genomie tego gatunku. Retrokopia *retro_ggal_15* należy do retrogenów specyficznych dla kury – nie posiada ortologów w genomach innych kręgowców (Rosikiewicz i in., 2017).

4.3. Analiza eksperymentalna retrogenów o zmienionej ekspresji w nowotworach

4.3.1. Ekspresja w nowotworach i zdrowych tkankach/narządach *RHOB*, *HSPA2* i *CALML5*

W celu zweryfikowania wyników analiz bioinformatycznych, ekspresję trzech retrogenów wspólnych dla nowotworów odległych ewolucyjnie gatunków kręgowców sprawdzono eksperymentalnie w ludzkich liniach komórkowych MCF-7 (linia komórkowa ludzkiego estrogenozależnego gruczolakoraka piersi), MCF-12F (linia komórkowa ludzkiego gruczołu mlekowego), NCI-H1703 (linia komórkowa niedrobnokomórkowego raka płuc) i MRC-5 (ludzkie fibroblasty płuca). Ekspresję retrokopii sprawdzono także w panelu ludzkiego cDNA pochodzącego z 8 narządów, panelu cDNA pochodzącego z 6 narządów psa i cDNA pochodzącego z 6 narządów kury oraz w świeżo mrożonych próbach nowotworowych i tkankach okalających pobranych od psów i szczurów. Pary prób do analiz pozyskano z kliniki weterynaryjnej „Klinika Małych Zwierząt Wąsiatycz” w Poznaniu. Pozyskano 2 pary prób gruczołu mlekowego szczura, parę prób guza z jamy brzusznej psa, parę prób guza śledziony psa, parę prób guza z obszaru zatoki okołoodbytniczej psa oraz 2 pary prób gruczołu mlekowego szczura. W związku z tym analizy rozszerzono o wyżej wymienione nowotwory. Nie udało się uzyskać materiału weterynaryjnego pochodzącego od kury. Podjęto także próbę potwierdzenia ekspresji retrogenów w próbach pochodzących z guzów nowotworowych oraz tkanek okalających kota i chomika. Ze względu na zbyt duże podobieństwo sekwencji retrokopii i ich genów rodzicielskich w przypadku tych dwóch gatunków nie udało się zaprojektować specyficznych dla retrokopii starterów.



Rycina 13. Wykresy pudełkowe dla trzech ewolucyjnie starych retrogenów, wspólnych dla nowotworów człowieka, psa i kury – *CALML5*, *HSPA2* oraz *RHOB*.

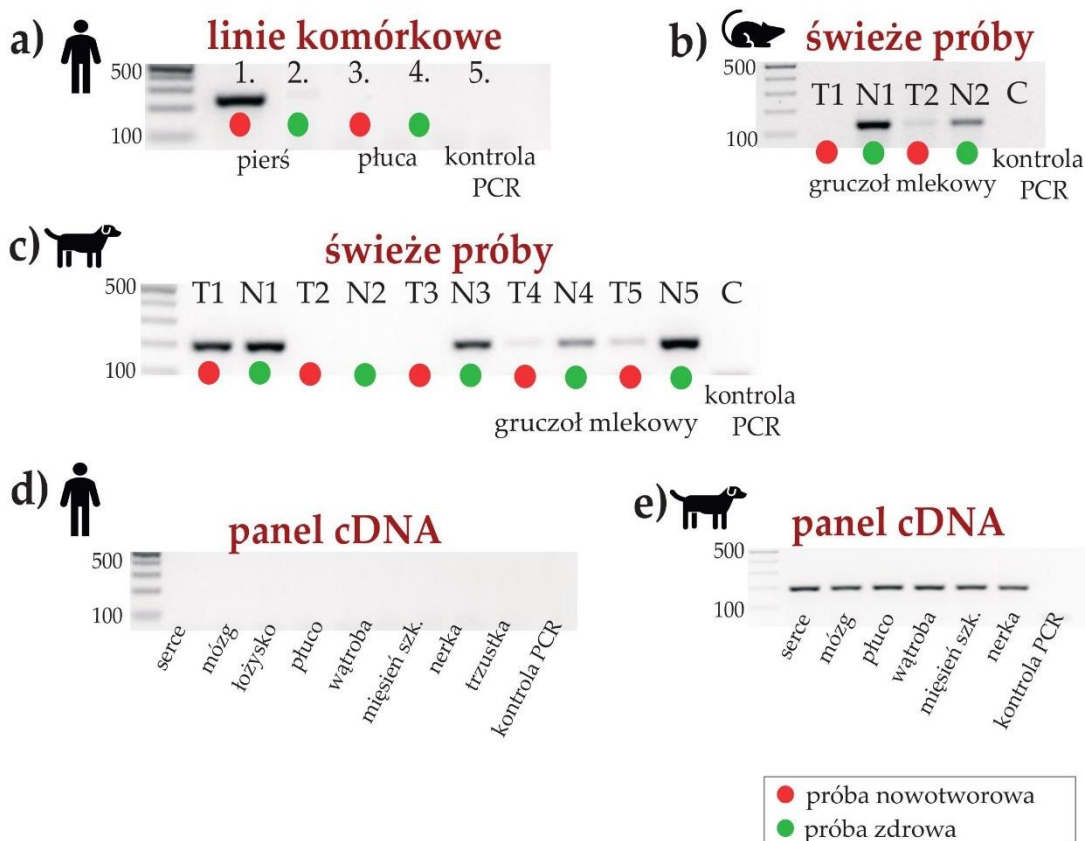
Z-próby zdrowe; N-próby nowotworowe; BRCA- rak piersi; CMC- rak gruczołu mlekowego psa; COAD-gruczołakorak jelita grubego; HCC- rak wątrobowokomórkowy; LUAD-gruczołakorak płuc; HB- wątrobiak zarodkowy.

*** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$, ns- $p > 0,05$

Pierwszy z analizowanych retrogenów – *CALML5* – zidentyfikowano jako retrogen o zmienionym poziomie ekspresji w ludzkim nowotworze piersi (Rycina 13a). Co ciekawe, ekspresja *CALML5* maleje w łagodniejszym nowotworze typu ER+, natomiast wzrasta w TNBC – bardziej agresywnym typie nowotworu piersi. Potwierdzono ekspresję *CALML5* w linii komórkowej nowotworu piersi MCF-7, w pozostałych liniach komórkowych nie wykazano ekspresji tego retrogenu (Rycina 14a). Nie potwierdza to analiz bioinformatycznych, w tym przypadku konieczna byłaby analiza tkanki guzów pobranych od pacjentek z różnymi typami nowotworu piersi – ER+ oraz TNBC. W danych transkryptomicznych nowotworu gruczołu

mlekowego psa *CALML5* został zidentyfikowany jako retrogen o zmniejszonym poziomie ekspresji (Rycina 13b). Potwierdzono to w świeżo mrożonych próbach, pochodzących od szczura i psa – widoczny był spadek poziomu ekspresji *CALML5* w próbach nowotworowych (Rycina 14bc, *szczur*, *T1*, *N1*, *T2*, *N2*, *pies*, *T3*, *N3*, *T4*, *N4*, *T5*, *N5*). W przypadku prób psa, *CALML5* nie wykazywał ekspresji w nowotworze śledziony oraz tkance okalającej (Rycina 14c, *T2*, *N2*). Dodatkowo sprawdzono ekspresję w próbach z nowotworu zlokalizowanego w jamie brzusznej psa oraz tkanki okalającej, gdzie *CALML5* wykazał ekspresję w obydwóch próbach (Rycina 14c, *T1*, *N1*). Nowotwór jamy brzusznej psa to mało precyzyjne określenie, niemniej jednak tylko taki opis otrzymano z Kliniki Weterynaryjnej. Ekspresję retrogenu sprawdzono w tkankowych panelach cDNA (Rycina 14de). *CALML5* jest specyficzny dla ssaków, nie występuje w genomie kury, dlatego jego ekspresję sprawdzano tylko w przypadku człowieka i psa. *CALML5* nie ulegał ekspresji w którejkolwiek z testowanych zdrowych narządów/tkanek człowieka. *CALML5* u człowieka należy do genów wysoce tkankowo-specyficznych i jak wskazują dane z projektu GTEx (ang. *The Genotype-Tissue Expression*) ulega ekspresji głównie w komórkach skóry (GTEx Consortium, 2013). Co ciekawe, w przeciwieństwie do człowieka, *CALML5* wykazuje szeroką ekspresję w narządach psa. Może to świadczyć o roli białka *CALML5* w podstawowych procesach fizjologicznych u psa. Z kolei ekspresja *CALML5* ograniczona do skóry może sugerować ewolucyjną specjalizację funkcji tego genu w ludzkim organizmie. To zróżnicowanie profilu ekspresji retrogenu, a co za tym idzie najprawdopodobniej także funkcji, znajduje odzwierciedlenie we wcześniej już wspomnianych stosunkowo dużych różnicach w samej sekwencji.

CALML5



Rycina 14. Wyniki analiz eksperymentalnych retrogeny *CALML5*.

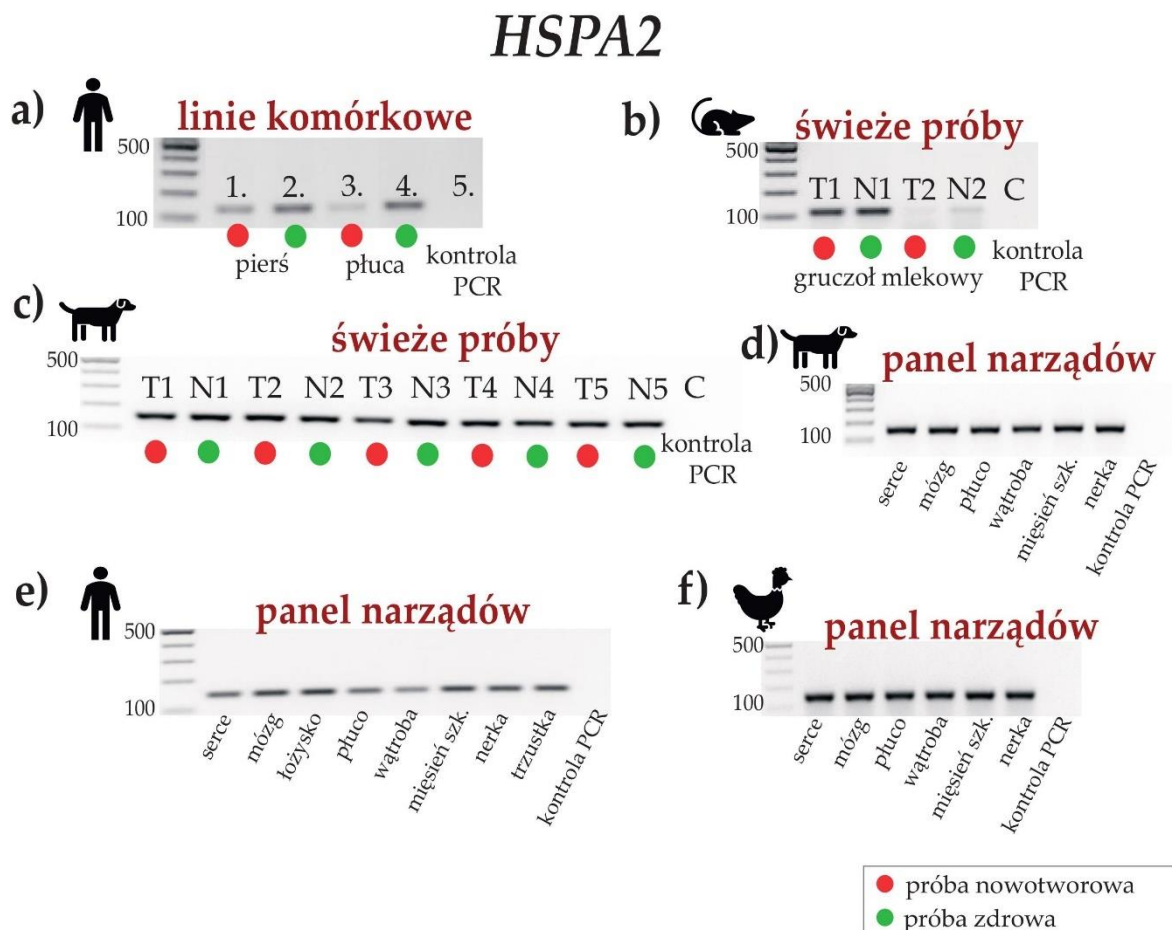
Człowiek: 1.MCF-7 rak piersi; 2. MCF-12F pierś, tkanka zdrowa; 3. NCI-H1703 rak płuc; 4.MRC-5 płuco, tkanka zdrowa; 5. Kontrola PCR.

Szczur: T1, N1, T2, N2 – próbki gruczołu mlekowego;

Pies: T1, N1 - jama brzuszna; T2, N2 - śledziona; T3, N3 - obszar zatoki okołoodbytniczej; T4, N4, T5, N5 - próbki gruczołu mlekowego.

Kolejnym starym ewolucyjnie retrogenem jest *HSPA2*, którego obniżony poziom ekspresji występował w przerzutach gruczolakoraka jelita grubego człowieka oraz w bardziej agresywnym typie raka piersi (Rycina 13cd). Analizy eksperymentalne również wykazały niższą ekspresję *HSPA2* w nowotworowych ludzkich liniach komórkowych w porównaniu z ekspresją w liniach pochodzących ze zdrowych tkanek (Rycina 15a). Z wykorzystaniem analizy PCR potwierdzono ekspresję *HSPA2* w próbach gruczołu mlekowego psa (Rycina 15c). Analizy eksperymentalne nie pokazały zmiany ekspresji tego retrogeny w przypadku prób nowotworowych. Warto zaznaczyć jednak, że rozkład zmienności ekspresji *HSPA2* w przypadku analiz bioinformatycznych psa jest spory w zdrowej tkance i pokrywa się częściowo z zakresem zmienności w nowotworze (Rycina 13e). Ekspresję *HSPA2* wykazano dodatkowo w próbach z guza gruczołu mlekowego oraz tkanki okalającej szczura (Rycina 15b). Ponadto z

wykorzystaniem paneli bibliotek cDNA z różnych narządów wykazano, że *HSPA2* należy do retrogenów o szerokiej ekspresji w tkankach człowieka, psa oraz kury (Rycina 15def).



Rycina 15. Wyniki analiz eksperymentalnych retrogenu *HSPA2*.

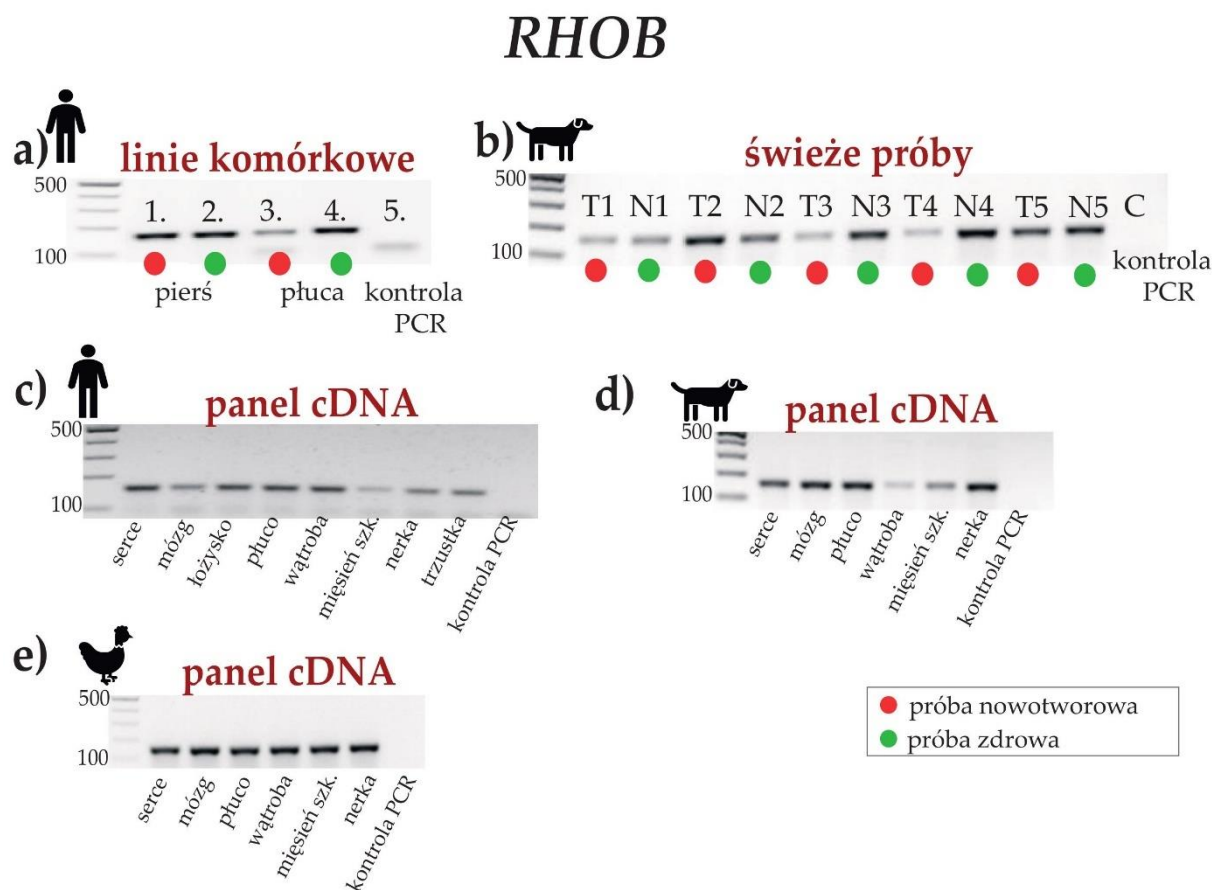
Człowiek: 1.MCF-7 rak piersi; 2. MCF-12F pierś, tkanka zdrowa; 3. NCI-H1703 rak płuca; 4.MRC-5 płuco, tkanka zdrowa; 5. Kontrola PCR.

Szczur: T1, N1, T2, N2 – próbki gruczołu mlekowego;

Pies: T1, N1 - jama brzuszna; T2, N2 - śledziona; T3, N3 - obszar zatoki okołoodbytniczej; T4, N4, T5, N5 - próbki gruczołu mlekowego.

Ostatnim retrogenem, wspólnym dla nowotworów kręgowców jest *RHOB*, który wykazał obniżony poziom ekspresji w raku piersi, gruczolakoraku płuc, raku wątrobowokomórkowym oraz wątrobiaku zarodkowym człowieka (Rycina 13ghij). Spadek ekspresji potwierdzono w linii komórkowej nowotworu płuc (Rycina 16a). PCR nie wskazał widocznej różnicy w ekspresji *RHOB* w liniach komórkowych piersi. Zmniejszoną ekspresję tego retrogenu wykazano także w analizach ekspresji różnicowej retrogenów psa i kury (Rycina 13kl). *RHOB* wykazał wyraźny spadek ekspresji w próbach z nowotworu jamy brzusznej, obszaru zatoki okołoodbytniczej oraz nowotworu gruczołu mlekowego psa (Rycina 16b, T1,

N1, T3, N3, T4, N4). Odmienny kierunek zmiany ekspresji zaobserwowano w próbach pochodzących ze śledziony psa (Rycina 16, *T2, N2*). *RHOB* należy do retrogenów charakteryzujących się ekspresją w szerokim zakresie tkanek człowieka, psa oraz kury (Rycina 16cde).



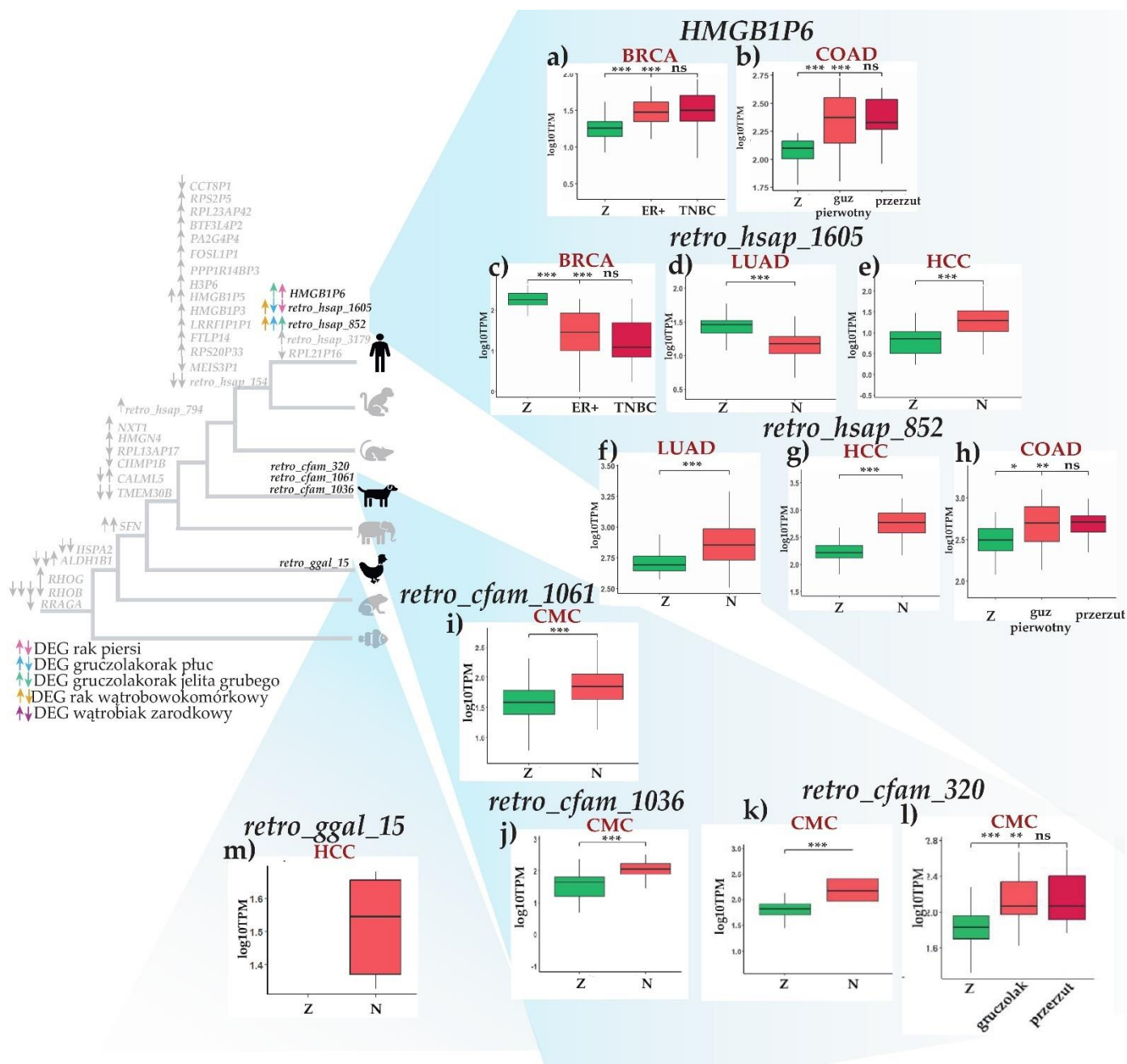
Rycina 16. Wyniki analiz eksperymentalnych retrogenu *RHOB*.

Człowiek: 1.MCF-7 rak piersi; 2. MCF-12F pierś, tkanka zdrowa; 3. NCI-H1703 rak płuc; 4.MRC-5 płuco, tkanka zdrowa; 5. Kontrola PCR.

Pies: T1, N1 - jama brzuszna; T2, N2 - śledziona; T3, N3 - obszar zatoki okołoodbytniczej; T4, N4, T5, N5 - próbki gruczołu mlekowego.

4.3.2. Ekspresja retrogenów specyficznych gatunkowo w nowotworach oraz zdrowych tkankach i narządach

Spośród 5 retrogenów specyficznych dla człowieka wybrano trzy *HMGB1P6*, *retro_hsap_1605* i *retro_hsap_852*, które wykazywały zmianę ekspresji w co najmniej dwóch nowotworach, a następnie sprawdzono ich ekspresję eksperymentalnie. Co ciekawe, żaden ze specyficznych dla człowieka retrogenów nie wykazywał zmienionej ekspresji w przypadku wątrobiaka zarodkowego (*hepatoblastoma*).



Rycina 17. Retrogeny specyficzne dla badanych gatunków kręgowców.

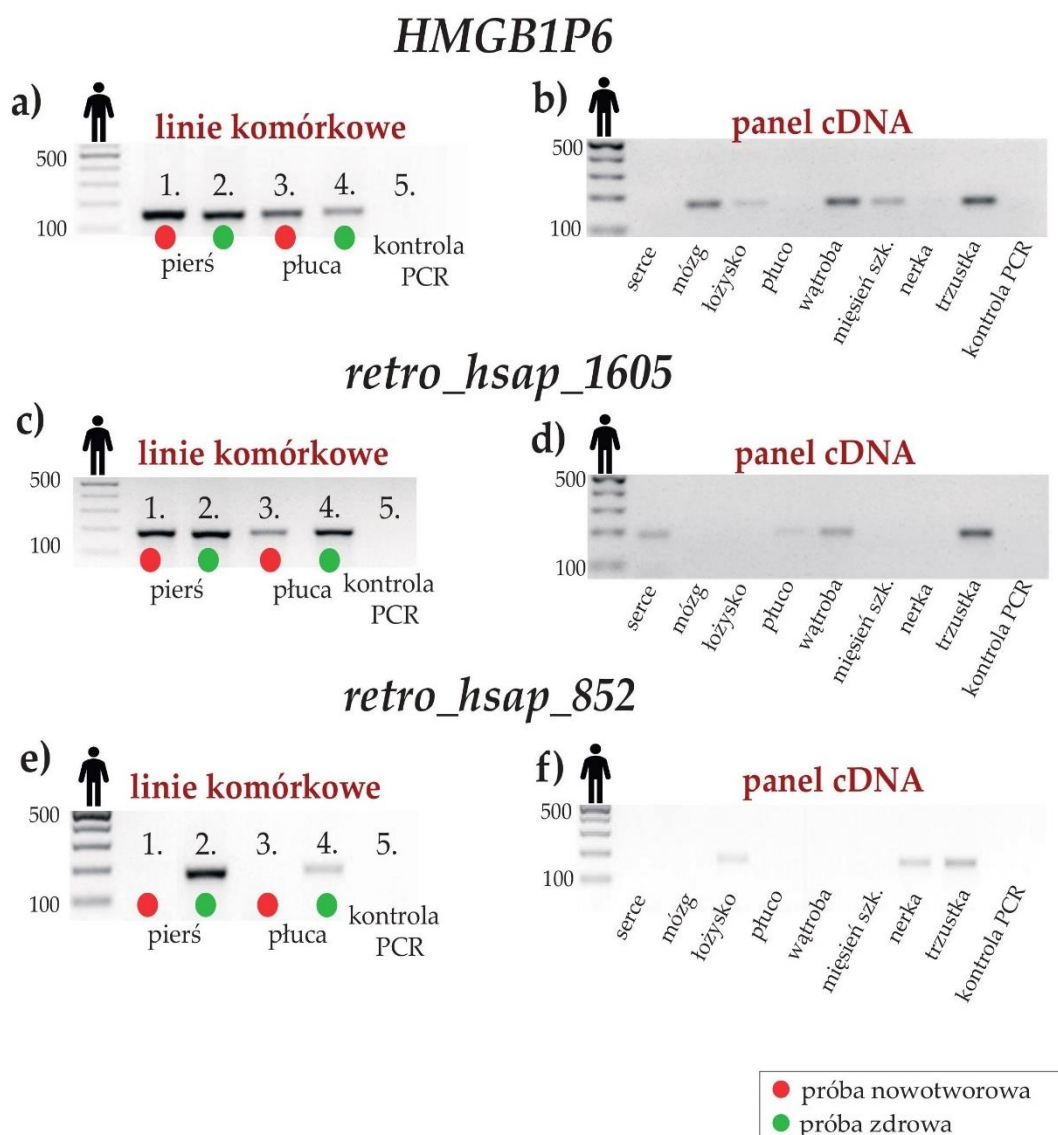
Z-próby zdrowe; N-próby nowotworowe; BRCA- rak piersi; CMC- rak gruczołu mlekowego psa; COAD-gruczolakorak jelita grubego; HCC- rak wątrobowokomórkowy; LUAD-gruczolakorak płuc; HB- wątrobiak zarodkowy.

***p<0,001, **p<0,01, *p<0,05, ns-p>0,05

HMGB1P6 wykazał podwyższony poziom ekspresji w raku piersi (Rycina 17a) i gruczolakoraku jelita grubego (Rycina 17b). Podobnie jak w przypadku innych badanych retrogenów człowieka, analizy eksperymentalne przeprowadzano w liniach komórkowych

pochodzących z nowotworów piersi i płuca oraz ich odpowiednich liniach nienowotworowych. *HMGBIP6* wykazywał ekspresję we wszystkich czterech liniach komórkowych (Rycina 18a). Analiza ta nie potwierdziła jednak zwiększonej ekspresji w nowotworze piersi. Pozwoliła natomiast na wykazanie wzrostu ekspresji *HMGBIP6* w linii nowotworu płuc na co nie wskazywały analizy bioinformatyczne. Drugi retrogen specyficzny dla naszego gatunku, *retro_hsap_1605*, również wykazał ekspresję we wszystkich ludzkich liniach komórkowych z widocznie niższym poziomem ekspresji w nowotworze płuc (Rycina 18c). Jest to zgodne z wynikami analiz bioinformatycznych, gdzie poziom ekspresji *retro_hsap_1605* był niższy w gruczolakoraku płuc (Rycina 17d). Nie udało się potwierdzić zmienionej ekspresji tego retrogenu w nowotworze piersi. Retrogen *retro_hsap_852* w wynikach analiz bioinformatycznych wykazał jednolity kierunek zmian – wzrost ekspresji w trzech ludzkich nowotworach: LUAD (Rycina 17f), HCC (Rycina 17g) i COAD (Rycina 17h). PCR nie potwierdził wyników analiz bioinformatycznych – *retro_hsap_852* wykazywał ekspresję tylko w liniach komórkowych pochodzących ze zdrowych tkanek (Rycina 18e).

Analizy z wykorzystaniem panelu bibliotek cDNA pozwoliły scharakteryzować profil ekspresji tychże retrogenów w narządach człowieka. Wszystkie trzy badane retrogeny, *HMGBIP6*, *retro_hsap_1605* oraz *retro_hsap_852* ulegają ekspresji w trzustce człowieka. Ponadto, *HMGBIP6* wykazał ekspresję w mózgu, łożysku, wątrobie oraz mięśniu szkieletowym (Rycina 18b). Poza wspomnianą trzustką, retrogen *retro_hsap_1605* ulegał ekspresji w sercu, płucu oraz wątrobie (Rycina 18d) a *retro_hsap_852* w łożysku oraz nerce człowieka (Rycina 18f). Co ciekawe, analiza ekspresji z wykorzystaniem komercyjnych bibliotek cDNA nie potwierdziła ekspresji *retro_hsap_852* w płucach na co wskazywały wyniki PCR w linii komórkowej pochodzącej ze zdrowej tkanki płucnej (Rycina 18e).

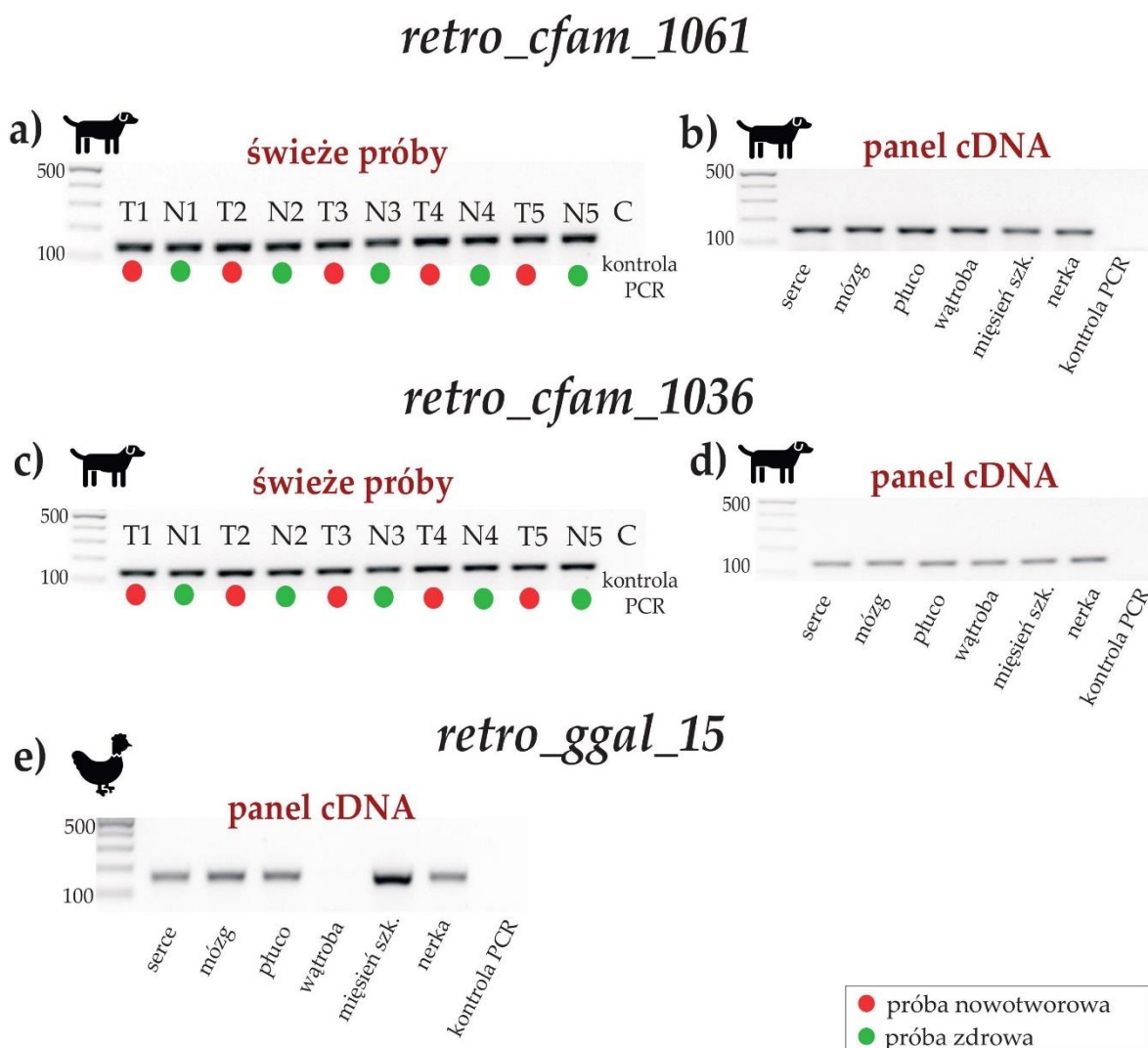


Rycina 18. Wyniki analiz eksperymentalnych retrogenów specyficznych dla człowieka, zidentyfikowanych jako retrogeny o zmienionej ekspresji w więcej niż jednym nowotworze człowieka.

1.MCF-7 rak piersi; 2. MCF-12F piersć, tkanka zdrowa; 3. NCI-H1703 rak płuc; 4.MRC-5 płuco, tkanka zdrowa; 5. Kontrola PCR.

Z przeprowadzonych analiz bioinformatycznych wynika, że retrogeny *retro_cfam_1061*, *retro_cfam_1036* oraz *retro_cfam_320*, które są specyficzne dla psa, cechują się zwiększoną ekspresją w nowotworze gruczołu mlekowego (Rycina 17ijkl). Ekspresję tych retrogenów poddano weryfikacji w świeżo mrożonych próbach pochodzących z guzów nowotworowych i tkanek okalających pobranych od psów oraz w panelu cDNA pochodzącego z 6 narządów psa (Rycina 19). Analizę przeprowadzono jedynie dla dwóch retrokopii gdyż nie udało się zaprojektować starterów specyficznych dla *retro_cfam_320*. Wykazano ekspresję retrogenów *retro_cfam_1061* i *retro_cfam_1036* we wszystkich próbach pochodzących z

guzów i tkanek okalających psa (odpowiednio Rycina 19a i Rycina 19c) bez wyraźnej różnicy w poziomach ekspresji pomiędzy próbkami z tkanki nowotworowej a próbkami z tkanki okalającej. Analizy z wykorzystaniem panelu cDNA wykazały ekspresję retrogenów *retro_cfam_1061* i *retro_cfam_1036* we wszystkich badanych narządach psa (odpowiednio Rycina 19b i 19d).



Rycina 19. Wyniki analiz eksperymentalnych retrogenów specyficznych dla psa lub kury, które wykazywały zmieniony poziom ekspresji w nowotworach.

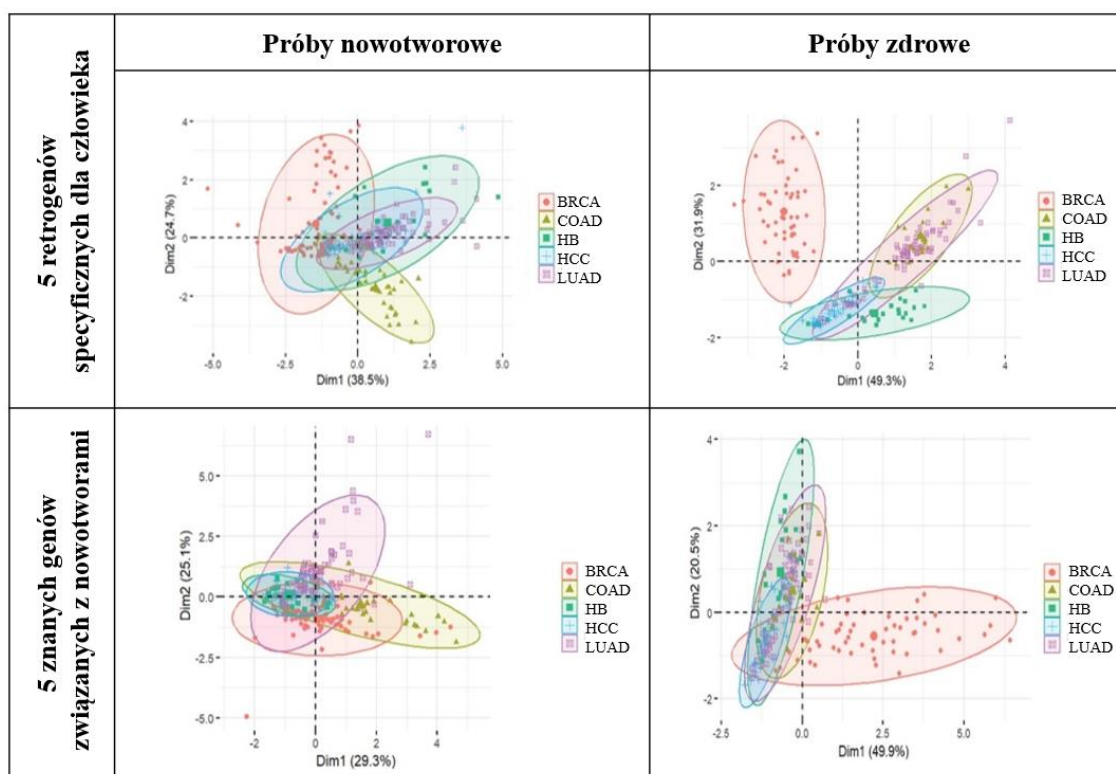
Pies: T1, N1 - jama brzuszna; T2, N2 - śledziona; T3, N3 - obszar zatoki okołoodbytniczej; T4, N4, T5, N5 - próbki gruczołu mlekowego.

Do retrogenów specyficznych dla kury należy *retro_ggal_15*, który wykazał ekspresję jedynie w próbach nowotworowych (Rycina 17m). Ze względu na brak możliwości sprawdzenia ekspresji w tkance nowotworowej lub pochodzącej od takiej tkanki linii komórkowej, ekspresję *retro_ggal_15* zbadano eksperymentalnie jedynie w panelu cDNA

pochodzącego z 6 narządów kury (Rycina 19e). Retrogen wykazał ekspresję we wszystkich tkankach z wyjątkiem wątroby, co jest zgodne z analizą bioinformatyczną, gdzie w próbach pochodzących ze zdrowej wątroby nie wykazano ekspresji tej retrokopii.

4.4. Profil ekspresji retrogenów a różnicowanie typów nowotworów

Poszczególne typy nowotworów różnią się znacząco pod względem molekularnym, histopatologicznym oraz klinicznym. Zrozumienie zróżnicowania ekspresji genów w różnych typach nowotworów może mieć kluczowe znaczenie dla precyzyjnej diagnostyki i doboru terapii onkologicznej. W przypadku niektórych z retrokopii widoczny jest spójny wzorzec zmiany ekspresji w różnych typach badanych nowotworów. Przykładem jest podwyższona ekspresja *retro_hsap_852* w 3 nowotworach czy zmniejszona ekspresja *RHOB* w 4 spośród 5 analizowanych nowotworów człowieka. Niejednokrotnie retrogeny zmieniają swój poziom ekspresji w różnym kierunku, w zależności od rodzaju nowotworu. Dobrym przykładem jest *retro_hsap_1605*, którego zmniejszony poziom ekspresji występował w raku piersi i gruczolakoraku płuc a podwyższony w raku wątrobowokomórkowym. Aby dowiedzieć się więcej na temat różnic pomiędzy analizowanymi typami nowotworów w oparciu o ekspresję retrogenów, przeprowadzono analizę składowych głównych (PCA, z ang. *principal component analysis*). Porównano zdolność różnicowania poszczególnych typów nowotworów lub tkanek zdrowych przez profil ekspresji 5 retrogenów specyficznych dla człowieka z ekspresją 5 najczęściej cytowanych genów związanych z nowotworzeniem (*TP53*, *MYC*, *KRAS*, *EGFR*, *ERBB2*; według CancerMine database). Analizy PCA wykazały, że młodsze ewolucyjnie retrogeny (specyficzne dla człowieka) różnicują 4 spośród badanych nowotworów nieco lepiej niż znane geny nowotworowe (*BRCA*, *COAD*, *HB*, *HCC*) (Rycina 20). Przeciwnie, gruczolakorak płuc (LUAD) jest lepiej różnicowany na podstawie ekspresji 5 genów, powiązanych z nowotworami. Wykresy wskazują, że poziom ekspresji retrogenów pozwala na częściowe oddzielenie prób raka piersi oraz raka jelita grubego. W przypadku znanych genów związanych z nowotworami widoczna jest większa kumulacja prób w centrum wykresu. W przypadku tkanek zdrowych, ekspresja wskazanych retrogenów pozwala na całkowite odróżnienie prób pochodzących z tkanek okalających raka piersi oraz częściowo tkanek okalających wątrobiaka zarodkowego.



Rycina 20. Profile ekspresji wybranych genów a zdolność różnicowania rodzajów nowotworów.

BRCA- rak piersi, COAD- gruczolakorak jelita grubego, HB- wątrobiak zarodkowy, HCC- rak wątrobowokomórkowy, LUAD- gruczolakorak płuc.

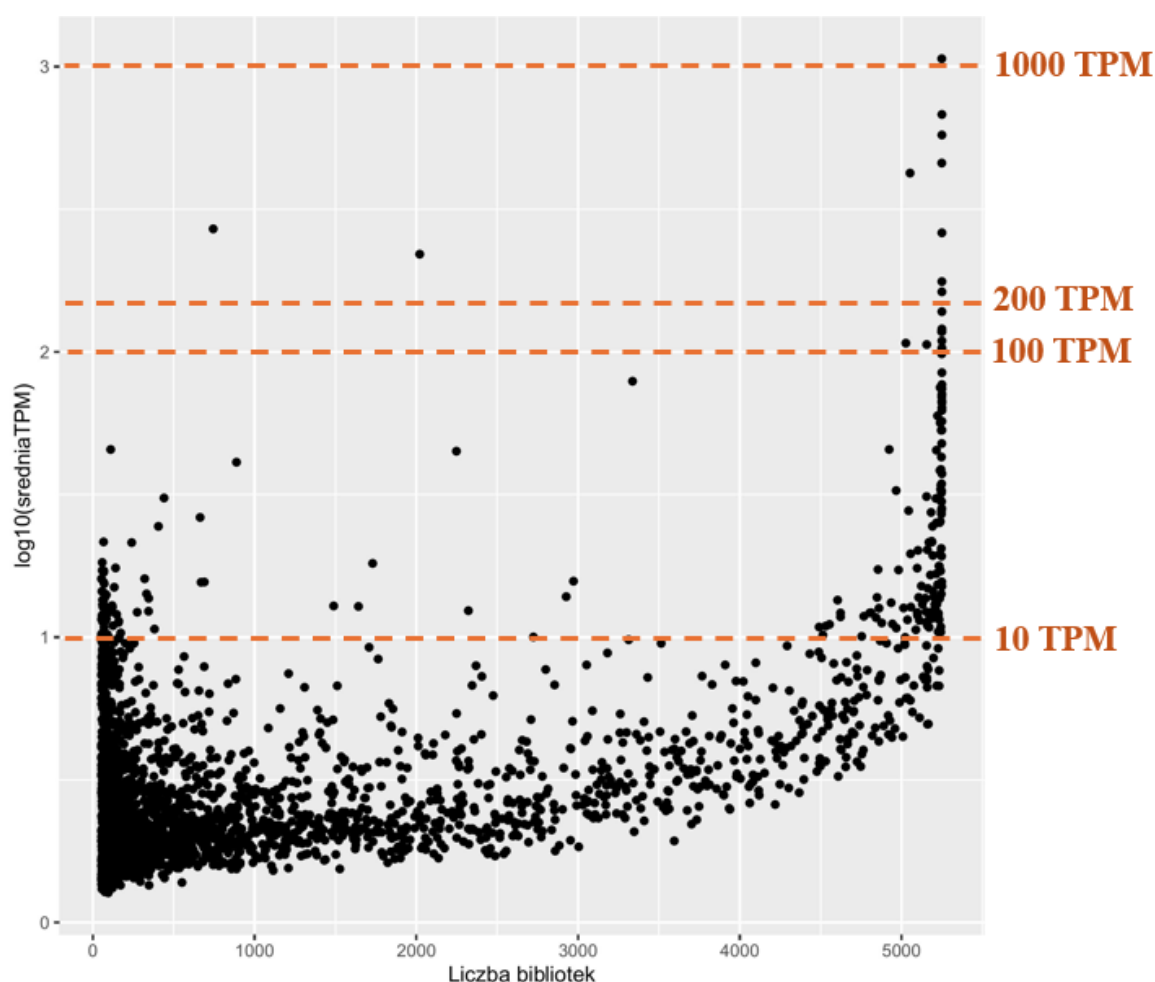
4.5. Wielkoskalowe analizy retrokopii

4.5.1. Ekspresja retrokopii w danych TCGA

W celu zwiększenia skali analiz ekspresji retrokopii w nowotworach człowieka, przeanalizowano dane dotyczące 10 rodzajów nowotworów znajdujące się w repozytorium TCGA (The Cancer Genome Atlas Program (TCGA) - NCI, 2022; Weinstein i in., 2013). Analizy oparto na bezpośrednio pobranych macierzach, które zawierały dane ekspresji 60 660 genów, w tym lncRNA (Tabela 3). 8 zestawów danych zawierało wystarczającą liczbę bibliotek pochodzących z tkanek okalających, by przeprowadzić analizę ekspresji różnicowej. Dane z glejaka wielopostaciowego (TCGA-GBM) oraz raka jajnika (TCGA-OV) z uwagi na małą ilość lub brak prób zdrowych zostały użyte tylko do analizy ogólnej ekspresji retrogenów, nie uwzględniono ich w analizach ekspresji różnicowej na tym etapie.

Ekspresję retrokopii sprawdzono w sumie w 5250 bibliotekach RNA-Seq. W celu wyeliminowania artefaktów wzięto pod uwagę tylko retrokopie, które wykazywały ekspresję minimum 1 TPM w co najmniej 1% bibliotek. W tym przypadku były to 53 biblioteki. Średnia

wartość TPM dla retrokopii uznanych za ulegające ekspresji została wyliczona tylko dla prób, w których ekspresja danego retrogenu wynosiła minimum 1 TPM. Analiza wskazała 2554 retrogenów wykazujących ekspresję oraz 2026 retrokopii niespełniających przyjętych warunków. 681 retrogenów wykazywało ekspresję w co najmniej 1000 bibliotek (Rycina 21). Ponadto, 469 retrokopii wykazywało ekspresję w ponad połowie prób. Do retrogenów o najwyższej ekspresji, średniej przekraczającej 100 TPM (101 – 1065 TPM) w ponad 5000 prób, należało 17 retrogenów: *RHOG* (ENSG00000177105), *RPS28P4* (ENSG00000237039), *retro_hsap_852* (ENSG00000244398), *RPL17P6* (ENSG00000226084), *RPL13AP5* (ENSG00000236552), *RPL13P12* (ENSG00000215030), *RPL18AP3* (ENSG00000213442), *HMGH4* (ENSG00000182952), *ARF6* (ENSG00000165527), *PPIAP22* (ENSG00000198618), *RRAGA* (ENSG00000155876), *RPS2P5* (ENSG00000240342), *SFN* (ENSG00000175793), *retro_hsap_1605* (ENSG00000237550), *RHOB* (ENSG00000143878), *RPL36AL* (ENSG00000165502), *RPS28P7* (ENSG00000227097).



Rycina 21. Ekspresja retrokopii w analizowanych 5250 próbach TCGA. Liczba bibliotek, w których retrokopia wykazywała ekspresję (minimum 1 TPM) (oś x) oraz średnia wartość ekspresji (oś y).

4.5.2. Retrogeny o zróżnicowanej ekspresji w nowotworach człowieka – dane TCGA

8 zestawów danych nowotworów człowieka zostało poddanych analizie ekspresji różnicowej, gdzie porównywano próby nowotworowe i próby zdrowe (BRCA, COAD, KIRC, LIHC, LUSC, PRAD, READ, UCEC). Wyniki analiz danych TCGA pozwoliły na identyfikację 119 retrogenów o zmienionym poziomie ekspresji w nowotworach człowieka (

Tabela 13). 20 spośród nich to retrogeny, które wskazano także w poprzedniej analizie danych NCBI SRA (Tabela 8). Wszystkie trzy wspomniane wcześniej jako „stare” ewolucyjnie retrogeny (*CALML5*, *HSPA2*, *RHOB*) zostały zidentyfikowane jako retrogeny o zmienionej ekspresji w nowotworach, także w przypadku wyników analiz TCGA. Wyniki analizy retrogenu *CALML5* potwierdziły kierunek zmiany ekspresji z poprzedniej analizy nowotworu typu TNBC – u pacjentów z rakiem piersi wykazał wzrost ekspresji (

Tabela 13). Zwiększony poziom ekspresji tego retrogenu wykazano także w raku płuc oraz raku endometrium trzonu macicy. Retrogen *HSPA2* wykazał spadek ekspresji w nowotworze jelita grubego, analogicznie do wcześniejszych wyników. W przypadku raka piersi jego ekspresja wzrosła, w przeciwieństwie do poprzedniej analizy. Ponadto poziom ekspresji *HSPA2* spadł w raku jasnokomórkowym nerki oraz gruczolakoraku odbytnicy. Z kolei, *RHOB* podobnie do wcześniejszych wyników, wykazał spadek ekspresji w nowotworze wątroby oraz nowotworze płuc. Ekspresja tego retrogenu spadła także w trzech innych nowotworach: gruczolakoraku prostaty, gruczolakoraku odbytnicy oraz raku endometrium trzonu macicy. Szczegółowe informacje dotyczące wszystkich retrogenów o zmienionej ekspresji (log2foldchange, wartości p oraz zmiany w wartościach TPM) znajdują się w tabelach na końcu niniejszej pracy (Aneks, Tabela 32 – Tabela 39).

Retrogen *retro_hsap_852* – specyficzny dla naszego gatunku – wykazał w nowotworze jelita grubego, nowotworze wątroby oraz płuc podobne zmiany ekspresji jak w poprzednich analizach (

Tabela 13). Jego ekspresja wzrosła także w raku jasnokomórkowym nerki, gruczolakeraku prostaty oraz gruczolakeraku odbytnicy. Drugi specyficzny dla człowieka retrogen *retro_hsap_1605* wykazał istotny statystycznie wzrost ekspresji w raku jasnokomórkowym nerki oraz spadek w raku endometrium trzonu macicy. *HMGB1P6* nie został zidentyfikowany jako retrogen o zmienionym poziomie ekspresji w wynikach analiz danych TCGA.

Tabela 13. Retrogeny o zmienionym poziomie ekspresji w 8 ludzkich nowotworach (dane TCGA). Czerwoną strzałką oznaczono zwiększony poziom ekspresji a zieloną zmniejszony poziom ekspresji. BRCA- rak piersi, COAD- gruczolakorak jelita grubego, KIRC- rak jasnokomórkowy nerki, LIHC- rak wątrobowokomórkowy, LUSC- rak płaskonabłonkowy płuc, PRAD- gruczolakorak prostaty, READ- gruczolakorak odbytnicy, UCEC- rak endometrium trzonu macicy.

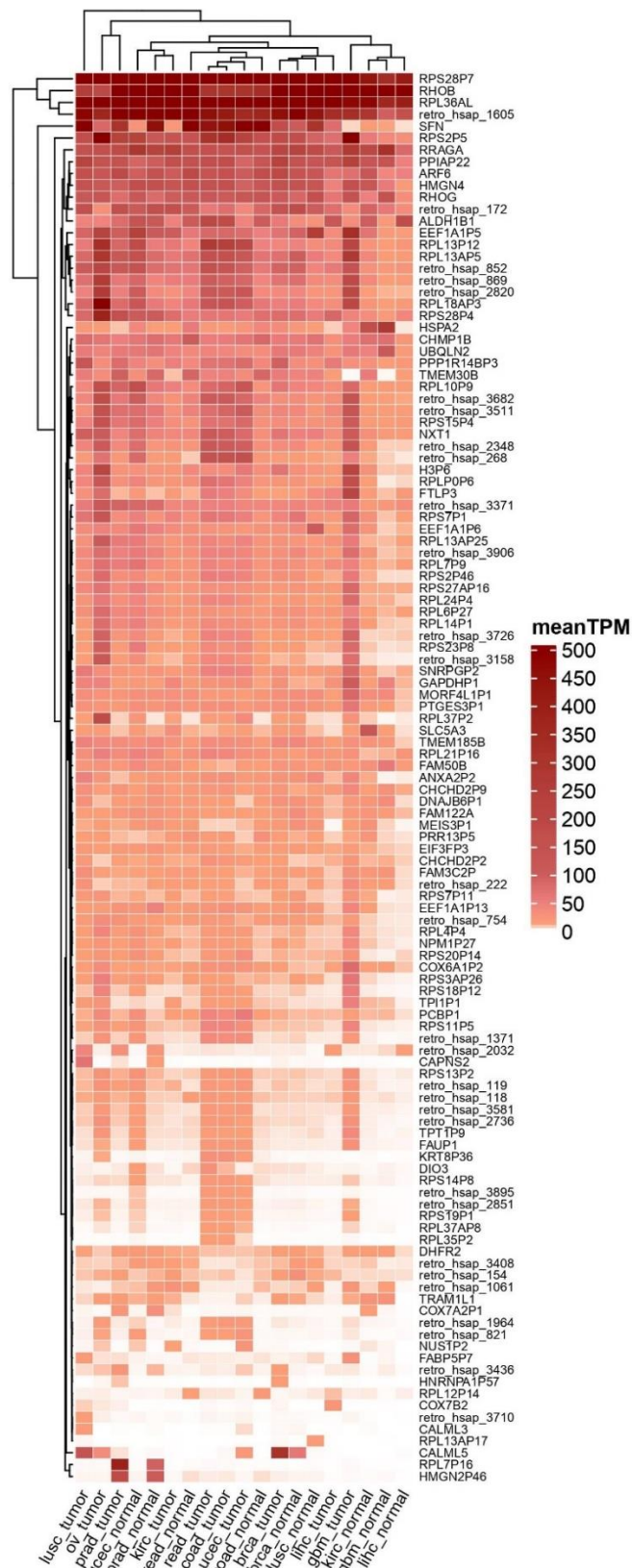
Retrogen	Gen rodzicielski	BRCA	COAD	KIRC	LIHC	LUSC	PRAD	READ	UCEC
ALDH1B1	ALDH2	↑	NA	↓	↓	↑	↓	NA	↓
ANXA2P2	ANXA2	NA	NA	NA	NA	NA	↓	NA	NA
ARF6	ARF3	NA	NA	NA	NA	NA	NA	↓	NA
CALML3	CALM3	NA	NA	NA	NA	↑	NA	NA	NA
CALML5	CALM3	↑	NA	NA	NA	↑	NA	NA	↑
CAPNS2	CAPNS1	NA	NA	NA	NA	↑	NA	NA	NA
CHCHD2P2	CHCHD2	NA	↑	NA	NA	↑	NA	↑	NA
CHCHD2P9	CHCHD2	↑	NA	NA	NA	↑	NA	NA	NA
CHMP1B	CHMP1A	NA	↓	NA	NA	NA	↓	↓	NA
COX6A1P2	COX6A1	NA	NA	NA	NA	NA	NA	NA	↑
COX7A2P1	COX7A2	NA	NA	↓	NA	NA	NA	NA	NA
COX7B2	COX7B	NA	NA	NA	↑	NA	NA	NA	NA
DHFR2	DHFR	NA	NA	NA	NA	NA	NA	NA	↓
DIO3	DIO1	NA	NA	NA	NA	NA	NA	NA	↓
DNAJB6P1	DNAJB6	NA	NA	NA	NA	↑	NA	NA	NA
EEF1A1P13	EEF1A1	NA	↑	NA	NA	NA	NA	NA	NA
EEF1A1P5	EEF1A1	NA	NA	NA	NA	↓	NA	NA	↓
EEF1A1P6	EEF1A1	NA	NA	NA	NA	↓	NA	NA	↓
EIF3FP3	EIF3F	NA	NA	NA	NA	NA	NA	NA	↓
FABP5P7	FABP5	NA	NA	NA	NA	↑	NA	NA	NA
FAM122A	FAM122C	↓	NA	NA	NA	NA	NA	NA	↓
FAM3C2P	FAM3C	NA	NA	NA	NA	↑	NA	NA	NA
FAM50B	FAM50A	NA	NA	NA	NA	NA	NA	NA	↓
FAUP1	FAU	NA	↑	NA	NA	NA	NA	↑	NA
FTLP3	FTL	NA	↑	NA	NA	NA	NA	↑	↑
GAPDHP1	GAPDH	↑	↑	↑	↑	↑	NA	NA	↑
H3P6	H3-3B	↑	NA	NA	NA	↑	NA	NA	↑
HMG2N2P46	HMG2	NA	NA	NA	NA	NA	↑	NA	NA
HMG4	HMG2	NA	NA	NA	↑	NA	↓	NA	↓
HNRNPA1P57	HNRNPA1	↑	NA	NA	NA	NA	NA	NA	NA
HSPA2	HSPA8	↑	↓	↓	NA	NA	NA	↓	NA
KRT8P36	KRT8	NA	↑	NA	NA	NA	NA	↑	↑
MEIS3P1	MEIS3	↓	NA	↓	NA	↓	↓	NA	↓
MORF4L1P1	MORF4L1	NA	NA	NA	NA	NA	NA	NA	↓
NPM1P27	NPM1	NA	NA	NA	NA	NA	NA	NA	↓
NUS1P2	NUS1	NA	NA	NA	NA	NA	NA	NA	↑
NXT1	NXT2	NA	↑	NA	NA	↑	NA	↑	NA
PCBP1	PCBP2	NA	↑	NA	NA	NA	NA	↑	↑
PPIAP22	PPIA	↑	NA	NA	↑	↑	NA	NA	NA
PPP1R14BP3	PPP1R14B	↑	NA	↑	↑	↑	↑	NA	↑
PRR13P5	PRR13	NA	↓	NA	NA	NA	NA	NA	NA
PTGES3P1	PTGES3	NA	NA	↑	NA	NA	NA	NA	NA
retro_hsap_1061	PDAP1	NA	NA	↑	NA	↓	↓	↓	NA
retro_hsap_118	RPS16	NA	↑	NA	NA	NA	NA	NA	NA
retro_hsap_119	RPS14	NA	↑	NA	NA	NA	NA	↑	NA

Retrogen	Gen rodzicielski	BRCA	COAD	KIRC	LIHC	LUSC	PRAD	READ	UCEC
retro_hsap_1371	RPS18	NA	↑	NA	NA	NA	NA	↑	↑
retro_hsap_154	TMA7	↓	NA	↑	NA	NA	NA	NA	NA
retro_hsap_1605	RPL9	NA	NA	↑	NA	NA	NA	NA	↓
retro_hsap_172	RPL17	NA	NA	↑	NA	NA	NA	NA	↓
retro_hsap_1964	RPL28	NA	↑	NA	NA	NA	NA	↑	↑
retro_hsap_2032	IPMK	NA	NA	NA	NA	↑	NA	NA	NA
retro_hsap_222	VOPP1	NA	NA	NA	NA	↑	NA	NA	NA
retro_hsap_2348	RPLP1	NA	↑	NA	↑	↑	↑	↑	NA
retro_hsap_268	RPS14	NA	↑	NA	NA	NA	NA	↑	NA
retro_hsap_2736	RPS3	NA	↑	NA	NA	NA	NA	↑	NA
retro_hsap_2820	RPS18	NA	↑	↑	↑	↑	NA	↑	NA
retro_hsap_2851	RPL8	NA	↑	NA	NA	NA	NA	↑	NA
retro_hsap_3158	RPS8	NA	↑	NA	NA	NA	NA	NA	NA
retro_hsap_3371	RPS25	NA	NA	↑	NA	NA	NA	NA	NA
retro_hsap_3408	PTS	NA	NA	NA	NA	NA	NA	NA	↓
retro_hsap_3436	RPS10	NA	NA	NA	NA	NA	↑	NA	NA
retro_hsap_3511	RPL29	NA	↑	NA	NA	NA	NA	↑	NA
retro_hsap_3581	RPL23	NA	↑	NA	NA	NA	NA	↑	NA
retro_hsap_3682	RPL27A	NA	↑	↑	↑	↑	↑	NA	NA
retro_hsap_3710	RPL23	NA	NA	NA	NA	↑	NA	NA	NA
retro_hsap_3726	RPL32	NA	↑	NA	NA	NA	NA	↑	NA
retro_hsap_3895	RPL18	NA	↑	NA	NA	NA	NA	↑	NA
retro_hsap_3906	EEF1G	NA	↑	NA	NA	↑	NA	NA	NA
retro_hsap_754	RPS2	NA	NA	NA	NA	NA	↑	NA	NA
retro_hsap_821	RPLP2	NA	↑	NA	NA	NA	NA	↑	NA
retro_hsap_852	RPL36A	NA	↑	↑	↑	↑	↑	↑	NA
retro_hsap_869	RPS10	NA	↑	NA	NA	↑	NA	NA	NA
RHOB	RHOA	NA	NA	↑	↓	↓	↓	↓	↓
RHOG	RAC1	NA	NA	↑	NA	↓	NA	NA	NA
RPL10P9	RPL10	↑	↑	↑	NA	NA	NA	↑	NA
RPL12P14	RPL12	NA	↓	NA	NA	NA	NA	NA	NA
RPL13AP17	RPL13A	NA	NA	NA	NA	↓	NA	NA	NA
RPL13AP25	RPL13A	NA	NA	↑	NA	NA	↑	NA	NA
RPL13AP5	RPL13A	NA	↑	NA	NA	NA	NA	NA	NA
RPL13P12	RPL13	NA	↑	↑	NA	↑	↑	↑	NA
RPL14P1	RPL14	NA	NA	NA	NA	NA	NA	NA	↓
RPL18AP3	RPL18A	NA	↑	↑	NA	↑	NA	NA	NA
RPL21P16	RPL21	NA	NA	↑	NA	NA	NA	NA	NA
RPL24P4	RPL24	NA	↑	NA	NA	↑	NA	NA	NA
RPL35P2	RPL35	NA	↑	NA	NA	NA	NA	↑	NA
RPL36AL	RPL36A	NA	NA	NA	NA	NA	NA	↓	NA
RPL37AP8	RPL37A	NA	↑	NA	NA	NA	NA	↑	NA
RPL37P2	RPL37	NA	↑	NA	NA	NA	NA	↑	↑
RPL4P4	RPL4	NA	↑	NA	NA	NA	NA	NA	NA
RPL6P27	RPL6	NA	NA	NA	NA	↑	NA	NA	NA

Retrogen	Gen rodzicielski	BRCA	COAD	KIRC	LIHC	LUSC	PRAD	READ	UCEC
RPL7P16	RPL7	NA	NA	NA	NA	NA	↑	NA	NA
RPL7P9	RPL7	NA	NA	NA	↑	↑	↑	NA	NA
RPLP0P6	RPLP0	NA	↑	NA	NA	↑	NA	NA	NA
RPS11P5	RPS11	NA	↑	NA	NA	NA	NA	↑	NA
RPS13P2	RPS13	NA	↑	NA	NA	NA	NA	↑	NA
RPS14P8	RPS14	NA	↑	NA	NA	NA	NA	↑	NA
RPS15P4	RPS15	NA	↑	↑	NA	↑	NA	↑	NA
RPS18P12	RPS18	NA	↑	NA	NA	NA	NA	NA	NA
RPS19P1	RPS19	NA	↑	NA	NA	NA	NA	↑	NA
RPS20P14	RPS20	NA	↑	NA	NA	NA	NA	↑	NA
RPS23P8	RPS23	NA	↑	NA	NA	NA	NA	↑	NA
RPS27AP16	RPS27A	NA	NA	NA	NA	↑	NA	NA	↓
RPS28P4	RPS28	NA	NA	↑	NA	NA	↑	NA	NA
RPS28P7	RPS28	NA	NA	↑	NA	NA	NA	NA	NA
RPS2P46	RPS2	NA	↑	↑	NA	↑	↑	NA	NA
RPS2P5	RPS2	NA	↑	↑	NA	↑	↑	↑	NA
RPS3AP26	RPS3A	NA	NA	NA	NA	NA	NA	NA	↓
RPS7P1	RPS7	NA	↑	↑	NA	↑	NA	NA	NA
RPS7P11	RPS7	NA	NA	NA	NA	↑	NA	NA	NA
RRAGA	RRAGB	NA	NA	NA	NA	NA	NA	NA	↓
SFN	YWHAQ	↑	↓	NA	↑	↑	NA	NA	↑
SLC5A3	SLC5A1	NA	NA	↓	NA	NA	NA	NA	NA
SNRPGP2	SNRPG	↑	↑	NA	NA	↑	NA	↑	↑
TMEM185B	TMEM185A	NA	↑	NA	NA	↑	NA	↑	NA
TMEM30B	TMEM30A	NA	↓	↓	NA	NA	NA	↓	NA
TPI1P1	TPI1	NA	NA	NA	NA	↑	NA	NA	NA
TPT1P9	TPT1	NA	NA	NA	NA	NA	NA	↑	NA
TRAM1L1	TRAM1	NA	NA	↓	NA	NA	NA	NA	NA
UBQLN2	UBQLN1	NA	NA	NA	NA	NA	NA	NA	↓

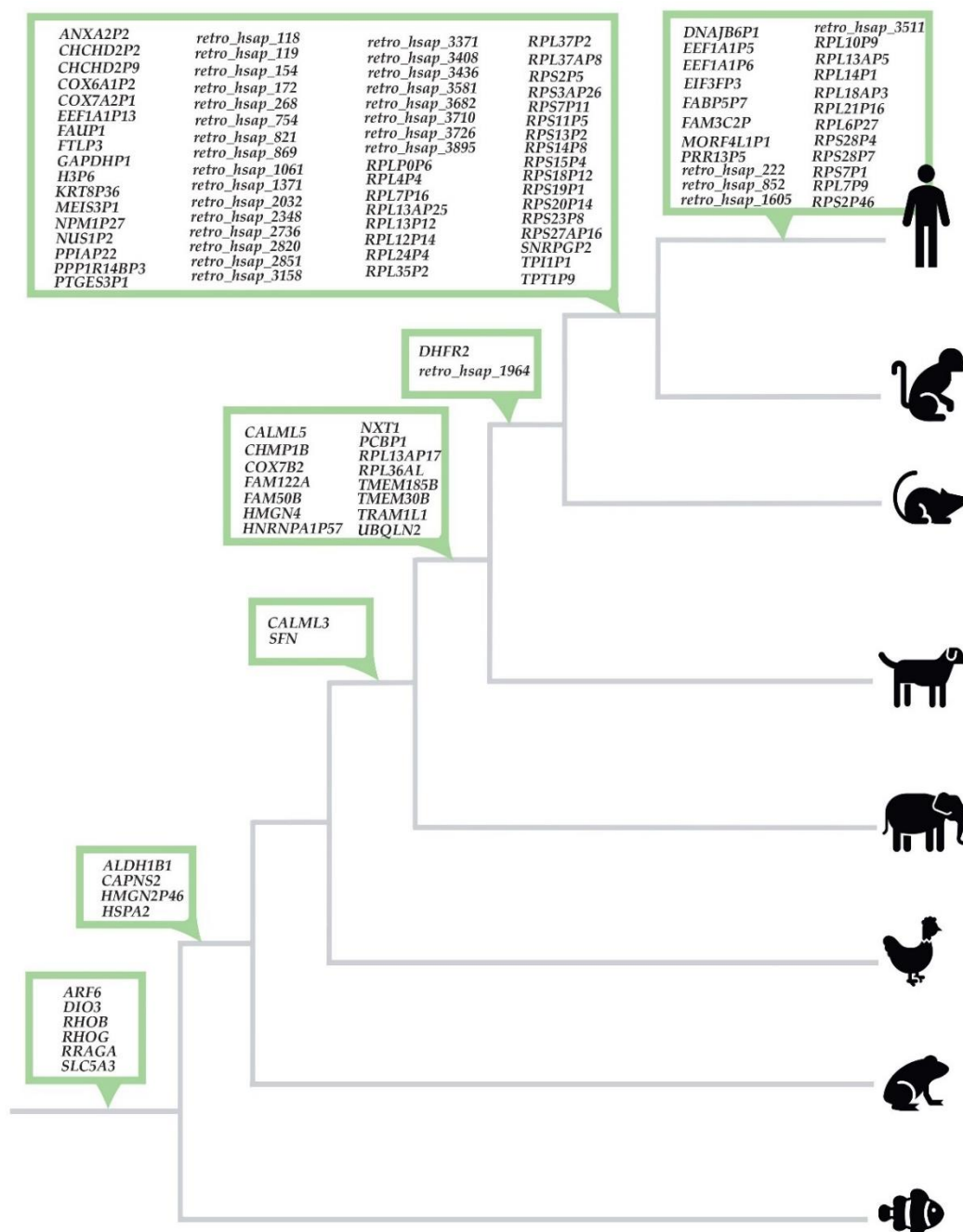
Do retrogenów najczęściej wykazujących zmiany poziomu ekspresji, w ponad połowie analizowanych nowotworów (minimum 5), należą: *ALDH1B1*, *GAPDHP1*, *MEIS3P1*, *PPP1R14BP3*, *retro_hsap_2348*, *retro_hsap_2820*, *retro_hsap_3682*, *retro_hsap_852*, *RHOB*, *RPL13P12*, *RPS2P5*, *SFN* oraz *SNRPGP2*. Na podstawie ekspresji zidentyfikowanych retrogenów przeprowadzono grupowanie hierarchiczne 119 retrogenów o zmienionych w nowotworach profilach ekspresji. Celem było uzyskanie klastrów retrogenów o podobnych profilach ekspresji w analizowanych próbach oraz klastrów nowotworów i zdrowych tkanek podobnych pod względem poziomu ekspresji retrogenów (Rycina 22). W klastrze grupującym retrogeny o największej ekspresji znalazły się: *RPS28P7*, *RHOB*, *RPL36AL* i *retro_hsap_1605*. Analizowane dane nie grupowały się w jakiś określony sposób, ani pod względem swojego pochodzenia ani pod względem tego czy są to próby ze zdrowej czy nowotworowej tkanki. Próby z tkanek nowotworowych grupowały się często z próbami pochodzącymi z innych

zdrowych narządów. Wyjątkiem są grupujące się razem i wyraźnie odróżniające się trzy grupy nowotworowe: gruczolakorak jelita grubego (COAD), gruczolakorak odbytnicy (READ) i rak endometrium trzonu macicy (UCEC).



Rycina 22. Ekspresja retrogenów o zmienionej ekspresji w 8 nowotworach człowieka (dane TCGA). Wiersze przedstawiają retrogeny i przedstawione za pomocą gradientu kolorów poziomy ich ekspresji. Kolor biały oznacza brak ekspresji (0 TPM), a ciemno czerwony najwyższą ekspresję (500+ TPM). Pomiędzy wierszami oraz kolumnami przeprowadzono grupowanie hierarchiczne w celu uzyskania odpowiednio klastrow genów oraz klastrow rodzajów prób.

Podobnie jak w przypadku analizy danych z repozytorium NCBI SRA, określono czas powstania retrokopii o zmienionej ekspresji w nowotworach (Rycina 23). Nie ustalono czasu powstania *retro_hsap_3906* ze względu na brak adnotacji w bazie danych GenTree (Shao i in., 2019). Większość spośród zidentyfikowanych retrokopii (89 spośród 118) związanych z nowotworzeniem u człowieka powstała podczas ewolucji naczelnych i należy do ewolucyjnie młodych genów. Spośród nich 66 jest wspólnych dla naczelnych, a 23 specyficznych dla człowieka.



Rycina 23. Czas powstania retrogenów o zmienionej ekspresji w nowotworach człowieka (dane TCGA, na podstawie GenTree <http://gentree.ioz.ac.cn>, brak *retro_hsap_3906* ze względu na brak adnotacji w bazie danych GenTree).

W ramach przedstawionej pracy sformułowano hipotezę, która zakładała, że retrogeny specyficzne dla naczelnych lub człowieka częściej wykazują powiązania z nowotworami w porównaniu z retrogenami ewolucyjnie starszymi. Analizy bioinformatyczne pozwoliły na zidentyfikowanie odpowiednio 36 i 119 retrogenów wykazujących istotne różnice w poziomie ekspresji w danych nowotworowych z NCBI SRIA i TCGA. Spośród nich 20 retrogenów zostało wspólnie wykrytych w danych z obydwóch repozytoriów. W sumie udało się zidentyfikować 135 retrogenów charakteryzujących się zmienioną ekspresją w ludzkich nowotworach. Za pomocą bazy danych GenTree ustalono czas powstania 132 retrogenów, spośród których grupa 30 retrogenów powstała na różnych etapach ewolucji kręgowców, a w 102 przypadkach retropozycja nastąpiła podczas ewolucji naczelnych. Ponad trzykrotnie więcej retrogenów związanych z nowotworami człowieka należy do stosunkowo nowych genów, co może sugerować, że młodsze retrogeny częściej odgrywają istotną rolę w procesach związanych z transformacją nowotworową u naszego gatunku.

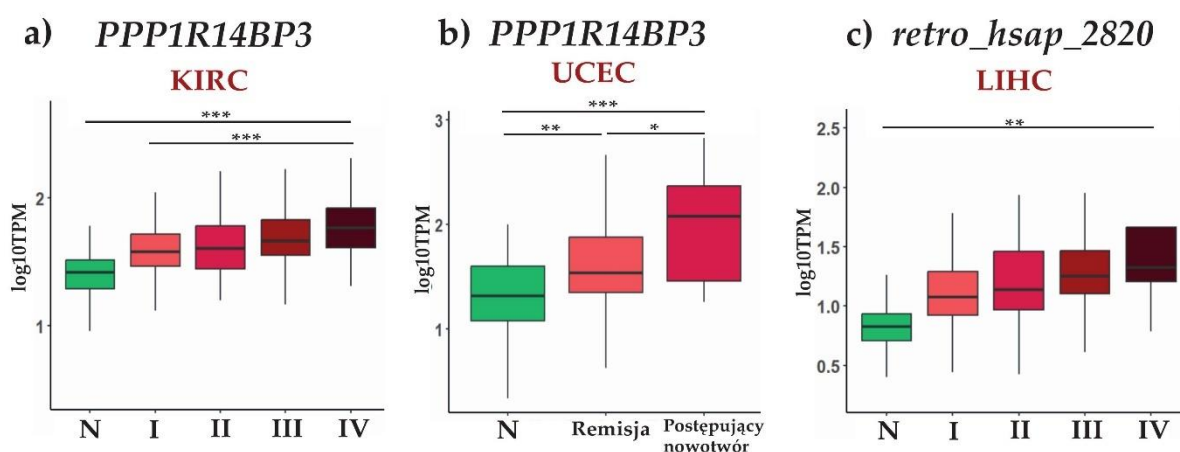
4.5.3. Retrogeny związane z ewolucją komórek w obrębie guza nowotworowego oraz z odpowiedzią na terapię

W celu identyfikacji retrogenów związanych z rozwojem guza nowotworowego przeprowadzono analizy ekspresji różnicowej pomiędzy próbkami nowotworowymi w różnych stadiach nowotworów. Wyniki, które wykazały istotność statystyczną otrzymano tylko w przypadku porównania stadium IV i I w raku jasnokomórkowym nerki (KIRC) oraz raku wątrobowokomórkowym (LIHC) (Tabela 14). Szczegółowe informacje dotyczące retrogenów o zmienionej ekspresji ($\log_2\text{foldchange}$, wartości p oraz zmiany w wartościach TPM) znajdują się w tabeli na końcu niniejszej pracy (Aneks, Tabela 40). Ze względu na zbyt małą liczbę prób, w niektórych nowotworach nie uzyskano założonej istotności statystycznej.

Tabela 14. Retrogeny o zmienionym poziomie ekspresji podczas ewolucji guza nowotworowego (dane TCGA). Czerwoną strzałką oznaczono zwiększony poziom ekspresji a zieloną zmniejszony poziom ekspresji. KIRC- rak jasnokomórkowy nerki, LIHC- rak wątrobowokomórkowy.

Retrogen	Gen rodzicielski	KIRC (stadium IV vs I)	LIHC (stadium IV vs I)
<i>RHOB</i>	<i>RHOA</i>	↓	NA
<i>PPP1R14BP3</i>	<i>PPP1R14B</i>	↑	NA
<i>retro_hsap_2820</i>	<i>RPS18</i>	NA	↑

Zidentyfikowano 3 retrogeny o zmienionym poziomie ekspresji podczas progresji nowotworów (Tabela 14). Pierwszy z nich, *RHOB* wykazał zmniejszoną ekspresję podczas rozwoju raka jasnokomórkowego nerki. *RHOB* to wskazany wcześniej retrogen wykazujący zmiany ekspresji w nowotworach odległych ewolucyjnie gatunków kręgowców. Spadek ekspresji *RHOB* odnotowano także w 9 nowotworach człowieka, w tym w bardziej agresywnym typie raka piersi (TNBC) oraz wątrobiaka zarodkowego (C2). Kolejny retrogen *PPP1R14BP3* (ang. *protein phosphatase 1 regulatory inhibitor subunit 14B pseudogene 3*) wykazał zwiększoną ekspresję wraz z progresją raka jasnokomórkowego nerki (Rycina 24a). W poprzednich analizach również wykazano wzrost ekspresji tego retrogeny w 5 nowotworach (BRCA, LIHC, LUSC, PRAD, UCEC). Do tej pory nie wykazano związku między ekspresją *PPP1R14BP3* a występowaniem nowotworów człowieka. Z kolei, rola jego genu rodzicielskiego w nowotworach została dość dobrze opisana. *PPP1R14B* (ang. *protein phosphatase 1 regulatory inhibitor subunit 14B*) wskazano jako marker diagnostyczny u pacjentek z rakiem endometrium trzonu macicy (He i in., 2023). Ekspresję *PPP1R14B* powiązano z progresją raka piersi TNBC (Liao i in., 2023).



Rycina 24. Wykresy pudełkowe dla wybranych retrogenów związanych z ewolucją w obrębie guza nowotworowego lub odpowiedzią na terapię.

KIRC- rak jasnokomórkowy nerki, LIHC- rak wątrobowokomórkowy, UCEC- rak endometrium trzonu macicy, N – próby zdrowe, I-IV stadia guzów nowotworowych.

*** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$, ns- $p > 0,05$

Retro_hsap_2820 zidentyfikowano jako retrokopię, której poziom ekspresji wzrasta wraz z progresją raka wątrobowokomórkowego (Rycina 24c). We wcześniejszych analizach *retro_hsap_2820* wykazał zwiększony poziom ekspresji w 5 badanych nowotworach (BRCA,

LIHC, LUSC, READ). Dotychczas nie powiązano ekspresji tej retrokopii z nowotworami człowieka. Natomiast wariant rs17215231 (C/T) genu rodzicielskiego *RPS18* (ang. *ribosomal protein S18*) wskazano jako czynnik zwiększający ryzyko występowania raka piersi (Bigge i in., 2024).

W celu identyfikacji retrogenów związanych z efektem działania terapii porównano ich ekspresję pomiędzy próbami, gdzie w wyniku terapii nowotwór postępował a próbami, gdzie terapia skutkowała remisją choroby nowotworowej. Zidentyfikowano trzy retrogeny o zmienionej ekspresji (Tabela 16). Szczegółowe informacje dotyczące retrogenów o zmienionej ekspresji znajdują się w tabeli na końcu niniejszej pracy (Aneks, Tabela 41).

Tabela 15. Retrogeny o zmienionym poziomie ekspresji w odpowiedzi na zastosowaną terapię (dane TCGA). Czerwoną strzałką oznaczono zwiększony poziom ekspresji a zieloną zmniejszony poziom ekspresji. LUSC- rak płaskonabłonkowy płuc, OV- rak jajnika, UCEC- rak endometrium trzonu macicy.

Retrogen	Gen rodzicielski	LUSC (postępujący nowotwór vs remisja)	OV (postępujący nowotwór vs remisja)	UCEC (postępujący nowotwór vs remisja)
RPL10P9	<i>RPL10</i>	↓	↑	NA
PPP1R14BP3	<i>PPP1R14B</i>	NA	NA	↑
RPS28P7	<i>RPS28</i>	NA	NA	↑

Spadek ekspresji retrogenu *RPL10P9* (ang. *ribosomal protein L10 pseudogene 9*) powiązano z postępującym rakiem płaskonabłonkowym płuc po zastosowanej terapii (Tabela 15). Odwrotnie, istotny statystycznie wzrost ekspresji *RPL10P9* zauważono w postępującym raku jajnika. Zwiększony poziom ekspresji tego retrogenu wykazano w 4 analizowanych wcześniej nowotworach (BRCA, COAD, KIRC, READ). Dotychczas nie powiązano *RPL10P9* z występowaniem nowotworów. Natomiast mutacje w obrębie genu rodzicielskiego *RPL10* powiązano z dziecięcą ostrą białaczką limfoblastyczną komórek T (De Keersmaecker i in., 2013).

Zwiększona ekspresja *PPP1R14BP3* została powiązana z postępującym rakiem endometrium trzonu macicy (UCEC) po zastosowanej terapii onkologicznej (Rycina 24b). Jak wspomniano wcześniej *PPP1R14BP3* nie został opisany w literaturze w kontekście nowotworów. Natomiast jego gen rodzicielski powiązano z odpowiedzią na terapię, dokładniej

z rozwojem oporności na paklitaksel u pacjentek z rakiem piersi TNBC (Liao i in., 2023). Drugim retrogenem, którego ekspresja wzrosła wraz z postępującym nowotworem po wdrożonym leczeniu UCEC był *RPS28P7* (ang. *ribosomal protein S28 pseudogene 7*). *RPS28P7* wykazał zwiększony poziom ekspresji w raku jasnokomórkowym nerki (KIRC). Salgado wraz z zespołem zidentyfikowali *RPS28P7* jako pseudogen, którego zwiększona ekspresja w raku trzustki wpływa na złe rokowanie i szybszą śmierć pacjentów (Salgado i in., 2024).

4.5.4. Retrogeny o zróżnicowanej ekspresji pochodzące od genów powiązanych z nowotworami

Spośród 1520 genów rodzicielskich opisanych w RetrogeneDB2 (Rosikiewicz i in., 2017) aż 562 jest powiązanych z nowotworami (na podstawie wyszukiwania w bazie danych CancerMine (Lever i in., 2019)). Z grupy 119 retrogenów o zróżnicowanej ekspresji w nowotworach z repozytorium TCGA, wyselekcjonowano 39 retrokopii, których geny rodzicielskie odgrywają rolę w nowotworach człowieka (Tabela 16). Większość retrokopii (26 z 39) powstała z genów wykazujących rolę onkogeną. 9 retrogenów powstała na drodze retropozycji z genów, które określa się jako kluczowe dla inicjacji i progresji guzów nowotworowych (ang. *driver*). Są to retrogeny: *RHOB*, *H3P6*, *RPS15P4*, *RPL10P9*, *NPM1P27*, *ANXA2P2*, *GAPDH1P1*, *FABP5P7*, *CAPNS2*. Warto zaznaczyć, że spośród genów rodzicielskich *EEF1A1* i *RPS14* dały po 3 retrokopie wykazujące zmiany ekspresji w nowotworach.

Tabela 16. Retrogeny o zróżnicowanej ekspresji wyselekcjonowane na podstawie roli ich genu rodzicielskiego w nowotworach człowieka.

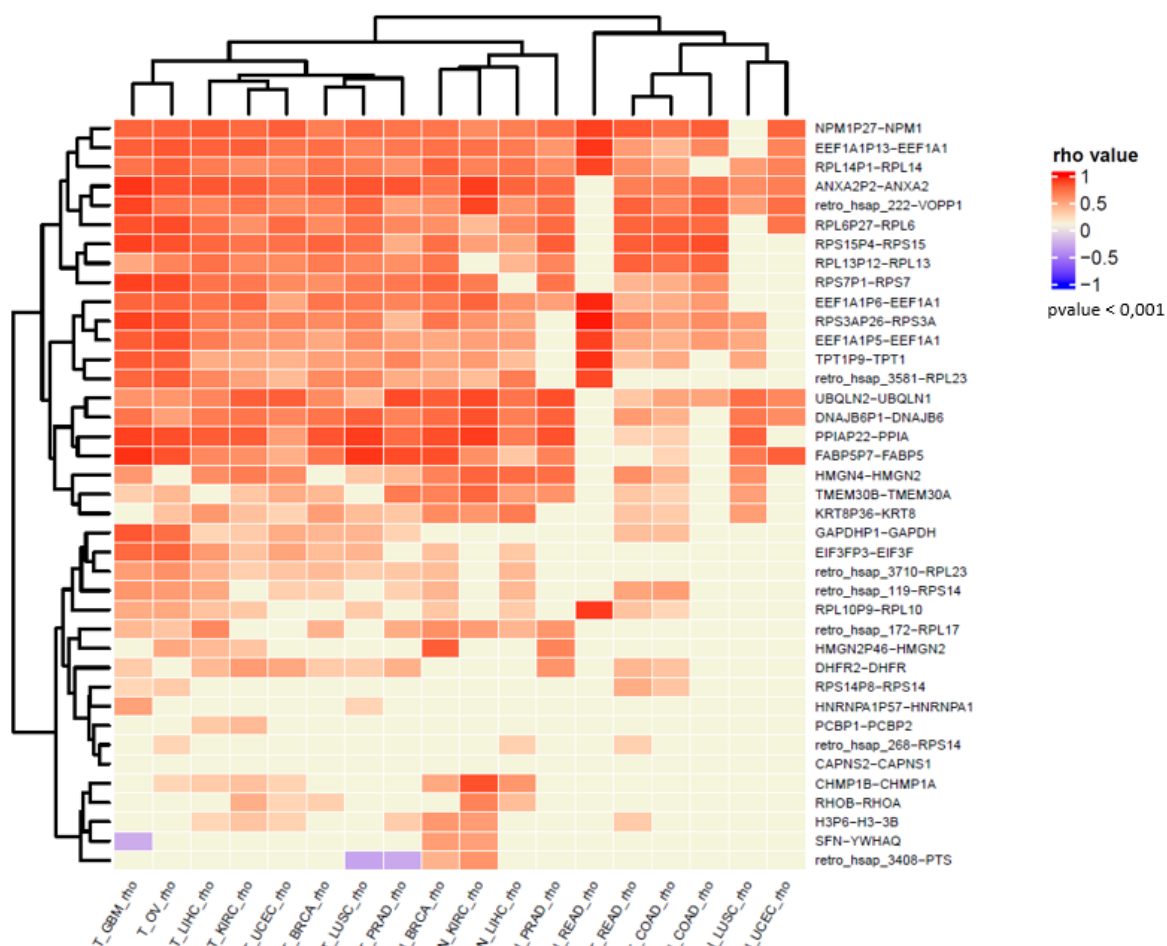
*onc – onkogen; sup – gen supresorowy; driv – driver

RetrogeneDB2 ID	Ensembl ID retrogeny	Nazwa retrogeny	Ensembl ID genu rodzicielskiego	Nazwa genu rodzicielskiego	Rola genu rodzicielskiego w nowotworach* (Lever i in., 2019)
retro_hsap_4150	ENSG00000231991	<i>ANXA2P2</i>	ENSG00000182718	<i>ANXA2</i>	onc, driv
retro_hsap_98	ENSG00000256812	<i>CAPNS2</i>	ENSG00000126247	<i>CAPNS1</i>	onc, driv
retro_hsap_75	ENSG00000255112	<i>CHMP1B</i>	ENSG00000131165	<i>CHMP1A</i>	sup, onc
retro_hsap_6	ENSG00000178700	<i>DHFR2</i>	ENSG00000228716	<i>DHFR</i>	onc
retro_hsap_939	ENSG00000254612	<i>DNAJB6P1</i>	ENSG00000105993	<i>DNAJB6</i>	onc
retro_hsap_4220	ENSG00000196205	<i>EEF1A1P5</i>	ENSG00000156508	<i>EEF1A1</i>	onc
retro_hsap_3825	ENSG00000233476	<i>EEF1A1P6</i>	ENSG00000156508	<i>EEF1A1</i>	onc

retro_hsap_3291	ENSG00000250182	EEF1A1P13	ENSG00000156508	EEF1A1	onc
retro_hsap_2106	ENSG00000233426	EIF3FP3	ENSG00000175390	EIF3F	sup
retro_hsap_880	ENSG00000234964	FABP5P7	ENSG00000164687	FABP5	sup, driv
retro_hsap_4806	ENSG00000228232	GAPDHP1	ENSG00000111640	GAPDH	onc, driv
retro_hsap_2180	ENSG00000235655	H3P6	ENSG00000132475	H3-3B	driv
retro_hsap_1492	ENSG00000179362	HMG2P46	ENSG00000198830	HMG2	onc
retro_hsap_100	ENSG00000182952	HMG4	ENSG00000198830	HMG2	onc
retro_hsap_2096	ENSG00000237442	HNRNPA1P57	ENSG00000135486	HNRNPA1	onc
retro_hsap_2842	ENSG00000240668	KRT8P36	ENSG00000170421	KRT8	onc
retro_hsap_3355	ENSG00000249353	NPM1P27	ENSG00000181163	NPM1	sup, onc, driv
retro_hsap_83	ENSG00000169564	PCBP1	ENSG00000197111	PCBP2	onc
retro_hsap_2497	ENSG00000198618	PPIAP22	ENSG00000196262	PPIA	onc
retro_hsap_119	ENSG00000226396	retro_hsap_119	ENSG00000164587	RPS14	sup
retro_hsap_172	ENSG00000226084	retro_hsap_172	ENSG00000265681	RPL17	sup
retro_hsap_222	ENSG00000223612	retro_hsap_222	ENSG00000154978	VOPP1	onc
retro_hsap_268	ENSG00000213058	retro_hsap_268	ENSG00000164587	RPS14	sup
retro_hsap_3408	ENSG00000253520	retro_hsap_3408	ENSG00000150787	PTS	sup, onc
retro_hsap_3581	ENSG00000180211	retro_hsap_3581	ENSG00000125691	RPL23	onc
retro_hsap_3710	ENSG00000232383	retro_hsap_3710	ENSG00000125691	RPL23	onc
retro_hsap_108	ENSG00000143878	RHOB	ENSG00000067560	RHOA	sup, onc, driv
retro_hsap_1905	ENSG00000235552	RPL6P27	ENSG00000089009	RPL6	onc
retro_hsap_3265	ENSG00000233913	RPL10P9	ENSG00000147403	RPL10	sup, onc, driv
retro_hsap_1810	ENSG00000215030	RPL13P12	ENSG00000167526	RPL13	onc
retro_hsap_1015	ENSG00000139239	RPL14P1	ENSG00000188846	RPL14	sup
retro_hsap_3879	ENSG00000214389	RPS3AP26	ENSG00000145425	RPS3A	onc
retro_hsap_1737	ENSG00000263266	RPS7P1	ENSG00000171863	RPS7	sup
retro_hsap_3232	ENSG00000239528	RPS14P8	ENSG00000164587	RPS14	sup
retro_hsap_2341	ENSG00000233762	RPS15P4	ENSG00000115268	RPS15	driv
retro_hsap_8	ENSG00000175793	SFN	ENSG00000134308	YWHAQ	sup
retro_hsap_12	ENSG00000182107	TMEM30B	ENSG00000112697	TMEM30A	sup
retro_hsap_4307	ENSG00000234782	TPT1P9	ENSG00000133112	TPT1	onc
retro_hsap_58	ENSG00000188021	UBQLN2	ENSG00000135018	UBQLN1	sup

Ze względu na mechanizm powstawania, retrokopie charakteryzują się wysokim stopniem podobieństwa sekwencji do swoich przodków, czyli genów rodzicielskich. Mimo, że potencjał kodowania białek nie zawsze jest zachowany, miejsca docelowe miRNA często pozostają w sekwencji retrokopii niezmienione. W związku z tym, posiadają miejsca docelowe dla tych samych miRNA co ich geny rodzicielskie i dlatego mogą działać jak gąbki miRNA regulujące ekspresję genów rodzicielskich, a także innych genów. Wcześniejsze analizy, wskazały m.in. *retro_hsap_119*, *RPS7P1* i *FABP5P7* jako retrogeny potencjalnie pełniące rolę gąbek (Kubiak i in., 2020). W celu zidentyfikowania retrogenów potencjalnie regulujących ich geny rodzicielskie obliczono współczynniki korelacji ekspresji. Korelację ekspresji obliczano

dla każdego z rodzajów nowotworów, osobno dla prób nowotworowych i zdrowych (Rycina 25). Szczegółowe dane dotyczące wartości współczynników korelacji oraz wartości p zawarte są w aneksie na końcu niniejszej pracy (Tabela 42).



Rycina 25. Retrogeny, które potencjalnie mogą regulować ekspresję genów rodzicielskich powiązanych z nowotworami.

Wykres typu mapa ciepła dla istotnych statystycznie współczynników korelacji ekspresji ($p < 0,001$ i $|\rho| > 0,25$). Nieistotne statystycznie wyniki oznaczono kolorem beżowym. Pomiedzy wierszami oraz kolumnami przeprowadzono grupowanie hierarchiczne w celu uzyskania odpowiednio klastrow genów oraz klastrow rodzajów nowotworów oraz tkanek okalających na podstawie korelacji ekspresji retrogenów i ich genów rodzicielskich.

T- próby nowotworowe, N-próby zdrowe

W przeważających przypadkach korelacja ekspresji pomiędzy retrokopią i genem rodzicielskim była korelacją pozytywną. Wiele retrogenów wykazywało taką korelację w większości analizowanych prób. Jedynie w przypadku dwóch retrogenów, *SFN* (ang. *stratifin*) oraz *retro_hsap_3408*, korelacja ekspresji z genem rodzicielskim przyjęła wartość ujemną (Rycina 25). Co ciekawe retrogeny te wykazywały jednocześnie w innych próbach korelację dodatnią. *SFN* i jego gen rodzicielski *YWHAQ* wykazały pozytywną korelację ekspresji w

próbach zdrowych piersi oraz nerki, natomiast negatywną w próbach nowotworowych z glejaka wielopostaciowego. *YWHAQ* (ang. *tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein theta*) wykazuje stosunkowo wysoką ekspresję w mózgu. Jak dotąd w literaturze nie opisano zależności poziomów ekspresji w przypadku *SFN* i jego genu rodzicielskiego *YWHAQ*.

Drugi retrogen, *retro_hsap_3408*, podobnie jak *SFN*, wykazał pozytywną korelację ekspresji z genem rodzicielskim *PTS* w próbach okalających raka piersi i raka jasnokomórkowego nerki (Rycina 25). Ujemną korelację ekspresji ze swoim genem rodzicielskim wykazywał jednak w innych niż *SFN* próbach – w próbach nowotworowych raka płaskonabłonkowego płuc oraz raka prostaty. Zjawisko zmiany wzorca korelacji ekspresji pomiędzy parami genów w nowotworach jest coraz częściej badane, ale dalekie od pełnego zrozumienia (Zhou i in., 2021).

W celu sprawdzenia mechanizmu regulacji ekspresji retrogen-gen rodzicielski sprawdzono, czy retrogeny które wykazały korelację ekspresji z genem rodzicielskim mają wspólne miejsca wiązania miRNA i co za tym idzie, czy retrogeny mogą działać jako potencjalne gąbki miRNA (Tabela 16). Okazało się, że 31 retrogenów, które wykazały korelację ekspresji ze swoim genem rodzicielskim może, przynajmniej teoretycznie, pełnić funkcję regulatorową poprzez wspólną pulę regulujących je miRNA. Dla przykładu, silną, pozytywną korelację ekspresji wykazała para *EEF1AIP13* i gen rodzicielski *EEF1A1*, która może być regulowana przez 4 wspólne miRNA. Retrogen *EEF1AIP13* (ang. *eukaryotic translation elongation factor 1 alpha 1 pseudogene 13*) wykazał zwiększony poziom ekspresji w raku jelita grubego (

Tabela 13). Wysoce prawdopodobne jest zatem, że zwiększenie ekspresji retrogeny, może skutkować zwiększonym wiązaniem miRNA, a w rezultacie wyższym poziomem ekspresji genu rodzicielskiego – onkogeny *EEF1A1* (ang. *eukaryotic translation elongation factor 1 alpha 1*).

Tabela 17. Potencjalne miejsca wiązania miRNA wspólne dla retrogenów i ich genów rodzicielskich (na podstawie miRDB.org (Y. Chen & Wang, 2020)).

Nazwa retrogeny	Nazwa genu rodzicielskiego	Wspólne miejsca wiązania miRNA (target score >60)
CAPNS2	CAPNS1	hsa-miR-5589-3p
CHMP1B	CHMP1A	hsa-miR-6721-5p
DHFR2	DHFR	hsa-miR-10523-5p, hsa-miR-1208, hsa-miR-146a-5p, hsa-miR-146b-5p, hsa-miR-200c-5p, hsa-miR-224-3p, hsa-miR-3182, hsa-miR-3191-5p, hsa-miR-3664-3p, hsa-miR-4437, hsa-miR-4477a, hsa-miR-451b, hsa-miR-5003-3p, hsa-miR-518c-5p, hsa-miR-522-3p, hsa-miR-550a-3p, hsa-miR-5699-5p, hsa-miR-6764-3p, hsa-miR-6824-3p, hsa-miR-7153-5p
DNAJB6P1	DNAJB6	hsa-miR-3664-3p, hsa-miR-4776-3p, hsa-miR-4786-3p, hsa-miR-500a-5p, hsa-miR-532-5p, hsa-miR-5590-5p, hsa-miR-670-5p
EIF3FP3	EIF3F	hsa-miR-335-3p, hsa-miR-4435, hsa-miR-6134, hsa-miR-6733-3p
EEF1A1P6	EEF1A1	hsa-miR-10397-5p, hsa-miR-1226-3p, hsa-miR-3064-5p, hsa-miR-3085-3p, hsa-miR-320a-5p, hsa-miR-3692-3p, hsa-miR-4303, hsa-miR-6504-5p
EEF1A1P13	EEF1A1	hsa-miR-3064-5p, hsa-miR-3085-3p, hsa-miR-320a-5p, hsa-miR-6504-5p
FABP5P7	FABP5	hsa-miR-10395-3p, hsa-miR-125b-2-3p, hsa-miR-4328, hsa-miR-4710, hsa-miR-501-3p, hsa-miR-502-3p, hsa-miR-552-3p, hsa-miR-578, hsa-miR-6738-3p, hsa-miR-6884-3p
GAPDHP1	GAPDH	hsa-miR-1226-5p, hsa-miR-1279, hsa-miR-196a-3p, hsa-miR-34a-5p, hsa-miR-34c-5p, hsa-miR-4443, hsa-miR-449a, hsa-miR-449b-5p, hsa-miR-4736, hsa-miR-6744-5p, hsa-miR-6745, hsa-miR-7150
HMG2P46	HMG2	hsa-miR-11181-3p, hsa-miR-5585-5p, hsa-miR-765
HMG2P4	HMG2	hsa-miR-11181-3p, hsa-miR-205-3p, hsa-miR-340-5p, hsa-miR-4287, hsa-miR-4446-5p, hsa-miR-4685-3p, hsa-miR-635, hsa-miR-6774-5p, hsa-miR-765
KRT8P36	KRT8	hsa-miR-4264, hsa-miR-4279, hsa-miR-4695-5p, hsa-miR-6736-3p, hsa-miR-766-3p
NPM1P27	NPM1	hsa-miR-1250-3p, hsa-miR-2113, hsa-miR-4327, hsa-miR-4694-5p, hsa-miR-5193, hsa-miR-5196-3p, hsa-miR-6072, hsa-miR-6740-3p, hsa-miR-6842-3p, hsa-miR-6885-3p, hsa-miR-6891-3p, hsa-miR-8069, hsa-miR-877-3p
PCBP1	PCBP2	hsa-miR-4767, hsa-miR-6777-3p
PPIAP22	PPIA	hsa-miR-4288, hsa-miR-4308, hsa-miR-4732-3p, hsa-miR-4776-3p, hsa-miR-629-3p, hsa-miR-632
retro_hsap_119	RPS14	hsa-miR-10524-5p, hsa-miR-15a-5p, hsa-miR-15b-5p, hsa-miR-16-5p, hsa-miR-195-5p, hsa-miR-424-5p, hsa-miR-4524a-5p, hsa-miR-4524b-5p, hsa-miR-497-5p, hsa-miR-518a-5p, hsa-miR-527, hsa-miR-616-3p, hsa-miR-6838-5p
retro_hsap_172	RPL17	hsa-miR-1207-5p, hsa-miR-2116-5p, hsa-miR-29b-1-5p, hsa-miR-3609, hsa-miR-3614-5p, hsa-miR-3686, hsa-miR-3942-3p, hsa-miR-3944-5p, hsa-miR-4677-5p, hsa-miR-548ah-5p, hsa-miR-619-5p, hsa-miR-6506-5p
retro_hsap_222	VOPP1	hsa-miR-1197
retro_hsap_268	RPS14	hsa-miR-518a-5p, hsa-miR-5190, hsa-miR-527
retro_hsap_3581	RPL23	hsa-miR-141-5p, hsa-miR-627-3p
RHOB	RHOA	hsa-miR-3148, hsa-miR-6785-5p
RPS7P1	RPS7	hsa-miR-1290, hsa-miR-1299, hsa-miR-22-3p, hsa-miR-326, hsa-miR-330-5p, hsa-miR-4279, hsa-miR-6736-3p, hsa-miR-6885-3p, hsa-miR-9902
RPS14P8	RPS14	hsa-miR-15a-5p, hsa-miR-15b-5p, hsa-miR-16-5p, hsa-miR-195-5p, hsa-miR-424-5p, hsa-miR-4524a-5p, hsa-miR-4524b-5p, hsa-miR-497-5p, hsa-miR-5190
RPS15P4	RPS15	hsa-miR-4291, hsa-miR-493-3p, hsa-miR-4778-3p, hsa-miR-4530
RPL10P9	RPL10	hsa-miR-1184, hsa-miR-3064-3p, hsa-miR-4466, hsa-miR-675-5p
RPL13P12	RPL13	hsa-miR-1248, hsa-miR-127-3p, hsa-miR-4287, hsa-miR-4685-3p, hsa-miR-6781-5p, hsa-miR-6852-5p
RPL14P1	RPL14	hsa-miR-1180-3p, hsa-miR-15a-5p, hsa-miR-15b-5p, hsa-miR-16-5p, hsa-miR-195-5p, hsa-miR-3192-3p, hsa-miR-3605-5p, hsa-miR-378g, hsa-miR-424-5p, hsa-miR-4303, hsa-miR-497-5p, hsa-miR-6838-5p

SFN	YWHAQ	<i>hsa-miR-626, hsa-miR-6852-3p, hsa-miR-6876-3p, hsa-miR-6878-3p, hsa-miR-8485</i>
TMEM30B	TMEM30A	<i>hsa-miR-2681-5p, hsa-miR-4282</i>
TPT1P9	TPT1	<i>hsa-miR-3074-5p, hsa-miR-3665, hsa-miR-636, hsa-miR-6516-3p, hsa-miR-6839-5p</i>
UBQLN2	UBQLN1	<i>hsa-miR-20a-3p, hsa-miR-33a-3p, hsa-miR-4698, hsa-miR-6835-5p</i>

Ponad 40% retrokopii jest zagnieżdżonych w intronach innych genów (2019 spośród 4611 retrogenów, na podstawie bazy danych RetrogeneDB2). Kiedy retrokopie występują na przeciwnej nici DNA i są aktywne transkrypcyjnie mogą regulować ekspresję gospodarza działając jako *cis*-NAT (Groen i in., 2014; Kubiak i in., 2020). Jednym z kilku możliwych mechanizmów może być interferencja transkrypcji (Rosikiewicz & Makałowska, 2016; Werner i in., 2024). Spośród genów gospodarzy tych zagnieżdżonych retrogenów ponad 500 powiązanych jest z występowaniem nowotworów u człowieka (według CancerMine, Lever i in., 2019). Z grupy 119 retrogenów o zróżnicowanej ekspresji w nowotworach wyselekcjonowano 9 retrokopii, których geny gospodarze odgrywają rolę w nowotworach człowieka (na podstawie bazy CancerMine) (Tabela 18).

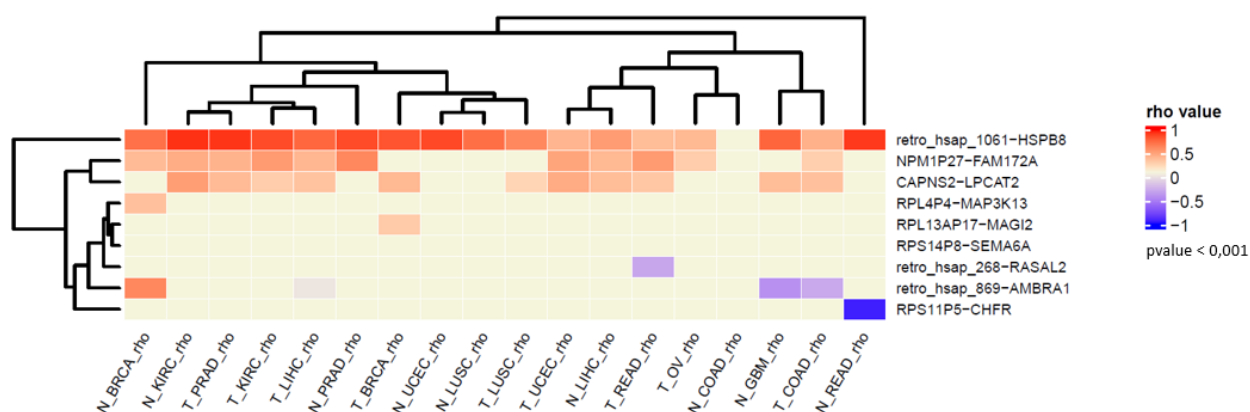
Tabela 18. Retrogeny o zróżnicowanej ekspresji wyselekcjonowane na podstawie roli ich genu gospodarza w nowotworach człowieka.

*onc – onkogen; sup – gen supresorowy; driv – driver

RetrogeneDB2 ID	Retrogen – Ensembl ID	Nazwa retrogeny	Gen rodzicielski – Ensembl ID	Nazwa genu rodzicielskiego	Nici DNA (retrogen / gospodarz)	Rola genu rodzicielskiego w nowotworach (Lever i in., 2019)
retro_hsap_98	ENSG00000256812	CAPNS2	ENSG00000087253	LPCAT2	+ / +	driv
retro_hsap_3355	ENSG00000249353	NPM1P27	ENSG00000113391	FAM172A	- / -	sup
retro_hsap_268	ENSG00000213058	retro_hsap_268	ENSG00000075391	RASAL2	+ / +	sup, onc, driv
retro_hsap_869	ENSG00000244313	retro_hsap_869	ENSG00000110497	AMBRA1	- / -	sup
retro_hsap_1061	ENSG00000213144	retro_hsap_1061	ENSG00000152137	HSPB8	+ / +	onc
retro_hsap_2898	ENSG00000229638	RPL4P4	ENSG00000073803	MAP3K13	+ / -	onc, driv
retro_hsap_3752	ENSG00000231322	RPL13AP17	ENSG00000187391	MAGI2	+ / -	sup, onc, driv
retro_hsap_1068	ENSG00000232888	RPS11P5	ENSG00000072609	CHFR	+ / -	sup, onc
retro_hsap_3232	ENSG00000239528	RPS14P8	ENSG00000092421	SEMA6A	+ / -	sup

W celu zidentyfikowania retrokopii potencjalnie mogących pełnić funkcję regulatorową, również obliczono korelację ekspresji pomiędzy retrogenem a genem gospodarza (Rycina 26). Istotną statystycznie korelację ekspresji w więcej niż jednym zestawie prób wykazały pary *retro_hsap_1061-HSPB8*, *NPM1P27-FAM172A* i *CAPNS2-LPCAT2*. Szczegółowe dane dotyczące wartości współczynników korelacji oraz wartości p zawarte są w aneksie na końcu niniejszej pracy (Tabela 43).

Podobnie jak w przypadku korelacji ekspresji retrokopii z ich genami rodzicielskimi, także w przypadku zagnieżdżonych retrogenów dominuje korelacja dodatnia. W wielu przypadkach korelacja ta jest bardzo silna i przekracza wartość 0.9 (Rycina 26, Tabela 43). *Retro_hsap_869* wykazał dodatnią wartość korelacji z genem gospodarzem *AMBRA1* w próbach zdrowych piersi (Rycina 26). Z kolei, ujemną wartość korelacji para tych genów wykazała w przypadku prób nowotworowych gruczolakoraka jelita grubego, glejaka wielopostaciowego oraz raka wątrobowokomórkowego. Zaobserwowany inny kierunek korelacji ekspresji może być powiązany z innymi mechanizmami regulatorowymi cechującymi *cis*-NATs (Rosikiewicz & Makołowska, 2016; Werner i in., 2024).



Rycina 26. Retrogeny jako regulatory genów gospodarzy, które powiązano z nowotworami. Wykres typu mapa ciepła dla istotnych statystycznie współczynników korelacji ekspresji ($p < 0,001$ i $|\rho| > 0,25$). Nieistotne statystycznie wyniki oznaczono kolorem beżowym. Pomiędzy wierszami oraz kolumnami przeprowadzono grupowanie hierarchiczne w celu uzyskania odpowiednio klastrow genów oraz klastrow rodzajów prób.

T- próby nowotworowe, N-próby zdrowe

W ramach niniejszej pracy postawiono hipotezę, która zakłada, że wśród retrogenów o zmienionej ekspresji w nowotworach przeważająca grupa to retrogeny powstałe z genów powiązanych z procesami nowotworowymi lub ulokowane w intronach takich genów. W celu weryfikacji tego założenia przeprowadzono test statystyczny chi-kwadrat (X^2). Wzięto pod uwagę łączną ilość retrogenów o zmienionej ekspresji, których geny rodzicielskie są związane z nowotworami (39) wraz z retrogenami ulokowanymi w tego typu genach (9). Cztery spośród nich powtarzały się, więc łączna liczba retrogenów wyniosła 44 (Tabela 19). Resztę spośród 119 retrogenów o zmienionej ekspresji w danych TCGA uznano za nie pochodzące od genów związanych z nowotworami (75). Z grupy retrogenów o niezmienionej ekspresji w nowotworach w danych pochodzących z TCGA (4461) wyodrębniono te, które pochodzą lub

są ulokowane w genach związanych z nowotworami (2816, na podstawie danych z CancerMine).

Tabela 19. Tablica kontyngencji dla testu statystycznego weryfikującego hipotezę 2.

	Retrogeny o zmienionej ekspresji w nowotworach	Retrogeny o niezmienionej ekspresji w nowotworach	Suma retrogenów
Retrogeny pochodzące od genów związanych z nowotworami lub ulokowane w nich	44	2816	2860
Retrogeny pochodzące od genów niezwiązanych z nowotworami	75	1645	1720
Suma retrogenów	119	4461	4580

Wynik testu tj. $X^2=32,692$, $df=1$, $p\text{-value}=1,08e-08$, wskazuje na istotną statystycznie różnicę pomiędzy badanymi grupami retrogenów, z tym, że nie potwierdza to jednak przyjętej hipotezy. Zmienioną w nowotworach ekspresję częściej wykazują retrokopie nie pochodzące od genów powiązanych z nowotworami i niezagnieżdżone w takowych.

Trzecia hipoteza badawcza zakładała, że retrogeny, które powstały z genów kodujących istotne dla procesów nowotworowych białka lub retrogeny ulokowane w intronach takich genów często regulują ekspresję tychże genów. Spośród 39 retrogenów o zmienionym poziomie ekspresji i pochodzących od genów rodzicielskich, które powiązano z nowotworami, 35 wykazało korelację ekspresji z genem rodzicielskim. Grupa 31 retrogenów dzieli wspólne miejsca wiązania miRNA ze swoim genem rodzicielskim, co może wskazywać na to, że regulują one ekspresję pełniąc rolę gąbek miRNA. Ponadto, w grupie 9 retrogenów ulokowanych w intronie innego genu, 8 wykazało istotną statystycznie korelację ekspresji z genem gospodarza.

4.6. Eksperymentalne analizy wpływu retrogenów na ekspresję genów rodzicielskich i żywotność komórek

Analizy bioinformatyczne pozwoliły na zidentyfikowanie 135 retrogenów o zmienionym poziomie ekspresji w nowotworach człowieka. Pozytywna korelacja ekspresji z genem rodzicielskim nasuwa sugestie regulatorowej roli tychże retrogenów. W związku z tym ostatnim etapem analiz było wyciszenie kilku wybranych retrogenów oraz sprawdzenie w jaki sposób obniżona ekspresja wpływa na poziom ekspresji genów rodzicielskich oraz na żywotność komórek. Eksperymenty przeprowadzono w liniach komórkowych MCF-7 (linia komórkowa ludzkiego estrogenozależnego gruczolakoraka piersi) oraz MCF-12F (linia

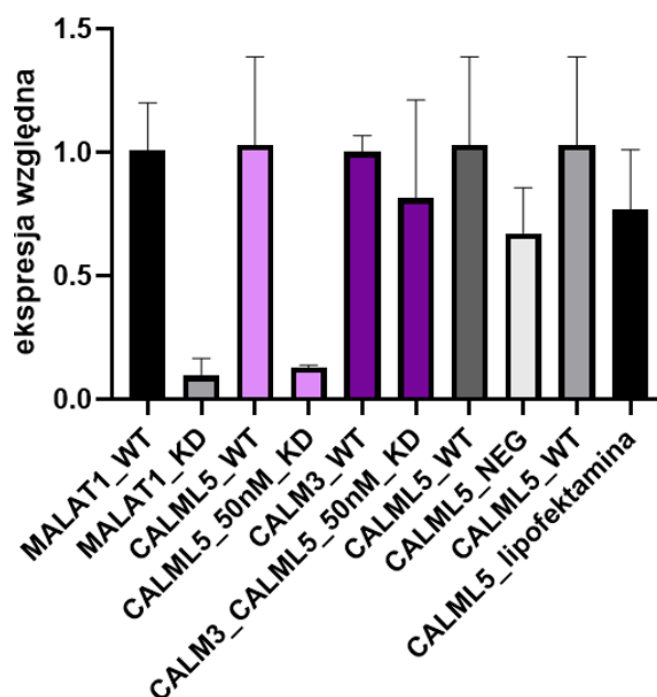
komórkowa ludzkiego gruczołu mlekowego). Ponieważ retrogeny mogą regulować inne geny zarówno w jądrze komórkowym na poziomie transkrypcji jak i w cytoplazmie, wyciszenie przeprowadzano za pomocą gapmeRów. siRNA nie są skuteczne w wyciszaniu transkryptów znajdujących się w jądrze komórkowym. Jako pozytywną kontrolę wykorzystano gen *MALAT1*. Sekwencje wykorzystanych w eksperymentach starterów znajdują się w Tabeli 5, a gapmeRów w Tabeli 7 w rozdziale „Materiały i metody”.

4.6.1. Wpływ retrogenów na poziom ekspresji genów rodzicielskich

Eksperymentalne badania związku pomiędzy ekspresją retrogenu i jego genu rodzicielskiego przeprowadzono dla 3 retrogenów. W dwóch przypadkach były to jednocześnie retrogeny zagnieżdżone w intronach innych genów. Zbadano zatem także wpływ wyciszenia retrogenu na gen gospodarza.

W przeprowadzonych analizach zidentyfikowano trzy ewolucyjnie stare retrogeny, których ekspresja różni się w komórkach nowotworowych w porównaniu z komórkami tkanki okalającej. Dwa z nich, *RHOB* i *HSPA2*, były już dobrze opisane w kontekście ich roli regulatorowej (Luis-Ravelo i in., 2014; Cao i in., 2019; Privat i in., 2020; Scieglinska i in., 2020). Do badań eksperymentalnych wybrano zatem jedynie *CALML5*, pomimo tego, że analizy korelacji nie wskazywały na powiązania pomiędzy ekspresją retrogenu i jego genu rodzicielskiego *CALM3*. Zarówno *CALML5* jak i *CALM3* zostały już wcześniej powiązane z nowotworami przez innych badaczy (Sunitha i in., 2020; Kanamori i in., 2023). Wzajemnej relacji pomiędzy tymi genami nigdzie jednak nie opisano.

Eksperyment wyciszenia przeprowadzono w linii komórkowej MCF-7, która jest linią komórkową gruczolaka piersi. Po wyciszeniu ekspresja retrogenu zmniejszyła się 8 razy (Rycina 27). Wyniki qPCR wykazały, że nawet tak duże wyciszenie nie miało wpływu na poziom ekspresji genu rodzicielskiego. Poziom ten obniżył się jedynie nieznacznie. Uzyskane wyniki potwierdzają wyniki analiz bioinformatycznych, które nie wskazywały na korelację ekspresji tych dwóch genów.

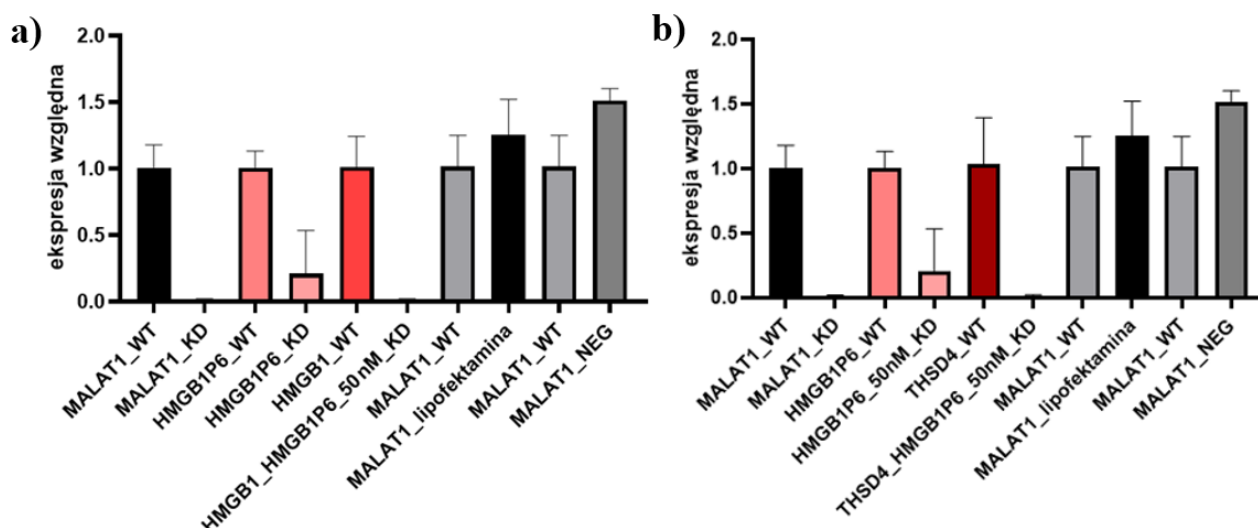


Rycina 27. Wpływ wyciszenia ekspresji retrogenu *CALML5* w linii nowotworowej MCF-7 na ekspresję genu rodzicielskiego *CALM3*.

Wyciszenie ekspresji *MALAT1* stanowiło kontrolę pozytywną gapmeRa. Kontrole negatywne obejmowały sprawdzenie ewentualnego wpływu lipofektaminy na ekspresję genów oraz wpływu samego gapmeRa – kontrola producenta (NEG).

Kolejnym retrogenem, który wyciszono w celu sprawdzenia, czy i w jakim stopniu wyciszenie to będzie wpływało na gen rodzicielski, był retrogen *HMGB1P6*. Analizy bioinformatyczne wykazały, że poziom ekspresji tego retrogenu ulega zwiększeniu w przypadku raka piersi i gruczolakoraka jelita grubego (Rycina 17ab). Zróżnicowana pomiędzy nowotworem i tkanką zdrową ekspresja została potwierdzona eksperymentem PCR w czterech liniach komórkowych (Rycina 18a). Dalsza analiza wykazała, że w obu typach nowotworów retrogen ten wykazuje silną, pozytywną korelację z genem rodzicielskim *HMGB1* (BRCA: $\rho=0,72$, $p=4,45e-23$; COAD: $\rho=0,92$; $p=1,8e-22$). *HMGB1* (ang. *high mobility group box*) należy do znanych genów nowotworowych. W raku piersi, opisano go jako supresor nowotworu (Martinotti i in., 2015) oraz gen związany z odpowiedzią na radioterapię (He et al. 2013). Taka silna korelacja ekspresji pomiędzy *HMGB1* a jego retrokopią *HMGB1P6* sugeruje potencjalną rolę regulatorową retrokopii. Możliwe, że *HMGB1P6* działa jako gąbka miRNA. Za pomocą bazy danych miRDB wskazano wspólne miejsca docelowe dla 15 miRNA (*hsa-miR-1184*, *hsa-miR-1248*, *hsa-miR-130b-5p*, *hsa-miR-149-3p*, *hsa-miR-374b-3p*, *hsa-miR-4468*, *hsa-miR-4728-5p*, *hsa-miR-4776-3p*, *hsa-miR-4778-3p*, *hsa-miR-5193*, *hsa-miR-6741-5p*, *hsa-miR-6785-5p*, *hsa-miR-6868-3p*, *hsa-miR-6883-5p*, *hsa-miR-877-3p*; miRDB, target score >60). A zatem, zwiększona ekspresja retrogeny może przyczyniać się do silniejszej sekwestracji miRNA i wyższego poziomu genu rodzicielskiego *HMGB1*. Podobnie jak w przypadku

CALML5, eksperyment wyciszenia przeprowadzono w ludzkiej linii nowotworowej piersi MCF-7. Jak wskazują wyniki qPCR, poziom ekspresji retrogenu *HMGBIP6* obniżył się pięciokrotnie, a efektem było niemal całkowite wyciszenie genu rodzicielskiego (Rycina 28a).



Rycina 28. Wpływ wyciszenia ekspresji retrokopii *HMGBIP6* w linii nowotworowej MCF-7 na ekspresję genu rodzicielskiego *HMGB1* (a) oraz na ekspresję genu gospodarza *THSD4* (b). Wyciszenie ekspresji *MALAT1* stanowiło kontrolę pozytywną gapmeRa. Kontrole negatywne obejmowały sprawdzenie ewentualnego wpływu lipofektaminy na ekspresję genów oraz wpływu samego gapmeRa – kontrola producenta (NEG).

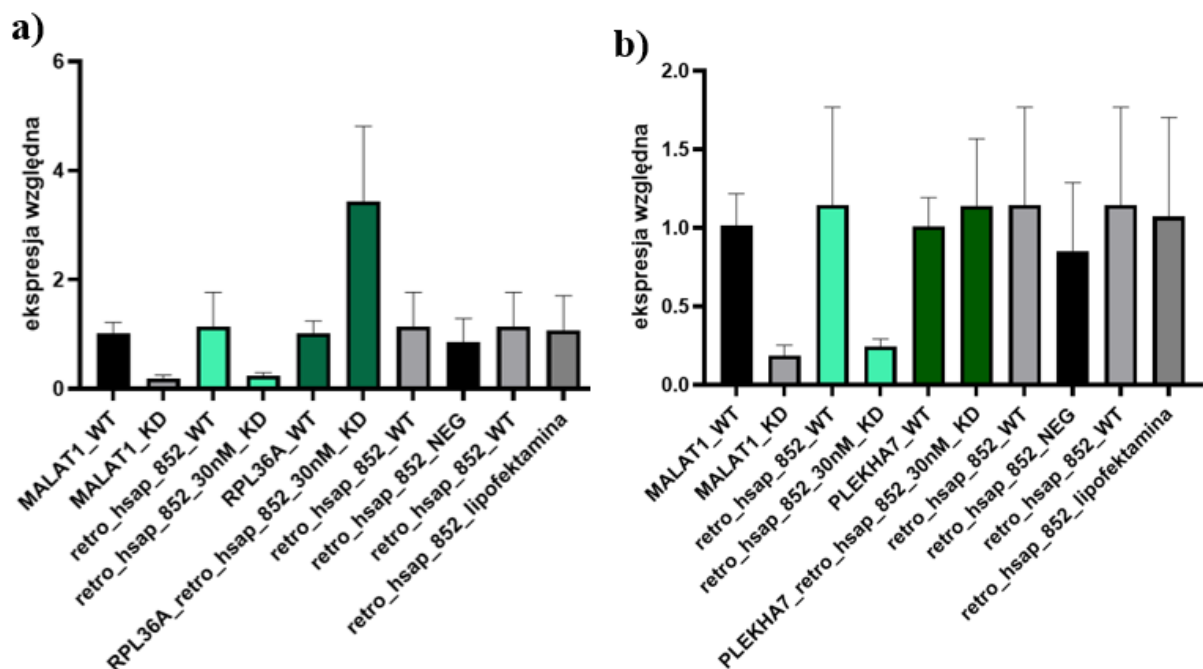
Retrogen *HMGBIP6* jest zagnieżdżony w intronie genu *THSD4* (ang. *thrombospondin type 1 domain containing 4*), którego zmienioną ekspresję powiązano z występowaniem raka prostaty (Wu i in., 2020) oraz z inwazyjnym rakiem piersi (The Human Protein Atlas). Podobnie jak w przypadku genu rodzicielskiego, wyciszenie *HMGBIP6* spowodowało spadek poziomu ekspresji genu *THSD4* niemal do zera (Rycina 28b).

Co ciekawe, inna retrokopia tego genu, *HMGBIP5*, która powstała wcześniej w ewolucji i jest specyficzna dla naczelnych, także wykazała zmianę w poziomie ekspresji, w raku piersi oraz dodatnią korelację ekspresji z genem rodzicielskim (BRCA: $\rho=0,68$, $p=4,19e-20$). Niestety ze względu na zbyt wysokie podobieństwo sekwencji retrogenu do sekwencji genu rodzicielskiego nie udało się zaprojektować starterów specyficznych dla każdego z tych genów, retrokopii i genu rodzicielskiego, co uniemożliwiło przeprowadzenie badań eksperymentalnych.

Trzeci retrogen, którego ekspresję poddano wyciszeniu to *retro_hsap_852*, należący do młodych ewolucyjnie retrogenów o zmienionej ekspresji w 8 nowotworach człowieka (Tabela 8, Tabela 13) oraz podczas ewolucji guza nowotworowego (Tabela 9). *Retro_hsap_852* powstał na drodze retropozycji jako kopia genu rodzicielskiego *RPL36A* (ang. *ribosomal protein L36a*).

Retro_hsap_852 wykazał stałą, silną pozytywną korelację z genem rodzicielskim (*RPL36A*) w 5 nowotworach człowieka (BRCA: $\rho=0,84$ $p<0,001$; COAD: $\rho=0,68$ $p<0,001$; LUAD= $0,7$ $p<0,001$; HCC= $0,76$ $p<0,001$; HB= $0,67$ $p<0,001$; na podstawie danych z bazy NCBI SRA). Najnowsze doniesienia wskazują, że zwiększony poziom ekspresji *RPL36A* jest związany ze wzrostem komórek raka okrężnicy, sugerując jednocześnie możliwość wykorzystania tego genu jako potencjalnego celu terapeutycznego w leczeniu nowotworu (Shi i in., 2025). Wyraźna zależność ekspresji pomiędzy *retro_hsap_852* a *RPL36A*, wskazuje na możliwą regulację ekspresji genu rodzicielskiego przez jego retrokopię. Zarówno retrokopia jak i gen rodzicielski posiadają wspólne miejsca wiązania dla trzech miRNA (*hsa-miR-4327*, *hsa-miR-600*, *hsa-miR-6886-3p*; miRDB, target score>60).

Retro_hsap_852 wyciszono w linii komórkowej ludzkiego gruczołu mlekowego MCF-12F. Spośród czterech badanych linii retrogen ten wykazywał ekspresję jedynie w liniach nienowotworowych MRC-5 i MCF-12F, w tym zdecydowanie większą w MCF-12F (Rycina 18). Wynik analizy qPCR wykazał, że po wyciszeniu ekspresja *retro_hsap_852* została obniżona ponad trzykrotnie. W rezultacie wyciszenia ekspresji *retro_hsap_852* zaobserwowano znacznie zwiększoną ekspresję genu rodzicielskiego *RPL36A* (Rycina 29a). Wynik ten jest zaskakujący, gdyż jest on odwrotny do wyników analizy korelacji ekspresji wskazujących na pozytywną, a nie negatywną korelację ekspresji, zarówno w próbach z tkanek nowotworowych jak i zdrowych. Należy jednak pamiętać, że eksperyment dotyczył wielokrotnie pasażowanej linii komórkowej. Wiadomym jest natomiast, że z czasem i rosnącą liczbą pasaży procesy biologiczne zachodzące w komórkach tychże linii ulegają zmianie. W związku z tym nie zawsze w prawidłowy sposób odzwierciedlają procesy zachodzące w żywym organizmie i nie zawsze, tak jak być może w tym przypadku, są najlepszym modelem. Z drugiej strony, są często jedynym modelem możliwym do wykorzystania.



Rycina 29. Wpływ wyciszenia ekspresji retrogenu *retro_hsap_852* w linii nienowotworowej MCF-12F na ekspresję genu rodzicielskiego *RPL36A* (a) oraz na ekspresję genu gospodarza *PLEKHA7* (b).

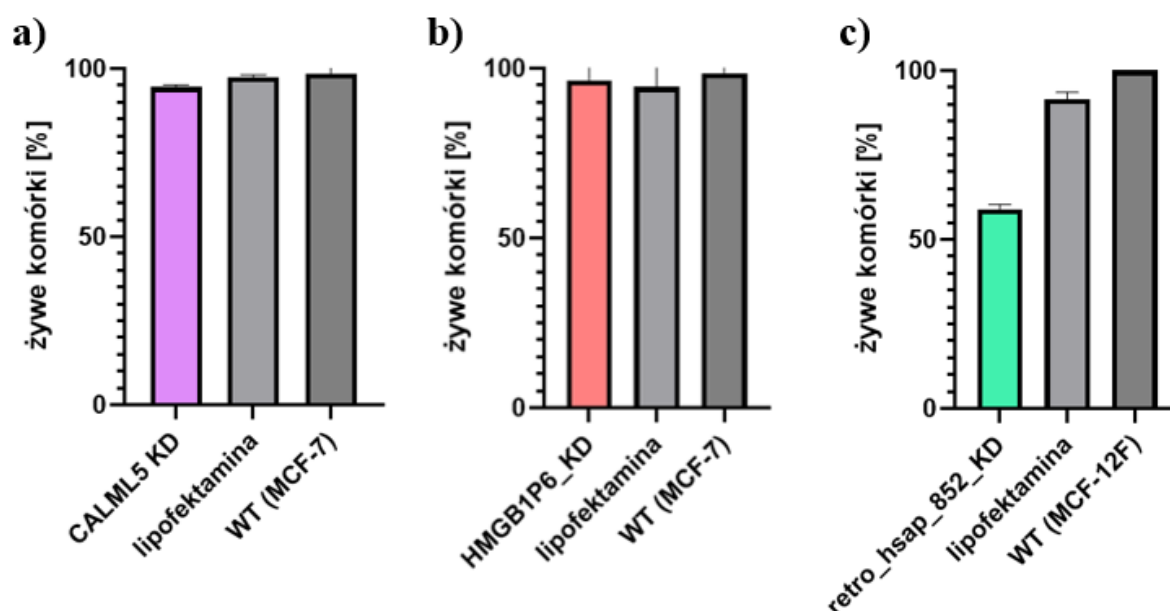
Wyciszenie ekspresji *MALAT1* stanowiło kontrolę pozytywną gapmeRa. Kontrole negatywne obejmowały sprawdzenie ewentualnego wpływu lipofektaminy na ekspresję genów oraz wpływu samego gapmeRa – kontrola producenta (NEG).

Retro_hsap_852 jest przykładem retrokopii ulokowanej w intronie genu gospodarza *PLEKHA7* (ang. *pleckstrin homology domain containing A7*). Nie zauważono wpływu wyciszenia retrogenu na ekspresję genu gospodarza *PLEKHA7* (Rycina 29b). Co ciekawe, zaobserwowano wzrost ekspresji *PLEKHA7* w miarę wzrostu stopnia zaawansowania raka jelita grubego (P. Li i in., 2024).

4.6.2. Wpływ wyciszenia ekspresji retrogenów na żywotność komórek

Różnice w poziomie ekspresji pomiędzy tkankami nowotworowymi i zdrowymi mogą być jedynie efektem ubocznym globalnych zmian metabolicznych jakie zachodzą w komórkach nowotworowych. Nie można jednak wykluczyć bezpośredniego wpływu na te procesy badanych retrogenów. Wykorzystując przeprowadzone już eksperymenty wyciszania retrogenów, zbadano w dalszej kolejności na ile wyciszanie to wpłynęło na żywotność komórek linii nowotworowych. Jak wskazują wyniki, procentowa zawartość żywych komórek nie zmieniła się w przypadku wyciszania *CALML5* i *HMGB1P6* w nowotworowej linii komórkowej MCF-7 (odpowiednio Rycina 30a i Rycina 30b). W przypadku retrogenu

retro_hsap_852, wyciszonego w komórkach pochodzących z gruczołu piersiowego (linia komórkowa MCF-12F) zaobserwowano niemal pięćdziesięcioprocentowy spadek zawartości komórek żywych (Rycina 30c). Retrogen ten wykazał zwiększoną ekspresję aż w 6 z 8 badanych nowotworów (Tabela 13). Spadek żywotności komórek nienowotworowych po obniżeniu poziomu ekspresji tego retrogeny może sugerować, iż *retro_hsap_852* pełni istotną rolę w funkcjonowaniu komórek.



Rycina 30. Wpływ wyciszenia ekspresji retrogeny *CALML5* (a), *HMBG1P6* (b) oraz *retro_hsap_852* (c) na żywotność komórek.

4.7. Identyfikacja retrogenów związanych zarówno z nowotworami jak i rozwojem łożyska człowieka

Procesy rozwoju łożyska wykazują wiele podobieństw z przerzutami nowotworowymi (W. Zhang & Huang, 2011). Sugeruje się, że gatunki charakteryzujące się bardziej inwazyjnym typem łożyska (np. człowiek) są bardziej podatne na nowotwory złośliwe niż gatunki o nieinwazyjnym typie łożyska, przykładowo przeżuwacze (Kshitiz i in., 2019). W związku z tym przeprowadzono analizy ekspresji różnicowej porównując dane z fibroblastów zrębu endometrium po inwazji trofoblastem z fibroblastami zrębu endometrium (Tabela 20). Celem było sprawdzenie czy retrogeny wykazujące zmiany w ekspresji w nowotworach również w tym przypadku będą je wykazywały.

Zidentyfikowano 6 retrogenów o zmniejszonej ekspresji oraz 18 retrogenów o zwiększonej ekspresji (Tabela 20). Co ciekawe, 14 retrogenów o istotnie zmienionym profilu ekspresji w procesie inwazji łożyska zidentyfikowano także w poprzednich analizach nowotworów człowieka. Należą do nich: *EEF1A1P5*, *FABP5P7*, *FTLP14*, *H3P6*, *PCBP1*, *retro_hsap_1605*, *retro_hsap_2297*, *retro_hsap_852*, *RHOG*, *RPL10P9*, *RPL13P12*, *RPL18AP3*, *RPL21P16* oraz *RPS2P5*. Część z nich wykazała kierunek zmiany ekspresji zgodny z wynikami poprzednich analiz (*FTLP14*, *H3P6*, *PCBP1*, *retro_hsap_1605*, *retro_hsap_852*, *RPL18AP3*, *RPS2P5*), dokładniej wzrost ekspresji zarówno w nowotworach jak i podczas rozwoju łożyska. Może to sugerować związek tych retrogenów ze zjawiskiem nabywania cech złośliwości przez guzy nowotworowe. Szczegółowe informacje dotyczące retrogenów o zmienionej ekspresji (log2foldchange, wartości p oraz zmiany w wartościach TPM) znajdują się w tabeli na końcu niniejszej pracy (Aneks, Tabela 44).

Tabela 20. Retrogeny o zmienionym poziomie ekspresji podczas rozwoju łożyska człowieka oraz o zmienionym poziomie ekspresji w nowotworach człowieka. W przypadku braku wspólnych retrogenów o zmienionej ekspresji danego porównania nie uwzględniono w tabeli. Czerwoną strzałką oznaczono zwiększony poziom ekspresji a zieloną zmniejszony poziom ekspresji. BRCA- rak piersi, ER+- rak piersi z pozytywnym statusem receptorów estrogenowych, TNBC- potrójnie ujemny rak piersi, COAD- gruczolakorak jelita grubego, HB- wątrobiak zarodkowy, C2, C1- podtypy molekularne wątrobiaka zarodkowego, HCC- rak wątrobowokomórkowy, KIRC- rak jasnokomórkowy nerki, LIHC- rak wątrobowokomórkowy, LUAD- gruczolakorak płuc, LUSC- rak płaskonabłonkowy płuc, OV- rak jajnika, PRAD- gruczolakorak prostaty, READ- gruczolakorak odbytnicy, UCEC- rak endometrium trzonu macicy.

Retrogen	Gen rodzicielski	NCBI SRA						TCGA									
		Rozwój łożyska	BRCA	BRCA (TNBC vs ER+)	COAD	HCC	LUAD	BRCA	COAD	KIRC	LIHC	LUSC	LUSC (postępujący guz vs remisja)	PRAD	READ	UCEC	OV (postępujący guz vs remisja)
EEF1A1P5	EEF1A1	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	↓	NA	NA	NA	↓	NA
EIF1AXP1	EIF1AX	↓	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
FABP5P7	FABP5	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	↑	NA	NA	NA	NA	NA
FTLP14	FTL	↑	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
GCSHP5	GCSH	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
H3P6	H3-3B	↑	↑	NA	NA	NA	NA	↑	NA	NA	NA	↑	NA	NA	NA	↑	NA
PCBP1	PCBP2	↑	NA	NA	NA	NA	NA	NA	↑	NA	NA	NA	NA	NA	↑	↑	NA
retro_hsap_1605	RPL9	↑	↓	NA	NA	↑	↓	NA	NA	↑	NA	NA	NA	NA	NA	↓	NA
retro_hsap_1849	AP1S2	↓	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
retro_hsap_1969	TMEM167A	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
retro_hsap_2297	PPIA	↑	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
retro_hsap_3402	TMA7	↓	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
retro_hsap_852	RPL36A	↑	NA	↑	↑	↑	↑	NA	↑	↑	↑	↑	NA	↑	↑	NA	NA
RHOG	RAC1	↓	NA	NA	NA	↑	NA	NA	NA	↑	NA	↓	NA	NA	NA	NA	NA
RPL10P9	RPL10	↓	NA	NA	NA	NA	NA	↑	↑	↑	NA	NA	↓	NA	↑	NA	NA
RPL13P12	RPL13	↓	NA	NA	NA	NA	NA	NA	↑	↑	NA	↑	NA	↑	↑	NA	↑
RPL18AP3	RPL18A	↑	NA	NA	NA	NA	NA	NA	↑	↑	NA	↑	NA	NA	NA	NA	NA
RPL21P16	RPL21	↑	↓	NA	NA	NA	NA	NA	NA	↑	NA	NA	NA	NA	NA	NA	NA
RPS23P8	RPS23	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
RPS2P5	RPS2	↑	NA	NA	NA	↑	NA	NA	↑	↑	NA	↑	NA	↑	↑	NA	NA
SNRPEP4	SNRPE	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SNRPGP15	SNRPG	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
TMEM183B	TMEM183A	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ZFP42	YY1	↑	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Na szczególną uwagę zasługuje *retro_hsap_852*, zidentyfikowany wcześniej jako retrogen specyficzny dla człowieka, który posiada istotnie statystycznie zwiększony poziom ekspresji w 8 badanych nowotworach. Ponadto, zwiększył swój poziom ekspresji także w bardziej agresywnym typie nowotworu piersi TNBC oraz u pacjentów z przerzutującym gruczolakorakiem jelita grubego, a jego wyciszenie wpłynęło negatywnie na żywotność komórek badanej linii komórkowej. Łącząc te wyniki ze zwiększoną ekspresją *retro_hsap_852* także podczas inwazji łożyska można wnioskować o jego istotnej roli podczas rozwoju nowotworów człowieka.

Retrogeny o zmienionej ekspresji podczas rozwoju łożyska oraz ich geny rodzicielskie sprawdzono pod kątem specyficzności tkankowej z użyciem bazy danych Gtex (GTEx Consortium, 2013). Retrogeny te tworzą dwie główne grupy różniące się specyfiką tkankową, co może odzwierciedlać ich potencjalne funkcje biologiczne podczas inwazji łożyska jak i rozwoju guza nowotworowego. Pierwsza obejmuje retrogeny, które są najbardziej aktywne transkrypcyjnie w jajnikach. Należą do nich: *retro_hsap_852*, *retro_hsap_1605*, *RPL21P16*, *RPS23P8* oraz *RPS2P5*. Co ciekawe ich geny rodzicielskie dzielą ten wzór ekspresji, także charakteryzują się najwyższą ekspresją w jajnikach człowieka. Wysoka ekspresja retrogenów w jajnikach może odzwierciedlać ich rolę w procesach związanych z reprodukcją, które wymagają intensywnej proliferacji i angiogenezy, takich jak rozwój łożyska. Te same procesy zachodzą w komórkach nowotworowych, co podkreśla funkcjonalne podobieństwa między procesami związanymi z rozwojem łożyska a patologicznymi procesami nowotworowymi.

Druga grupa retrogenów wykazuje bardzo wysoką ekspresję w komórkach krwi, co może świadczyć o ich zaangażowaniu w procesy hematopoetyczne oraz immunologiczne. Do tej grupy zaliczają się retrogeny o zróżnicowanym pochodzeniu ewolucyjnym, takie jak *RHOG* (ang. *ras homolog family member G*), który powstał u wspólnego przodka kręgowców, *PCBP1* (ang. *poly(rC) binding protein 1*), który powstał w czasie ewolucji ssaków łożyskowych, oraz *H3P6* (ang. *H3 histone pseudogene 6*), który pojawił się w trakcie ewolucji naczelnych. Wysoka ekspresja wspomnianych retrogenów we krwi może mieć znaczenie w procesach ukrwienia, istotnych zarówno dla rozwoju łożyska, jak i w angiogenezie, która odgrywa kluczową rolę w progresji nowotworów. Bogate unaczynienie zarówno w przypadku łożyska, jak i w przypadku nowotworów pozwala na efektywne dostarczenie składników pokarmowych potrzebnych do intensywnych procesów metabolicznych. Jednoznaczne stwierdzenie czy wysoka ekspresja badanych retrogenów istotnie ma znaczenie w tych procesach wymaga oczywiście dużo bardziej szczegółowych badań.

5. DYSKUSJA

Zrozumienie mechanizmów powstawania nowotworów, rozpoznanie istotnych w tym procesie genów i ich roli są kluczowe dla poprawnej diagnostyki, a co za tym idzie wyboru odpowiednich terapii. Efektywność diagnostyki onkologicznej nadal pozostaje niesatysfakcjonująca, dlatego konieczne jest opracowanie nowych biomarkerów (F. Liu i in. 2019). Badania z wykorzystaniem hodowli komórkowych, modeli zwierzęcych i próbek klinicznych potwierdzają rolę pseudogenów w procesie nowotworzenia (Kalyana-Sundaram i in. 2012). Jednym z rodzajów genów zaliczanych najczęściej do pseudogenów są retrogeny. Pomimo powszechnego uważania je za niefunkcjonalne, geny powstałe w wyniku retropozycji pełnią szereg funkcji w komórce. W wyniku procesów ewolucyjnych początkowo nieaktywne retrokopie nabierają funkcjonalności i mogą zarówno kodować białka jak pełnić role regulatorowe (Kubiak, Szcześniak, i Makałowska 2020). Bardzo często regulują poziom ekspresji swoich genów rodzicielskich działając jako gąbki miRNA czy antysensowne transkrypty oraz mogą brać udział w regulacji epigenetycznej, modyfikacji alternatywnego splicingu czy w formacji transkryptów fuzyjnych (Kubiak, Szcześniak, i Makałowska 2020). Pomimo, że udało się powiązać wiele retrogenów z procesami nowotworowymi (Fairbanks i in., 2012; Rohozinski i in., 2012; Yu i in., 2014; J. A. Ju & Gilkes, 2018), wiedza na temat roli retrogenów podczas kancerogenezy jest nadal mocno ograniczona.

Jak wskazują wyniki analiz, proces retropozycji był szczególnie wzmożony podczas ewolucji naczelnych, więc bardzo duża część retrokopii w genomie człowieka należy do stosunkowo młodych (Navarro & Galante, 2015). Rola nowo powstałych genów stanowi fascynujący obszar badań z punktu widzenia ewolucji (Cai & Petrov, 2010). Zaproponowano, że nowe geny mogą pełnić istotną rolę w procesach, które pozostają pod silną presją selekcyjną. Do tego rodzaju procesów należą m.in. spermatogeneza, mechanizmy odpowiedzi immunologicznej czy procesy związane z rozwojem narządów, takich jak inwazyjne łożysko czy rozwinięty mózg (Shao i in., 2019). W literaturze opisano, że ekspresja niektórych ewolucyjnie młodych genów występuje specyficznie lub preferencyjnie w nowotworach (tzw. TSEEN, ang. *tumor-specifically expressed, evolutionarily novel*) (Kozlov, 2016). Identyfikacja retrogenów specyficznych dla człowieka, które są związane z nowotworami, oraz różnice w repertuarze retrokopii w przypadku nowotworów różnych gatunków kręgowców, może być kluczowa dla lepszego zrozumienia wysokiej częstości nowotworów u naszego gatunku. Ponadto retrogeny należące do TSEEN, ze względu na swoją gatunkową i tkankową

specyficzność, mogą mieć ogromny potencjał jako nowe biomarkery nowotworowe (Staszak & Makałowska, 2021)

W niniejszej pracy przeprowadzono analizy ekspresji retrogenów w celu identyfikacji tych, które potencjalnie mogą być powiązane z powstawaniem i rozwojem nowotworów. Na zagadnienie to spojrzano z perspektywy ewolucyjnej porównując młode retrogeny specyficzne dla naczelnych z tymi, które powstały na wcześniejszych etapach ewolucji. Przeprowadzono także analizy ekspresji retrogenów psa i kury ze szczególną uwagą poświęconą retrogenom zmieniającym ekspresję we wszystkich trzech gatunkach i tym gatunkowo specyficznym. Sprawdzone także czy retrogeny, które powstały z genów opisanych jako onkogeny lub supresory mogą również w jakimś stopniu być powiązane z nowotworami i czy mogą regulować ekspresję swoich genów rodzicielskich. W trakcie retropozycji wiele kopii genów lokowanych jest w intronach innych genów. Zadano sobie zatem pytanie czy retrogeny zagnieżdżone w intronach genów istotnych w procesach nowotworowych mogą pełnić rolę regulatorów ekspresji tychże genów. Ostatni aspekt ewolucyjny dotyczył powiązań pomiędzy rozwojem inwazyjnego łóżyska i powstawaniem nowotworów. Szukano odpowiedzi na pytanie czy retrogeny zmieniające poziom ekspresji w nowotworach zmieniają też swoją ekspresję wraz z rozwojem łóżyska.

Przeprowadzone analizy retrogenów są wyjątkowe zarówno w swoim rozmiarze jak i w kwestii podejmowanych problemów. Nie zbadano dotychczas profilu ekspresji retrokopii w nowotworach na tak dużą skalę jak w niniejszej pracy.

Retrogeny o zmienionym poziomie ekspresji w nowotworach człowieka

Aby zbadać różnice w ekspresji retrogenów pomiędzy tkankami nowotworowymi a zdrowymi przeanalizowano dane RNA-Seq pochodzące z 5 nowotworów człowieka z repozytorium NCBI SRA oraz 8 nowotworów z repozytorium TCGA. Na podstawie uzyskanych wyników analiz bioinformatycznych udało się zidentyfikować odpowiednio 36 oraz 119 retrogenów o istotnie zróżnicowanym poziomie ekspresji. Grupa 20 retrogenów została wskazana w wynikach analiz danych z obydwóch źródeł. A zatem, łącznie zidentyfikowano 135 retrokopii o zmienionym poziomie ekspresji w nowotworach człowieka. Otrzymane wyniki obejmują dobrze znane retrogeny, o bogatym opisie literaturowym, tj. *RHOB* (Connolly i in., 2010; Ridley, 2013; Luis-Ravelo i in., 2014; J. A. Ju & Gilkes, 2018) czy *SFN* (Hu i in., 2019; Du i in., 2024), które mogą stanowić pozytywną kontrolę przeprowadzonych analiz. Jednak w większości zidentyfikowano nieopisane dotąd retrogeny. Wyjaśnieniem tego, że wielu retrokopii nie wskazano w badaniach prowadzonych przez inne

grupy może być to, że retrokopie są często pomijane w analizach genów kandydujących. Nie bez znaczenia jest również fakt, że analiza retrokopii jest znacznym wyzwaniem ze względu na niski poziom ekspresji w porównaniu z genami kodującymi białka oraz wysoki poziom podobieństwa sekwencji do genów rodzicielskich. Z reguły analizy danych transkryptomicznych skupiają się na genach o wyższej ekspresji, do których należą głównie geny kodujące białka. Poziom ten często jest wyznaczany na minimum 5 TPM co w przypadku wielu retrogenów, podobnie jak wielu długich niekodujących RNA (lncRNA) jest on zdecydowanie za wysoki (Q. Xu i in., 2017; Grammatikakis & Lal, 2022). Duża część dostępnych danych RNA-Seq zawiera informacje o ekspresji jedynie ok. 30 tysięcy genów. W takich bibliotekach nie znajdziemy informacji o większości genów, które nie kodują białek, w tym również większości retrogenów. Dla przykładu duża część zestawów danych w repozytorium TCGA zawiera macierze ekspresji 20531 genów, więc w analizach z ich użyciem nie zidentyfikujemy retrogenów.

Analiza ewolucyjna retrogenów o zróżnicowanej ekspresji w nowotworach człowieka

Aspekt ewolucyjny przeprowadzonych badań opierał się między innymi na analizie danych transkryptomicznych nowotworów różnych gatunków kręgowców oraz określeniu czasu powstania retrogenów o zmienionej ekspresji w nowotworach człowieka. Analiza filogenetyczna wykazała, że większość retrogenów związanych z nowotworami człowieka powstała podczas ewolucji naczelnych. W przypadku danych NCBI SRA jest to grupa 21 spośród 34 (61,8%), a w danych TCGA 89 spośród 118 (75,4%). A zatem wśród zidentyfikowanych retrokopii wyraźnie dominują retrogeny specyficzne dla człowieka lub naczelnych, co potwierdza przyjętą hipotezę.

Przeprowadzono także analizy ekspresji różnicowej dwóch zestawów danych z gruczolu mlekowego psa oraz zestaw danych z linii komórkowej raka wątrobowokomórkowego kury. Na tej podstawie zidentyfikowano trzy ewolucyjnie stare retrokopie *RHOB* i *HSPA2*, które wykazały zmianę ekspresji w danych nowotworowych wszystkich trzech gatunków oraz retrogen specyficzny dla ssaków *CALML5*, który wykazał zmianę ekspresji w nowotworach człowieka i psa. Wskazanie tej grupy jako retrogeny związane z nowotworami odległych ewolucyjnie gatunków może sugerować ich rolę w kancerogenezie nie tylko u człowieka co wielokrotnie opisano, ale być może u wszystkich kręgowców.

Pierwszy z nich *RHOB* to przykład retrogenu kodującego białko. *RHOB* jest zaangażowane w regulację cytoszkieletu aktynowego oraz migracji komórek (Ghafouri-Fard i

in., 2021). Retrogen *RHOB* powstał u wspólnego przodka kręgowców. Zachowanie ORF w tym starym pod względem ewolucyjnym retrogenie może wskazywać na to, że bardzo szybko nabył on zdolności transkrypcyjne, a skłonność do akumulacji mutacji, typowa dla retrokopii, została „zablokowana” ze względu na funkcjonalne znaczenie białka. *RHOB* posiada bogaty opis literaturowy, w którym jest wskazywany głównie jako supresor nowotworowy (W. Chen i in., 2016). Gen został opisany także jako czynnik promujący kancerogenezę. W przypadku tworzenia się guzów piersi wykazano, że *RHOB* działa jako supresor do momentu ustabilizowania się unaczynienia guza, wówczas sprzyja progresji komórek nowotworowych (J. A. Ju & Gilkes, 2018). W literaturze opisano eksperymenty, które wykazały związek wyciszenia ekspresji *RHOB* ze wzrostem ekspresji genu rodzicielskiego *RHOA* (Privat i in., 2020). Choć funkcja kodowanego przez ten retrogen białka jest dość dobrze poznana, jak dotychczas, nie opisano jego związku z nowotworami u psa i kury. Badano go jednak w przypadku ślepego kretoszczura podziemnego (łac. *Spalax ehrenbergi*), gatunku określanego jako odporny na nowotwory. Odkryto istotny wzrost ekspresji *RHOB* w warunkach niedotlenienia u dwóch allogatunków: *S.galili* i *S.judaei* (Avivi i in., 2006). Badania wykonane w przedstawionej pracy wskazują natomiast, że gen ten może być u wszystkich kręgowców stałocieplnych równie istotny w procesach nowotworowych jak u człowieka.

Retrogen *HSPA2* koduje białko należące do rodziny białek szoku cieplnego o masie 70 kDa (HSP70) i pełniące rolę białka opiekuńczego. Wysoki poziom ekspresji *HSPA2* stwierdzono szczególnie w jądrach co jest uzasadnione funkcją kodowanego białka. Białko *HSPA2* odgrywa istotną rolę w spermatogenezie, chroniąc przed negatywnym wpływem wysokiej temperatury na plemniki. Jest to szczególnie ważne u ssaków z jądrami zstępującymi (Padhi i in., 2016). W analizach przeprowadzonych w ramach niniejszej pracy spadek ekspresji *HSPA2* zaobserwowano dopiero w późniejszych stadiach progresji nowotworu jelita grubego i bardziej inwazyjnego podtypu raka piersi. Można spekulować, że obniżenie poziomu ekspresji retrogenu *HSPA2* może być w przypadku tych dwóch nowotworów wskaźnikiem ich zaawansowania. Klimczak i in. wykazali supresyjną rolę *HSPA2*, korelując jego obniżoną ekspresję z większymi i bardziej zaawansowanymi guzami piersi (Klimczak i in., 2019), co jest zgodne z otrzymanymi w tej pracy wynikami. Podobnie jak *RHOB*, *HSPA2* może odgrywać różne role w różnych typach nowotworów. W przeprowadzonych analizach *HSPA2* wykazał zwiększony poziom ekspresji w raku piersi w danych z repozytorium TCGA. W literaturze można znaleźć informacje, że jego podwyższona ekspresja została powiązana ze złym rokowaniem w przypadku raka płaskonabłonkowego przełyku (H. Zhang i in., 2013) oraz

wczesnym pooperacyjnym nawrotem raka trzustki (Zhai i in., 2021). Kompleksy HSPA2-BAG3 opisano jako powszechne regulatory sygnalizacji w komórkach nowotworowych i potencjalne cele terapeutyczne w niedrobnokomórkowym raku płuc (Sojka i in., 2019). Wyciszenie ekspresji *HSPA2* powiązano ze zmniejszoną proliferacją komórek nowotworowych (Cao i in., 2019). Dane RNA-Seq z tych rodzajów nowotworów są dostępne w repozytorium TCGA, lecz zawierają dane ekspresji, które nie obejmują wszystkich genów. Z reguły w tego typu danych nie znajdziemy informacji o ekspresji większości genów niekodujących białek. W literaturze nie opisano dotychczas związku *HSPA2* z nowotworami psa i kury. Jednym z istotnych odkryć analiz przeprowadzonych w ramach niniejszej pracy jest właśnie wykazanie tego związku.

Białko *CALML5* odgrywa istotną rolę w różnicowaniu komórek naskórka i przekazywaniu sygnału w komórce wiążąc jony wapnia. Stwierdzono, że podczas kancerogenezy bierze udział w proliferacji, różnicowaniu i migracji komórek nowotworowych (Villalobo i in., 2019). Przeprowadzone analizy wykazały, że ekspresja *CALML5* wzrasta w trzech nowotworach człowieka, w tym w bardziej agresywnym raku piersi TNBC. Uzyskane wyniki są zgodne z doniesieniami literaturowymi. Wysoka ekspresja *CALML5* została powiązana także ze złym rokowaniem w przypadku przerzutującego czerniaka skóry (Jia i in., 2021). Co więcej, zwiększenie ekspresji *CALML5* powiązano z inwazyjnością guzów piersi, dokładniej z naciekiem do naczyń limfatycznych (Kurozumi i in., 2019). Podkreślono potencjalną rolę *CALML5* jako markera diagnostycznego dysplazji różnicującej, a także przedinwazyjnych lub inwazyjnych nowotworów szyjki macicy (Kitazawa i in., 2021). Debald i in. wskazali, że ubiquitynacja k63 w *CALML5* jest powiązana z rakiem piersi u kobiet przed menopauzą (Debald i in., 2013). Interesującym jest, że choć *CALML5* ulega ekspresji głównie w skórze, jego aktywność wykazano w nowotworach wielu narządów, w których w normalnych warunkach nie ulega ekspresji. Brak ekspresji w tkankach pochodzących z niektórych z tych narządów potwierdziły wykonane w ramach pracy analizy bibliotek cDNA. Przeprowadzone w ramach tej pracy badania pokazały, że *CALML5* ma zmniejszony poziom ekspresji w nowotworze gruczołu mlekowego psa, co wykazano po raz pierwszy. Kierunek zmiany jest zatem odwrotny niż w przypadku człowieka. Nie jest to jedyna zaobserwowana różnica. W przeciwieństwie do człowieka, gdzie ekspresja *CALML5* nie została wykryta w którejkolwiek z badanych bibliotek cDNA pochodzących z różnych zdrowych narządów, u psa ulegał on ekspresji we wszystkich badanych zdrowych tkankach. Podobnie jak w przypadku profilu ekspresji, duże różnice widoczne są także w sekwencji retrogenu. Identyfikacja pomiędzy

białkiem CALML5 człowieka i psa wynosi zaledwie 57%. Może to wskazywać na dość zróżnicowane funkcje.

Na podstawie analiz bioinformatycznych wskazano specyficzne dla naszego gatunku retrogeny, które zmieniały swój poziom ekspresji w nowotworach. Grupa ta liczy 25 retrogenów, odpowiednio 2 w wyniku analiz danych z repozytorium NCBI SRA, 20 w wyniku analiz danych z bazy TCGA oraz 3, które zidentyfikowano w obydwóch analizach (*retro_hsap_1605*, *retro_hsap_852*, *RPL21P16*). Najciekawsze przypadki obejmują retrogeny, których do tej pory nie opisano jako powiązane z nowotworami i które zmieniają poziom ekspresji w co najmniej dwóch rodzajach nowotworów. Zmiana poziomu ekspresji w kilku różnych nowotworach może wskazywać, że wynik nie jest przypadkowy i retrogen jest istotnie powiązany z procesami nowotworowymi. Do retrogenów takich należą, między innymi, trzy specyficzne dla człowieka retrokopie, które zidentyfikowano na początkowych etapach badań (dane NCBI SRA): *retro_hsap_852*, *retro_hsap_1605* oraz *HMGB1P6*. Retrogen *retro_hsap_852* – duplikat genu rodzicielskiego *RPL36A* oraz *retro_hsap_1605* – retrokopia genu *RPL9* (ang. *ribosomal protein L9*), to przykłady retrogenów, które są kopiami genów kodujących białka rybosomalne. Co ciekawe, u człowieka aż 1340 retrokopii powstało na drodze retropozycji genów kodujących białka rybosomalne (Kozłowska-Masłoń i in., 2024). Związek między zmienioną biogenezą rybosomów a procesem nowotworzenia został już wcześniej opisany. Wykazano, że inhibicja biogenezy rybosomów prowadzi do blokady proliferacji komórek nowotworowych (Penzo i in., 2019). Wyciszenie genu rodzicielskiego *RPL9* powiązano z zahamowaniem wzrostu komórek raka jelita grubego (Baik i in., 2016). W badanych nowotworach retrogen *retro_hsap_1605* pochodzący od *RPL9* wykazał zwiększoną ekspresję w raku wątrobowokomórkowym oraz raku jasnokomórkowym nerki. Retrogen ten zarówno w tkankach nowotworowych jak i zdrowych wykazał bardzo silną pozytywną korelację ekspresji ze swoim genem rodzicielskim. Jego zwiększona ekspresja oznacza zatem także zwiększony poziom ekspresji genu *RPL9*, a to z kolei może być powiązane ze zwiększoną proliferacją i rozwojem nowotworu.

Nadekspresję genu *RPL36A*, który dał początek retrogenowi *retro_hsap_852* skorelowano z proliferacją komórek raka wątroby (J.-H. Kim i in., 2004), (J.-H. Kim i in., 2004). W przeprowadzonych badaniach retrogen wykazywał zwiększoną ekspresję w 8 nowotworach, w tym także w raku wątrobowokomórkowym. Zmiana ekspresji retrogenu *retro_hsap_852* jest zatem podobna do zmiany jaką zaobserwowano w przypadku jego genu rodzicielskiego. Mając dodatkowo na uwadze, że wyciszenie *retro_hsap_852* spowodowało

istotny spadek żywotności komórek oraz fakt, że jego ekspresja jest zwiększona łożysku inwazyjnym można przypuszczać, że retrogen ten odgrywa istotną rolę w rozwoju nowotworów. Podobnie jak w przypadku *retro_hsap_1605*, analizy bioinformatyczne wykazały silną pozytywną korelację ekspresji *retro_hsap_852* z genem rodzicielskim. Jest to zgodne z obserwacją, że ekspresja *RPL36A* i jego retrogeny ulega zwiększeniu w raku wątrobowokomórkowym. Intrygującym jest zatem fakt, że wyciszenie retrogeny spowodowało wzrost poziomu ekspresji genu rodzicielskiego. Skutek był zatem odwrotny od tego, jakiego spodziewano się na podstawie analiz bioinformatycznych. Jedną z przyczyn mogą być zmiany w procesach biologicznych będące następstwem hodowli w warunkach laboratoryjnych materiału pobranego od donora wiele lat temu. Pełne wyjaśnienie będzie wymagało dalszych badań, w tym powtórzenia eksperymentu w tej i innych liniach komórkowych. Niemniej retrogen ten jest zdecydowanie jednym z mocnych kandydatów do dalszych badań, szczególnie, że o retrogenie tym i jego funkcji do tej pory praktycznie niewiele nie było wiadomo.

Trzeci retrogen specyficzny dla naszego gatunku *HMGBIP6* stanowi kopię genu *HMGB1*, opisanego w wielu pracach jako związanego z nowotworami. W raku piersi, opisano go jako supresor nowotworu (Martinotti i in., 2015) oraz jako gen związany z odpowiedzią na radioterapię (He et al. 2013). Co ciekawe, białko HMGB1 jest zaangażowane w metabolizm leków przeciwnowotworowych zawierających platynę (Whirl-Carrillo i in., 2021). W analizowanych danych *HMGBIP6* wykazał zwiększony poziom ekspresji w próbach z przerzutów gruczolaka jelita grubego oraz w raku piersi. Wyniki analiz bioinformatycznych wskazują na korelację ekspresji retrogeny z poziomem ekspresji genu rodzicielskiego. Jego wyciszenie w linii nowotworowej MCF-7 spowodowało spadek poziomu ekspresji genu rodzicielskiego. *HMGBIP6* jest zagnieżdżony w intronie genu *THSD4*, którego zmienioną ekspresję powiązano z występowaniem raka prostaty (Wu i in., 2020) oraz z inwazyjnym rakiem piersi (The Human Protein Atlas). Podobnie jak w przypadku genu rodzicielskiego, wyciszenie *HMGBIP6* spowodowało spadek poziomu ekspresji genu *THSD4*.

Przeprowadzone analizy danych RNA-Seq pozwoliły także na zidentyfikowanie retrogenów związanych z nowotworami, które są specyficznie dla genomu psa lub kury. Jednym z nich jest retrogen psa *retro_cfam_320*, który posiada adnotację jedynie w bazie RetrogeneDB2 (Rosikiewicz i in., 2017). Jak dotąd, nie opisano roli tego retrogeny ani jego genu rodzicielskiego *ENO2* (ang. *enolase 2*) w procesach nowotworzenia u psa. Zwiększony poziom ekspresji *ENO2* stwierdzono w nowotworze płuc człowieka i powiązano z niższą przeżywalnością pacjentów (D. Liu i in., 2020). Ponadto, zwiększony poziom ekspresji tego

genu skorelowano z promowaniem wzrostu komórek nowotworowych, glikolizy oraz oporności na glikokortykoidy, wskazując *ENO2* jako marker monitorujący skuteczność chemioterapii i nawrotów w przypadku ostrej białaczki limfoblastycznej (C.-C. Liu i in., 2018).

Na podstawie danych zawartych w bazie RetrogeneDB2 można zauważyć istotną różnicę w nasileniu procesu retropozycji u analizowanych gatunków (od 4611 retrokopii w genomie człowieka do 25 retrokopii w genomie kury) (Rosikiewicz i in., 2017). Biorąc pod uwagę, że kura posiada tylko 25 retrokopii, identyfikacja pięciu z nich jako związanych z nowotworami, na podstawie analizy tylko jednego typu nowotworu jest dość zaskakująca. Retrogeny te zostały po raz pierwszy wskazane jako potencjalnie powiązane z nowotworami u kury. Jeden z nich to *retro_ggal_15*, retrogen specyficzny dla kury, który powstał jako retrokopia *C12orf45*. Adnotacje tego genu u kury nie zawierają informacji o jego funkcji, ale najprawdopodobniej tak jak w przypadku człowieka i innych eukariontów, ten ewolucyjnie bardzo stary gen koduje białko NOPCHAP1 (ang. *NOP protein chaperon 1*). Jak dotąd nie powiązano *C12orf45*, ani tym bardziej jego retrogeny, z występowaniem nowotworów.

Wyniki przeprowadzonych analiz pozwoliły na identyfikację retrogenów o zmienionym w nowotworach poziomie ekspresji. W przypadku większości z nich to pierwsze takie doniesienia. A fakt, że aż pięć z zaledwie 25 retrogenów kury wykazały zmiany w raku wątrobowokomórkowym może wskazywać, że rola retrokopii w funkcjonowaniu komórek może być dużo bardziej istotna niż przewidywano.

Retrogeny regulujące ekspresję genów związanych z procesem nowotworzenia

Wiele retrogenów o zmienionych w nowotworach poziomach ekspresji reprezentuje ważne niekodujące RNA, które mogą regulować geny związane z nowotworami. Wysoki wskaźnik nowotworzenia u ludzi może wynikać z dużej liczby takich retrokopii. Uzyskane wyniki dostarczają nowych danych odnośnie ich potencjalnej roli regulatorowej. Analizy korelacji ekspresji wykazały, że 35 genów rodzielskich związanych z nowotworami może być regulowanych poprzez swoje retrokopie. W większości tych przypadków retrogeny zachowały w swojej sekwencji przynajmniej część odziedziczonych miejsc docelowych wiązania miRNA. W związku z tym jest całkiem prawdopodobnym, że retrogeny te działają jako gąbki miRNA.

Retrogeny znajdujące się w intronie innego genu mogą regulować jego ekspresję jako *cis*-NAT. Może to się odbywać na przykład poprzez interferencję transkrypcji (Rosikiewicz & Makułowska, 2016). Analizy umożliwiły wskazanie 8 retrogenów wykazujących korelację ekspresji ze swoim genem gospodarzem, spośród których 3 leżą na przeciwnej nici DNA. Należą do nich pary: *RPS11P5-CHFR*, *RPL13AP17-MAGI2*, a także *RPL4P4-MAP3K13*.

W ramach pracy postawiono hipotezę, zakładającą, że wśród retrogenów wykazujących powiązania z nowotworami przeważającą grupą są retrogeny powstałe z genów kodujących istotne dla procesów nowotworowych białka lub ulokowane w intronach takich genów. Wyniki testu statystycznego nie potwierdziły jednak powyższej hipotezy, gdyż retrogeny takie stanowią istotnie mniejszą grupę w porównaniu z retrogenami pochodzącymi od innych genów.

W celu potwierdzenia ewentualnej funkcji regulatorowej retrogenów wyciszono ekspresję trzech wybranych retrogenów w ludzkich liniach komórkowych, a następnie sprawdzono wpływ wyciszenia na ekspresję genu rodzicielskiego lub genu gospodarza. Dodatkowo sprawdzono wpływ wyciszenia na żywotność komórek. Spośród grupy retrogenów o zmienionym poziomie ekspresji w nowotworach wyciszono ekspresję *HMGB1P6*, *retro_hsap_852* oraz *CALML5*. Jak już opisano powyżej, jedynie w przypadku retrogenu *HMGB1P6* wyciszenie wpłynęło negatywnie na poziom ekspresji genu rodzicielskiego *HMGB1* oraz genu gospodarza *THSD4*. Wyciszenie retrogenu *retro_hsap_852* w linii komórkowej MCF-12F spowodowało zwiększoną ekspresję genu rodzicielskiego *RPL36A*, ale nie miało wpływu na ekspresję genu gospodarza *PLEKHA7*. Wyciszenie *retro_hsap_852*, jako jedyne z badanych, wpłynęło na znaczne obniżenie żywotności komórek MCF-12F. Wyciszenie genu *CALML5* nie skutkowało ani zmianą poziomu ekspresji genu rodzicielskiego ani też zmianą stopnia żywotności komórek.

Retrogeny związane z ewolucją w obrębie guza

Wyniki analiz danych pochodzących z bazy NCBI SRA wskazały 10 retrogenów o zmienionej ekspresji w próbach pochodzących z bardziej agresywnego typu raka piersi (TNBC) oraz 2 retrogeny wśród prób z bardziej inwazyjnego subtypu molekularnego wątrobiaka zarodkowego (C2). Ponadto w porównaniu ekspresji pomiędzy przerzutami a próbami zdrowymi gruczolaka jelita grubego zidentyfikowano 8 retrogenów, które zmieniły ekspresję w tej inwazyjnej formie nowotworu. W przypadku danych pochodzących z projektu TCGA analiza ekspresji różnicowej pomiędzy nowotworami w stadium IV a I pozwoliła na zidentyfikowanie *RHOB*, *PPP1R14BP3* oraz *retro_hsap_2820* jako retrogenów o zmienionym poziomie ekspresji podczas progresji nowotworu.

Interesujący przypadek spośród tych retrogenów stanowi retrogen *PPP1R14BP3*, którego zwiększona ekspresja występowała w 6 analizowanych nowotworach (BRCA, KIRC, LIHC, LUSC, PRAD, UCEC), w różnych stadiach raka jasnokomórkowego nerki oraz w różnych rodzajach odpowiedzi na terapię. Dotychczas nie powiązano retrogenu *PPP1R14BP3*

z występowaniem chorób człowieka. Natomiast gen rodzicielski *PPP1R14B* wskazano w literaturze jako marker diagnostyczny u pacjentek z rakiem endometrium trzonu macicy (He i in., 2023). Ponadto opisano rolę tego genu w progresji raka TNBC oraz występowaniu oporności na leczenie paklitakselem, wskazując *PPP1R14B* jako cel terapeutyczny w leczeniu tego rodzaju nowotworu piersi (Liao i in., 2023).

Identyfikacja retrogenów o zmienionym poziomie ekspresji w nowotworach i jednocześnie związanych z rozwojem łożyska człowieka

Nowotwory są w stanie doprowadzić do śmierci organizmu, między innymi dlatego, że wykorzystują mechanizmy oszukiwania naszego układu odpornościowego, co pozwala im rosnąć i rozprzestrzeniać się. Są one jedną z głównych przyczyn zgonów na świecie, a mechanizmy ewolucyjne prowadzące do podatności na nowotwory mogą być powiązane z pojawianiem się cech adaptacyjnych u poszczególnych gromad kręgowców. Jedną z nich jest występowanie łożyska u ssaków łożyskowych. Występują bardzo duże podobieństwa pomiędzy procesem inwazji łożyska a inwazją komórek nowotworowych i sugeruje się, że gatunki charakteryzujące się bardziej inwazyjnym typem łożyska (np. człowiek) są bardziej podatne na nowotwory złośliwe (Kshitiz i in., 2019). Przypuszcza się, że ewolucja genów związanych z występowaniem nowotworów złośliwych nie była ukierunkowana na zapobieganie lub powodowanie nowotworów, lecz na umożliwienie tolerancji immunologicznej między płodem a matką, co jest kluczowe dla natury ssaków łożyskowych. Agresywne nowotwory nabłonkowe mogą być zatem niefortunnym "efektem ubocznym", który nie był poddany silnej presji ewolucyjnej, ponieważ średnia długość życia człowieka jeszcze niedawno wynosiła 45-50 lat, co było wystarczające dla reprodukcji (Bronchud, 2018).

Jednym z ciekawszych wyników prezentowanych badań jest zidentyfikowanie grupy 14 retrogenów o zmienionej ekspresji w nowotworach, które zostały także wskazane jako retrokopie o zróżnicowanej ekspresji podczas rozwoju łożyska człowieka. Na szczególną uwagę zasługują *retro_hsap_852* oraz *RPS2P5*. Inne przeprowadzone analizy wskazywały już *retro_hsap_852* jako retrogen o istotnej roli w procesach nowotworowych. Ten specyficzny dla człowieka retrogen wykazał zwiększony poziom ekspresji w 8 rodzajach nowotworów i okazuje się być także powiązany z procesami rozwoju łożyska. Drugi retrogen, *RPS2P5*, oprócz zwiększonego poziomu ekspresji w procesie rozwoju łożyska wykazał także zwiększoną ekspresję w 7 nowotworach człowieka. Dane przedkliniczne na myszach, wskazały gen rodzicielski *RPS2* jako obiecujący cel terapeutyczny w przypadku raka prostaty (M. Wang i in., 2009). Nie można wykluczyć, że pochodzący od niego *RPS2P5* jest także ważnym

graczem w rozwoju nowotworów, być może szczególnie w przypadku tych złośliwych. Dlatego, podobnie jak *retro_hsap_852* może być potencjalnym biomarkerem nabywania cech złośliwości lub też celem terapeutycznym.

Podsumowanie

Wyniki uzyskane w ramach niniejszej pracy umożliwiły po raz pierwszy w tak dużej skali spojrzeć na profil ekspresji retrogenów w nowotworach. Biorąc pod uwagę rozwijający się obszar jakim jest medycyna ewolucyjna, analiza retrogenów wydaje się obiecującym kierunkiem badań. Dotychczas otrzymane wyniki wskazują na ważne w procesach nowotworowych retrokopie i interesujących kandydatów na biomarkery. Dalsze badania tych retrogenów mogą dostarczyć dodatkowych danych i przyczynić się do zwiększenia efektywności diagnostyki oraz leczenia poszczególnych nowotworów, a tym samym lepszego rokowania pacjentów onkologicznych. Ponadto, ewolucyjne spojrzenie na proces nowotworzenia może dostarczyć informacji o cechach charakterystycznych dla człowieka i kierunku ewolucji naszego gatunku. Informacje na temat ekspresji retrogenów mogą być wykorzystane w przypadku badań przedklinicznych w celu lepszego zrozumienia różnic między organizmem człowieka a organizmami modelowymi. Mogą także pomóc w wyjaśnieniu częstych problemów związanych z przenoszeniem badań na organizmach modelowych np. na myszach, na człowieka. Niniejsza praca dostarczyła dowodów na to, że geny pochodzące z retropozycji mogą być obiecującymi kandydatami do badań nad procesami nowotworzenia.

6. WNIOSKI

Przedstawiona praca dostarczyła wielu nowych i interesujących wyników dotyczących retrogenów i ich potencjalnych powiązań z nowotworami. Poniżej wymienione są najważniejsze wnioski i wyniki tych przeprowadzanych w kontekście ewolucyjnym badań:

✓ Zarówno młode, jak i stare ewolucyjnie retrogeny są powiązane z procesem nowotworzenia. Niemniej jednak, w przypadku człowieka większość spośród nich to retrokopie powstałe podczas ewolucji naczelnych.

✓ *HSPA2*, *RHOB* i *CALML5* to retrogeny po raz pierwszy wskazane jako powiązane z nowotworami odległych ewolucyjnie gatunków kręgowców.

✓ *PPP1R14BP3*, *RHOB* i *retro_hsap_2820* to retrogeny, które istotnie zmieniają poziom ekspresji podczas rozwoju guza nowotworowego.

✓ Retrogeny *PPP1R14BP3*, *RPL10P9* oraz *RPS28P7* wykazały istotną zmianę ekspresji w odpowiedzi na zastosowane leczenie onkologiczne.

✓ Kilkadziesiąt retrogenów wykazujących zmiany ekspresji w nowotworach, może działać jako ważne niekodujące RNA regulujące geny znane już jako pełniące role w rozwoju nowotworów.

✓ Jak wskazuje przeprowadzony eksperyment, wyciszenie ekspresji *retro_hsap_852* wpływa na zmianę żywotności komórek MCF-12F.

✓ Poziom ekspresji retrogenu *HMGB1P6* ma wpływ na ekspresję genu rodzicielskiego *HMGB1*, który jest zaangażowany w metabolizm leków onkologicznych opartych na platynie.

✓ 14 retrogenów wykazujących zmianę poziomu ekspresji w nowotworach wykazuje podobne zmiany w trakcie rozwoju łożyska człowieka i może być z procesem tym powiązanych.

✓ Retrogen *retro_hsap_852* z dużym prawdopodobieństwem ma istotne znaczenie w procesach nowotworowych człowieka i jest najbardziej obiecującym kandydatem do dalszych badań.

SPIS RYCIN

Rycina 1. Mechanizm powstawania retrokopii genów kodujących białka (a) oraz możliwe funkcje pełnione w komórce (b)	8
Rycina 2. Schematyczne ujęcie związków pomiędzy retrogenami, procesem nowotworzenia a ewolucją	12
Rycina 3. Mechanizmy ewolucyjne a nowotwory.	14
Rycina 4. Inwazja łożyska a proces inwazji komórek nowotworowych.	16
Rycina 5. Retrogeny związane z procesem nowotworzenia jako gąbki miRNA i <i>cis</i> -NATs w komórce zdrowej i nowotworowej.	19
Rycina 6. Kontrola jakości odczytów pochodzących z próby SRR3129838 (rak wątrobowokomórkowy) przed (a) i po oczyszczeniu z adapterów, odczytów niskiej jakości oraz rRNA (b)	26
Rycina 7. Ekspresja retrokopii człowieka w analizowanych próbach (NCBI SRA)	33
Rycina 8. Ekspresja retrokopii psa w analizowanych próbach (NCBI SRA)	34
Rycina 9. Ekspresja 11 retrokopii kury w 12 analizowanych bibliotekach (NCBI SRA).	35
Rycina 10. Przykład identyfikacji prób odstających na podstawie analizy głównych składowych dla gruczolakoraka płuc PRJEB2784..	36
Rycina 11. Czas powstania retrogenów o zmienionej ekspresji w nowotworach człowieka.	42
Rycina 12. Porównanie struktury i lokalizacji retrokopii <i>RHOB</i> (a) <i>HSPA2</i> (b) i <i>CALML5</i> (c) z ich genami rodzicielskimi.	44
Rycina 13. Wykresy pudełkowe dla trzech ewolucyjnie starych retrogenów, wspólnych dla nowotworów człowieka, psa i kury – <i>CALML5</i> , <i>HSPA2</i> oraz <i>RHOB</i>	47
Rycina 14. Wyniki analiz eksperymentalnych retrogenu <i>CALML5</i>	49
Rycina 15. Wyniki analiz eksperymentalnych retrogenu <i>HSPA2</i>	50
Rycina 16. Wyniki analiz eksperymentalnych retrogenu <i>RHOB</i>	51
Rycina 17. Retrogeny specyficzne dla badanych gatunków kręgowców.	52
Rycina 18. Wyniki analiz eksperymentalnych retrogenów specyficznych dla człowieka, zidentyfikowanych jako retrogeny o zmienionej ekspresji w więcej niż jednym nowotworze człowieka.	54
Rycina 19. Wyniki analiz eksperymentalnych retrogenów specyficznych dla psa lub kury, które wykazywały zmieniony poziom ekspresji w nowotworach.	55
Rycina 20. Profile ekspresji wybranych genów a zdolność różnicowania rodzajów nowotworów.	57
Rycina 21. Ekspresja retrokopii w analizowanych 5250 próbach TCGA..	58
Rycina 22. Ekspresja retrogenów o zmienionej ekspresji w 8 nowotworach człowieka (dane TCGA).	64
Rycina 23. Czas powstania retrogenów o zmienionej ekspresji w nowotworach człowieka (dane TCGA).	65
Rycina 24. Wykresy pudełkowe dla wybranych retrogenów związanych z ewolucją w obrębie guza nowotworowego lub odpowiedzią na terapię.	67
Rycina 25. Retrogeny, które potencjalnie mogą regulować ekspresję genów rodzicielskich powiązanych z nowotworami.	71

Rycina 26. Retrogeny jako regulatory genów gospodarzy, które powiązano z nowotworami.	75
Rycina 27. Wpływ wyciszenia ekspresji retrogeny <i>CALML5</i> w linii nowotworowej MCF-7 na ekspresję genu rodzicielskiego <i>CALM3</i>	77
Rycina 28. Wpływ wyciszenia ekspresji retrokopii <i>HMGB1P6</i> w linii nowotworowej MCF-7 na ekspresję genu rodzicielskiego <i>HMGB1</i> (a) oraz na ekspresję genu gospodarza <i>THSD4</i> (b)	78
Rycina 29. Wpływ wyciszenia ekspresji retrogeny <i>retro_hsap_852</i> w linii nienowotworowej MCF-12F na ekspresję genu rodzicielskiego <i>RPL36A</i> (a) oraz na ekspresję genu gospodarza <i>PLEKHA7</i> (b)	80
Rycina 30. Wpływ wyciszenia ekspresji retrogeny <i>CALML5</i> (a) , <i>HMGB1P6</i> (b) oraz <i>retro_hsap_852</i> (c) na żywotność komórek.	81

SPIS TABEL

Tabela 1. Zestawy danych nowotworowych z repozytorium NCBI SRA użyte do analiz bioinformatycznych.....	22
Tabela 2. Zestaw danych z eksperymentu inwazji łożyska z repozytorium NCBI SRA użyty do analiz bioinformatycznych.....	23
Tabela 3. Zestawy danych z repozytorium TCGA użyte do analiz bioinformatycznych.....	23
Tabela 4. Przykładowe podsumowanie przetwarzania danych RNA-Seq z repozytorium NCBI SRA.....	25
Tabela 5. Startery PCR stosowane w analizach ekspresji genów człowieka.....	29
Tabela 6. Startery PCR stosowane w analizach ekspresji genów psa, kury i szczura.....	30
Tabela 7. Sekwencje gapmeRów użyte do wyciszenia ekspresji poszczególnych retrogenów.....	31
Tabela 8. Retrogeny o zmienionym poziomie ekspresji w 5 ludzkich nowotworach (dane NCBI SRA).	37
Tabela 9. Retrogeny o zmienionym poziomie ekspresji w różnych typach lub stadiach 3 ludzkich nowotworów (dane NCBI SRA).....	39
Tabela 10. Retrogeny o zmienionym poziomie ekspresji w nowotworach psa.....	40
Tabela 11. Retrogeny o zmienionym poziomie ekspresji w nowotworze kury.....	41
Tabela 12. Porównanie podobieństwa sekwencji aminokwasowej <i>RHOB</i> , <i>HSPA2</i> oraz <i>CALML5</i> (BLASTP) między białkami człowieka i innych gatunków kręgowców.....	45
Tabela 13. Retrogeny o zmienionym poziomie ekspresji w 8 ludzkich nowotworach (dane TCGA). ..	60
Tabela 14. Retrogeny o zmienionym poziomie ekspresji podczas ewolucji guza nowotworowego (dane TCGA). ..	66
Tabela 15. Retrogeny o zmienionym poziomie ekspresji w odpowiedzi na zastosowaną terapię (dane TCGA). ..	68
Tabela 16. Retrogeny o zróżnicowanej ekspresji wyselekcjonowane na podstawie roli ich genu rodzicielskiego w nowotworach człowieka.....	69
Tabela 17. Potencjalne miejsca wiązania miRNA wspólne dla retrogenów i ich genów rodzicielskich (na podstawie miRDB.org (Y. Chen & Wang, 2020)).....	72
Tabela 18. Retrogeny o zróżnicowanej ekspresji wyselekcjonowane na podstawie roli ich genu gospodarza w nowotworach człowieka.....	74
Tabela 19. Tablica kontyngencji dla testu statystycznego weryfikującego hipotezę 2.....	75
Tabela 20. Retrogeny o zmienionym poziomie ekspresji podczas rozwoju łożyska człowieka oraz o zmienionym poziomie ekspresji w nowotworach człowieka.....	83
Tabela 21. Wyniki analizy ekspresji różnicowej retrogenów dla raka piersi z zestawu danych o numerze PRJNA251383.....	112
Tabela 22. Wyniki analizy ekspresji różnicowej retrogenów dla typów raka piersi – ER+ oraz TNBC z zestawu danych o numerze PRJNA251383.....	113
Tabela 23. Wyniki analizy ekspresji różnicowej retrogenów dla gruczolaka płuc z zestawu danych o numerze PRJEB2784.....	113
Tabela 24. Wyniki analizy ekspresji różnicowej retrogenów dla raka jelita grubego z zestawu danych o numerze PRJNA218851.....	114
Tabela 25. Wyniki analizy ekspresji różnicowej retrogenów dla raka jelita grubego – przerzuty vs guzy pierwotne z zestawu danych o numerze PRJNA218851.....	114

Tabela 26. Wyniki analizy ekspresji różnicowej retrogenów dla raka wątrobowokomórkowego z zestawu danych o numerze PRJNA310012.....	115
Tabela 27. Wyniki analizy ekspresji różnicowej retrogenów dla wątrobiaka zarodkowego z zestawu danych o numerze PRJNA413799.....	115
Tabela 28. Wyniki analizy ekspresji różnicowej retrogenów dla subtypów molekularnych wątrobiaka zarodkowego – C1 oraz C2 (PRJNA413799).....	116
Tabela 29. Wyniki analizy ekspresji różnicowej retrogenów dla raka gruczołu mlekowego psa z zestawu danych o numerze PRJNA489087.....	116
Tabela 30. Wyniki analizy ekspresji różnicowej retrogenów dla raka gruczołu mlekowego psa z zestawu danych o numerze PRJNA561580.....	117
Tabela 31. Wyniki analizy ekspresji różnicowej retrogenów dla linii komórkowej raka wątrobowokomórkowego kury (PRJNA239321) oraz próbek zdrowych wątroby kury (PRJNA362606).	117
Tabela 32. Wyniki analizy ekspresji różnicowej retrogenów dla raka piersi (TCGA-BRCA).	117
Tabela 33. Wyniki analizy ekspresji różnicowej retrogenów dla gruczolaka jelita grubego (TCGA-COAD).....	118
Tabela 34. Wyniki analizy ekspresji różnicowej retrogenów dla raka jasnokomórkowego nerki (TCGA-KIRC).....	120
Tabela 35. Wyniki analizy ekspresji różnicowej retrogenów dla raka wątrobowokomórkowego (TCGA-LIHC).....	122
Tabela 36. Wyniki analizy ekspresji różnicowej retrogenów dla raka płaskonabłonkowy płuc (TCGA-LUSC).....	122
Tabela 37. Wyniki analizy ekspresji różnicowej retrogenów dla gruczolaka prostaty (TCGA-PRAD).	124
Tabela 38. Wyniki analizy ekspresji różnicowej retrogenów dla gruczolaka odbytnicy (TCGA-READ).	125
Tabela 39. Wyniki analizy ekspresji różnicowej retrogenów dla raka endometrium trzonu macicy (TCGA-UCEC).	127
Tabela 40. Wyniki analizy ekspresji różnicowej retrogenów podczas rozwoju guza (stadium IV vs stadium I) (TCGA-KIRC, TCGA-LIHC).	128
Tabela 41. Wyniki analizy ekspresji różnicowej retrogenów pomiędzy próbami o odmiennej odpowiedzi na zastosowane leczenie (postępujący nowotwór vs remisja) (TCGA-LUSC, TCGA-OV, TCGA-UCEC).	129
Tabela 42. Korelacja ekspresji retrogenów i genów rodzicielskich w badanych nowotworach oraz tkankach okalających.	130
Tabela 43. Korelacja ekspresji retrogenów i genów gospodarzy w badanych nowotworach oraz tkankach okalających.....	135
Tabela 44. Wyniki analizy ekspresji różnicowej retrogenów dla eksperymentu inwazji łożyska (PRJNA562073).....	136

LITERATURA

- Abegglen, L. M., Caulin, A. F., Chan, A., Lee, K., Robinson, R., Campbell, M. S., Kiso, W. K., Schmitt, D. L., Waddell, P. J., Bhaskara, S., Jensen, S. T., Maley, C. C., & Schiffman, J. D. (2015). Potential Mechanisms for Cancer Resistance in Elephants and Comparative Cellular Response to DNA Damage in Humans. *JAMA*, 314(17), 1850–1860. <https://doi.org/10.1001/jama.2015.13134>
- Ahlmann-Eltze, C., & Patil, I. (2021). ggsignif: R Package for Displaying Significance Brackets for „ggplot2”. OSF. <https://doi.org/10.31234/osf.io/7awm6>
- Avivi, A., Brodsky, L., Nevo, E., & Band, M. R. (2006). Differential expression profiling of the blind subterranean mole rat *Spalax ehrenbergi* superspecies: Bioprospecting for hypoxia tolerance. *Physiological Genomics*, 27(1), 54–64. <https://doi.org/10.1152/physiolgenomics.00001.2006>
- Baik, I. H., Jo, G.-H., Seo, D., Ko, M. J., Cho, C. H., Lee, M. G., & Lee, Y.-H. (2016). Knockdown of RPL9 expression inhibits colorectal carcinoma growth via the inactivation of Id-1/NF- κ B signaling axis. *International Journal of Oncology*, 49(5), 1953–1962. <https://doi.org/10.3892/ijo.2016.3688>
- Bigge, J., Koebbe, L. L., Giel, A.-S., Bornholdt, D., Buerfent, B., Dasmeh, P., Zink, A. M., Maj, C., & Schumacher, J. (2024). Expression quantitative trait loci influence DNA damage-induced apoptosis in cancer. *BMC Genomics*, 25(1), 1168. <https://doi.org/10.1186/s12864-024-11068-6>
- Bronchud, M. H. (2018). Are aggressive epithelial cancers ‘a disease’ of Eutherian mammals? *ecancermedicalsecience*, 12, 840. <https://doi.org/10.3332/ecancer.2018.840>
- Brosius, J. (1991). Retroposons—Seeds of evolution. *Science*, 251(4995), 753–753. <https://doi.org/10.1126/science.1990437>
- Bryzghalov, O., Szcześniak, M. W., & Makołowska, I. (2016). Retroposition as a source of antisense long non-coding RNAs with possible regulatory functions. *Acta Biochimica Polonica*, 63(4), 825–833. https://doi.org/10.18388/abp.2016_1354
- Cai, J. J., & Petrov, D. A. (2010). Relaxed purifying selection and possibly high rate of adaptation in primate lineage-specific genes. *Genome Biology and Evolution*, 2, 393–409. <https://doi.org/10.1093/gbe/evq019>
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10, 421. <https://doi.org/10.1186/1471-2105-10-421>
- Cao, L., Yuan, X., Bao, F., Lv, W., He, Z., Tang, J., Han, J., & Hu, J. (2019). Downregulation of HSPA2 inhibits proliferation via ERK1/2 pathway and endoplasmic reticular stress in lung adenocarcinoma. *Annals of Translational Medicine*, 7(20), 540. <https://doi.org/10.21037/atm.2019.10.16>
- Carelli, F. N., Hayakawa, T., Go, Y., Imai, H., Warnefors, M., & Kaessmann, H. (2016). The life history of retrocopies illuminates the evolution of new mammalian genes. *Genome Research*, 26(3), 301–314. <https://doi.org/10.1101/gr.198473.115>
- Chen, C.-L., Tseng, Y.-W., Wu, J.-C., Chen, G.-Y., Lin, K.-C., Hwang, S.-M., & Hu, Y.-C. (2015). Suppression of hepatocellular carcinoma by baculovirus-mediated expression of long non-coding RNA PTENP1 and MicroRNA regulation. *Biomaterials*, 44, 71–81. <https://doi.org/10.1016/j.biomaterials.2014.12.023>
- Chen, S., Zhang, Y. E., & Long, M. (2010). New genes in *Drosophila* quickly become essential. *Science* (New York, N.Y.), 330(6011), 1682–1685. <https://doi.org/10.1126/science.1196380>
- Chen, W., Niu, S., Ma, X., Zhang, P., Gao, Y., Fan, Y., Pang, H., Gong, H., Shen, D., Gu, L., Zhang, Y., & Zhang, X. (2016). RhoB Acts as a Tumor Suppressor That Inhibits Malignancy of Clear Cell

- Renal Cell Carcinoma. *PloS One*, 11(7), e0157599. <https://doi.org/10.1371/journal.pone.0157599>
- Chen, X., Zhu, H., Wu, X., Xie, X., Huang, G., Xu, X., Li, S., & Xing, C. (2016). Downregulated pseudogene CTNNAP1 promote tumor growth in human cancer by downregulating its cognate gene CTNNA1 expression. *Oncotarget*, 7(34), 55518–55528. <https://doi.org/10.18632/oncotarget.10833>
- Chen, Y., & Wang, X. (2020). miRDB: An online database for prediction of functional microRNA targets. *Nucleic Acids Research*, 48(D1), D127–D131. <https://doi.org/10.1093/nar/gkz757>
- Ciomborowska, J., Rosikiewicz, W., Szklarczyk, D., Makiłowski, W., & Makiłowska, I. (2013). „Orphan” retrogenes in the human genome. *Molecular Biology and Evolution*, 30(2), 384–396. <https://doi.org/10.1093/molbev/mss235>
- Ciomborowska-Basheer, J., Staszak, K., Kubiak, M. R., & Makiłowska, I. (2021). Not So Dead Genes—Retrocopies as Regulators of Their Disease-Related Progenitors and Hosts. *Cells*, 10(4), Article 4. <https://doi.org/10.3390/cells10040912>
- Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T. S., Malta, T. M., Pagnotta, S. M., Castiglioni, I., Ceccarelli, M., Bontempi, G., & Noushmehr, H. (2016). TCGAAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Research*, 44(8), e71. <https://doi.org/10.1093/nar/gkv1507>
- Connolly, E. C., Van Doorslaer, K., Rogler, L. E., & Rogler, C. E. (2010). Overexpression of miR-21 promotes an in vitro metastatic phenotype by targeting the tumor suppressor RHOB. *Molecular Cancer Research: MCR*, 8(5), 691–700. <https://doi.org/10.1158/1541-7786.MCR-09-0465>
- De Keersmaecker, K., Atak, Z. K., Li, N., Vicente, C., Patchett, S., Girardi, T., Gianfelici, V., Geerdens, E., Clappier, E., Porcu, M., Lahortiga, I., Lucà, R., Yan, J., Hulselmans, G., Vranckx, H., Vandepoel, R., Sweron, B., Jacobs, K., Mentens, N., ... Cools, J. (2013). Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. *Nature Genetics*, 45(2), 186–190. <https://doi.org/10.1038/ng.2508>
- Debald, M., Schildberg, F. A., Linke, A., Walgenbach, K., Kuhn, W., Hartmann, G., & Walgenbach-Brünagel, G. (2013). Specific expression of k63-linked ubiquitination of calmodulin-like protein 5 in breast cancer of premenopausal patients. *Journal of Cancer Research and Clinical Oncology*, 139(12), 2125–2132. <https://doi.org/10.1007/s00432-013-1541-y>
- Du, N., Li, D., Zhao, W., & Liu, Y. (2024). Stratifin (SFN) Regulates Cervical Cancer Cell Proliferation, Apoptosis, and Cytoskeletal Remodeling and Metastasis Progression Through LIMK2/Cofilin Signaling. *Molecular Biotechnology*, 66(11), 3369–3381. <https://doi.org/10.1007/s12033-023-00946-1>
- Efeyan, A., & Serrano, M. (2007). p53: Guardian of the genome and policeman of the oncogenes. *Cell Cycle (Georgetown, Tex.)*, 6(9), 1006–1010. <https://doi.org/10.4161/cc.6.9.4211>
- Enriquez-Navas, P. M., Wojtkowiak, J. W., & Gatenby, R. A. (2015). Application of Evolutionary Principles to Cancer Therapy. *Cancer Research*, 75(22), 4675–4680. <https://doi.org/10.1158/0008-5472.CAN-15-1337>
- Esposito, F., De Martino, M., Petti, M. G., Forzati, F., Tornincasa, M., Federico, A., Arra, C., Pierantoni, G. M., & Fusco, A. (2014). HMGA1 pseudogenes as candidate proto-oncogenic competitive endogenous RNAs. *Oncotarget*, 5(18), 8341–8354.
- Fablet, M., Bueno, M., Potrzebowski, L., & Kaessmann, H. (2009). Evolutionary origin and functions of retrogene introns. *Molecular Biology and Evolution*, 26(9), 2147–2156. <https://doi.org/10.1093/molbev/msp125>
- Fairbanks, D. J., Fairbanks, A. D., Ogden, T. H., Parker, G. J., & Maughan, P. J. (2012). NANOGP8: Evolution of a Human-Specific Retro-Oncogene. *G3: Genes|Genomes|Genetics*, 2(11), 1447–1457. <https://doi.org/10.1534/g3.112.004366>

- Fortunato, A., Boddy, A., Mallo, D., Aktipis, A., Maley, C. C., & Pepper, J. W. (2017). Natural Selection in Cancer Biology: From Molecular Snowflakes to Trait Hallmarks. *Cold Spring Harbor Perspectives in Medicine*, 7(2), a029652. <https://doi.org/10.1101/cshperspect.a029652>
- Ghafouri-Fard, S., Noroozi, R., Abak, A., Taheri, M., & Salimi, A. (2021). Emerging role of lncRNAs in the regulation of Rho GTPase pathway. *Biomedicine & Pharmacotherapy*, 140, 111731. <https://doi.org/10.1016/j.biopha.2021.111731>
- Giacomelli, A. O., Yang, X., Lintner, R. E., McFarland, J. M., Duby, M., Kim, J., Howard, T. P., Takeda, D. Y., Ly, S. H., Kim, E., Gannon, H. S., Hurhula, B., Sharpe, T., Goodale, A., Fritchman, B., Steelman, S., Vazquez, F., Tsherniak, A., Aguirre, A. J., ... Hahn, W. C. (2018). Mutational processes shape the landscape of TP53 mutations in human cancer. *Nature Genetics*, 50(10), 1381–1387. <https://doi.org/10.1038/s41588-018-0204-y>
- Goldman, M. J., Craft, B., Hastie, M., Repčeka, K., McDade, F., Kamath, A., Banerjee, A., Luo, Y., Rogers, D., Brooks, A. N., Zhu, J., & Haussler, D. (2020). Visualizing and interpreting cancer genomics data via the Xena platform. *Nature Biotechnology*, 38(6), Article 6. <https://doi.org/10.1038/s41587-020-0546-8>
- Gorbunova, V., Hine, C., Tian, X., Ablueva, J., Gudkov, A. V., Nevo, E., & Seluanov, A. (2012). Cancer resistance in the blind mole rat is mediated by concerted necrotic cell death mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 109(47), 19392–19396. <https://doi.org/10.1073/pnas.1217211109>
- Graim, K., Gorenshteyn, D., Robinson, D. G., Carriero, N. J., Cahill, J. A., Chakrabarti, R., Goldschmidt, M. H., Durham, A. C., Funk, J., Storey, J. D., Kristensen, V. N., Theesfeld, C. L., Sorenmo, K. U., & Troyanskaya, O. G. (2020). Modeling molecular development of breast cancer in canine mammary tumors. *Genome Research*. <https://doi.org/10.1101/gr.256388.119>
- Grammatikakis, I., & Lal, A. (2022). Significance of lncRNA abundance to function. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, 33(2), 271–280. <https://doi.org/10.1007/s00335-021-09901-4>
- Groen, J. N., Capraro, D., & Morris, K. V. (2014). The emerging role of pseudogene expressed non-coding RNAs in cellular functions. *The International Journal of Biochemistry & Cell Biology*, 54, 350–355. <https://doi.org/10.1016/j.biocel.2014.05.008>
- Grossman, R. L., Heath, A. P., Ferretti, V., Varmus, H. E., Lowy, D. R., Kibbe, W. A., & Staudt, L. M. (2016). Toward a Shared Vision for Cancer Genomic Data. *The New England Journal of Medicine*, 375(12), 1109–1112. <https://doi.org/10.1056/NEJMp1607591>
- GTEx Consortium. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, 45(6), 580–585. <https://doi.org/10.1038/ng.2653>
- Gu, Z. (2022). Complex heatmap visualization. *iMeta*, 1(3), e43. <https://doi.org/10.1002/imt2.43>
- Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, 32(18), 2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>
- Hall, T. A. (1999). BioEdit A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98. - References—Scientific Research Publishing. [https://www.scirp.org/\(S\(351jmbntvnsjtl aadkposzje\)\)/reference/ReferencesPapers.aspx?ReferenceID=1383440](https://www.scirp.org/(S(351jmbntvnsjtl aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=1383440)
- Hawkins, P. G., & Morris, K. V. (2010). Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5. *Transcription*, 1(3), 165–175. <https://doi.org/10.4161/trns.1.3.13332>
- He, K., Wang, T., Huang, X., Yang, Z., Wang, Z., Zhang, S., Sui, X., Jiang, J., & Zhao, L. (2023). PPP1R14B is a diagnostic prognostic marker in patients with uterine corpus endometrial

- carcinoma. *Journal of Cellular and Molecular Medicine*, 27(6), 846–863. <https://doi.org/10.1111/jcmm.17697>
- Herrero, J., Muffato, M., Beal, K., Fitzgerald, S., Gordon, L., Pignatelli, M., Vilella, A. J., Searle, S. M. J., Amode, R., Brent, S., Spooner, W., Kulesha, E., Yates, A., & Flicek, P. (2016). Ensembl comparative genomics resources. *Database: The Journal of Biological Databases and Curation*, 2016, bav096. <https://doi.org/10.1093/database/bav096>
- Hooks, K. B., Audoux, J., Fazli, H., Lesjean, S., Ernault, T., Dugot-Senart, N., Leste-Lasserre, T., Hagedorn, M., Rousseau, B., Danet, C., Branchereau, S., Brugières, L., Taquet, S., Guettier, C., Fabre, M., Rullier, A., Buendia, M.-A., Commes, T., Grosset, C. F., & Raymond, A.-A. (2018). New insights into diagnosis and therapeutic options for proliferative hepatoblastoma. *Hepatology (Baltimore, Md.)*, 68(1), 89–102. <https://doi.org/10.1002/hep.29672>
- Hu, Y., Zeng, Q., Li, C., & Xie, Y. (2019). Expression profile and prognostic value of SFN in human ovarian cancer. *Bioscience Reports*, 39(5), BSR20190100. <https://doi.org/10.1042/BSR20190100>
- Jia, G., Song, Z., Xu, Z., Tao, Y., Wu, Y., & Wan, X. (2021). Screening of gene markers related to the prognosis of metastatic skin cutaneous melanoma based on Logit regression and survival analysis. *BMC Medical Genomics*, 14(1), 96. <https://doi.org/10.1186/s12920-021-00923-0>
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: A better web interface. *Nucleic Acids Research*, 36(Web Server issue), W5–9. <https://doi.org/10.1093/nar/gkn201>
- Ju, J. A., & Gilkes, D. M. (2018). RhoB: Team Oncogene or Team Tumor Suppressor? *Genes*, 9(2), Article 2. <https://doi.org/10.3390/genes9020067>
- Ju, Y. S., Lee, W.-C., Shin, J.-Y., Lee, S., Bleazard, T., Won, J.-K., Kim, Y. T., Kim, J.-I., Kang, J.-H., & Seo, J.-S. (2012a). A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Research*, 22(3), 436–445. <https://doi.org/10.1101/gr.133645.111>
- Ju, Y. S., Lee, W.-C., Shin, J.-Y., Lee, S., Bleazard, T., Won, J.-K., Kim, Y. T., Kim, J.-I., Kang, J.-H., & Seo, J.-S. (2012b). A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Research*, 22(3), 436–445. <https://doi.org/10.1101/gr.133645.111>
- Kaer, K., Branovets, J., Hallikma, A., Nigumann, P., & Speek, M. (2011). Intronic L1 retrotransposons and nested genes cause transcriptional interference by inducing intron retention, exonization and cryptic polyadenylation. *PloS One*, 6(10), e26099. <https://doi.org/10.1371/journal.pone.0026099>
- Kaessmann, H., Vinckenbosch, N., & Long, M. (2009). RNA-based gene duplication: Mechanistic and evolutionary insights. *Nature Reviews. Genetics*, 10(1), 19–31. <https://doi.org/10.1038/nrg2487>
- Kanamori, K., Suina, K., Shukuya, T., Takahashi, F., Hayashi, T., Hara, K., Saito, T., Mitsuishi, Y., Shimamura, S. S., Winardi, W., Tajima, K., Ko, R., Mimori, T., Asao, T., Itoh, M., Kawaji, H., Suehara, Y., Takamochi, K., Suzuki, K., & Takahashi, K. (2023). CALML5 is a novel diagnostic marker for differentiating thymic squamous cell carcinoma from type B3 thymoma. *Thoracic Cancer*, 14(12), 1089–1097. <https://doi.org/10.1111/1759-7714.14853>
- Keane, M., Semeiks, J., Webb, A. E., Li, Y. I., Quesada, V., Craig, T., Madsen, L. B., van Dam, S., Brawand, D., Marques, P. I., Michalak, P., Kang, L., Bhak, J., Yim, H.-S., Grishin, N. V., Nielsen, N. H., Heide-Jørgensen, M. P., Oziolor, E. M., Matson, C. W., ... de Magalhães, J. P. (2015). Insights into the evolution of longevity from the bowhead whale genome. *Cell Reports*, 10(1), 112–122. <https://doi.org/10.1016/j.celrep.2014.12.008>
- Kim, J.-H., You, K.-R., Kim, I. H., Cho, B.-H., Kim, C.-Y., & Kim, D.-G. (2004). Over-expression of the ribosomal protein L36a gene is associated with cellular proliferation in hepatocellular carcinoma. *Hepatology (Baltimore, Md.)*, 39(1), 129–138. <https://doi.org/10.1002/hep.20017>

- Kim, S.-K., Kim, S.-Y., Kim, J.-H., Roh, S. A., Cho, D.-H., Kim, Y. S., & Kim, J. C. (2014). A nineteen gene-based risk score classifier predicts prognosis of colorectal cancer patients. *Molecular Oncology*, 8(8), 1653–1666. <https://doi.org/10.1016/j.molonc.2014.06.016>
- Kim, T.-M., Yang, I. S., Seung, B.-J., Lee, S., Kim, D., Ha, Y.-J., Seo, M., Kim, K.-K., Kim, H. S., Cheong, J.-H., Sur, J.-H., Nam, H., & Kim, S. (2020). Cross-species oncogenic signatures of breast cancer in canine mammary tumors. *Nature Communications*, 11(1), 3616. <https://doi.org/10.1038/s41467-020-17458-0>
- Kitazawa, S., Takaoka, Y., Ueda, Y., & Kitazawa, R. (2021). Identification of calmodulin-like protein 5 as tumor-suppressor gene silenced during early stage of carcinogenesis in squamous cell carcinoma of uterine cervix. *International Journal of Cancer*, 149(6), 1358–1368. <https://doi.org/10.1002/ijc.33687>
- Kitsoulis, C. V., Baxevanis, A. D., & Abatzopoulos, T. J. (2020). The occurrence of cancer in vertebrates: A mini review. *Journal of Biological Research-Thessaloniki*, 27(1), 9. <https://doi.org/10.1186/s40709-020-00119-0>
- Klimczak, M., Biecek, P., Zylicz, A., & Zylicz, M. (2019). Heat shock proteins create a signature to predict the clinical outcome in breast cancer. *Scientific Reports*, 9(1), 7507. <https://doi.org/10.1038/s41598-019-43556-1>
- Kozlov, A. P. (2016). Expression of evolutionarily novel genes in tumors. *Infectious Agents and Cancer*, 11. <https://doi.org/10.1186/s13027-016-0077-6>
- Kozłowska-Masłoń, J., Ciomborowska-Basheer, J., Kubiak, M. R., & Makałowska, I. (2024). Evolution of retrocopies in the context of HUSH silencing. *Biology Direct*, 19(1), 60. <https://doi.org/10.1186/s13062-024-00507-9>
- Kshitiz, null, Afzal, J., Maziarz, J. D., Hamidzadeh, A., Liang, C., Erkenbrack, E. M., Kim, H. N., Haeger, J.-D., Pfarrer, C., Hoang, T., Ott, T., Spencer, T., Pavličev, M., Antczak, D. F., Levchenko, A., & Wagner, G. P. (2019). Evolution of placental invasion and cancer metastasis are causally linked. *Nature Ecology & Evolution*, 3(12), 1743–1753. <https://doi.org/10.1038/s41559-019-1046-4>
- Kubiak, M. R., & Makałowska, I. (2017). Protein-Coding Genes' Retrocopies and Their Functions. *Viruses*, 9(4), E80. <https://doi.org/10.3390/v9040080>
- Kubiak, M. R., Szcześniak, M. W., & Makałowska, I. (2020). Complex Analysis of Retroposed Genes' Contribution to Human Genome, Proteome and Transcriptome. *Genes*, 11(5). <https://doi.org/10.3390/genes11050542>
- Kurozumi, S., Joseph, C., Sonbul, S., Alsaeed, S., Kariri, Y., Aljohani, A., Raafat, S., Alsaleem, M., Ogden, A., Johnston, S. J., Aleskandarany, M. A., Fujii, T., Shirabe, K., Caldas, C., Ashankyty, I., Dalton, L., Ellis, I. O., Desmedt, C., Green, A. R., ... Rakha, E. A. (2019). A key genomic subtype associated with lymphovascular invasion in invasive breast cancer. *British Journal of Cancer*, 120(12), 1129–1136. <https://doi.org/10.1038/s41416-019-0486-6>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Leinonen, R., Sugawara, H., & Shumway, M. (2011). The Sequence Read Archive. *Nucleic Acids Research*, 39(Database issue), D19–D21. <https://doi.org/10.1093/nar/gkq1019>
- Lever, J., Zhao, E. Y., Grewal, J., Jones, M. R., & Jones, S. J. M. (2019). CancerMine: A literature-mined resource for drivers, oncogenes and tumor suppressors in cancer. *Nature Methods*, 16(6), 505–507. <https://doi.org/10.1038/s41592-019-0422-y>
- Li, M., Sun, Q., & Wang, X. (2017). Transcriptional landscape of human cancers. *Oncotarget*, 8(21), 34534–34551. <https://doi.org/10.18632/oncotarget.15837>

- Li, P., Ju, X., & Yang, G. (2024). Concomitant double-fusion of PLEKHA7-ALK and INPP5D-ALK reveals favorable alectinib sensitivity in lung adenocarcinoma: A case report and literature review. *Discover Oncology*, 15(1), 43. <https://doi.org/10.1007/s12672-024-00899-0>
- Liao, L., Zhang, Y.-L., Deng, L., Chen, C., Ma, X.-Y., Andriani, L., Yang, S.-Y., Hu, S.-Y., Zhang, F.-L., Shao, Z.-M., & Li, D.-Q. (2023). Protein Phosphatase 1 Subunit PPP1R14B Stabilizes STMN1 to Promote Progression and Paclitaxel Resistance in Triple-Negative Breast Cancer. *Cancer Research*, 83(3), 471–484. <https://doi.org/10.1158/0008-5472.CAN-22-2709>
- Lin, J., Cao, S., Wang, Y., Hu, Y., Liu, H., Li, J., Chen, J., Li, P., Liu, J., Wang, Q., & Zheng, L. (2018). Long non-coding RNA UBE2CP3 enhances HCC cell secretion of VEGFA and promotes angiogenesis by activating ERK1/2/HIF-1 α /VEGFA signalling in hepatocellular carcinoma. *Journal of Experimental & Clinical Cancer Research: CR*, 37(1), 113. <https://doi.org/10.1186/s13046-018-0727-1>
- Liu, C.-C., Wang, H., Wang, W., Wang, L., Liu, W.-J., Wang, J.-H., Geng, Q.-R., & Lu, Y. (2018). ENO2 Promotes Cell Proliferation, Glycolysis, and Glucocorticoid-Resistance in Acute Lymphoblastic Leukemia. *Cellular Physiology and Biochemistry: International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology*, 46(4), 1525–1535. <https://doi.org/10.1159/000489196>
- Liu, D., Mao, Y., Chen, C., Zhu, F., Lu, W., & Ma, H. (2020). Expression patterns and clinical significances of ENO2 in lung cancer: An analysis based on Oncomine database. *Annals of Translational Medicine*, 8(10), 639. <https://doi.org/10.21037/atm-20-3354>
- Liu, G., Hou, G., Li, L., Li, Y., Zhou, W., & Liu, L. (2016). Potential diagnostic and prognostic marker dimethylglycine dehydrogenase (DMGDH) suppresses hepatocellular carcinoma metastasis in vitro and in vivo. *Oncotarget*, 7(22), 32607–32616. <https://doi.org/10.18632/oncotarget.8927>
- Liu, W., & Wang, X. (2019). Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biology*, 20(1), 18. <https://doi.org/10.1186/s13059-019-1629-z>
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method. *Methods*, 25(4), 402–408. <https://doi.org/10.1006/meth.2001.1262>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Luis-Ravelo, D., Antón, I., Zanduetta, C., Valencia, K., Pajares, M.-J., Agorreta, J., Montuenga, L., Vicent, S., Wistuba, I. I., De Las Rivas, J., & Lecanda, F. (2014). RHOB influences lung adenocarcinoma metastasis and resistance in a host-sensitive manner. *Molecular Oncology*, 8(2), 196–206. <https://doi.org/10.1016/j.molonc.2013.11.001>
- Macaulay, E. C., Chatterjee, A., Cheng, X., Baguley, B. C., Eccles, M. R., & Morison, I. M. (2017). The Genes of Life and Death: A Potential Role for Placental-Specific Genes in Cancer. *BioEssays*, 39(11), 1700091. <https://doi.org/10.1002/bies.201700091>
- Magadum, S., Banerjee, U., Murugan, P., Gangapur, D., & Ravikesavan, R. (2013). Gene duplication as a major force in evolution. *Journal of Genetics*, 92(1), 155–161. <https://doi.org/10.1007/s12041-013-0212-8>
- Marques, A. C., Dupanloup, I., Vinckenbosch, N., Reymond, A., & Kaessmann, H. (2005). Emergence of Young Human Genes after a Burst of Retroposition in Primates. *PLoS Biology*, 3(11), e357. <https://doi.org/10.1371/journal.pbio.0030357>
- Mazieres, J., Antonia, T., Daste, G., Muro-Cacho, C., Berchery, D., Tillement, V., Pradines, A., Sebti, S., & Favre, G. (2004). Loss of RhoB expression in human lung cancer progression. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 10(8), 2742–2750. <https://doi.org/10.1158/1078-0432.ccr-03-0149>

- McPherson, A. W., Chan, F. C., & Shah, S. P. (2018). Observing Clonal Dynamics across Spatiotemporal Axes: A Prelude to Quantitative Fitness Models for Cancer. *Cold Spring Harbor Perspectives in Medicine*, 8(2), a029603. <https://doi.org/10.1101/cshperspect.a029603>
- Navarro, F. C. P., & Galante, P. A. F. (2015). A Genome-Wide Landscape of Retrocopies in Primate Genomes. *Genome Biology and Evolution*, 7(8), 2265–2275. <https://doi.org/10.1093/gbe/evv142>
- O’Leary, N. A., Wright, M. W., Brister, J. R., Ciufu, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., ... Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, 44(D1), D733–745. <https://doi.org/10.1093/nar/gkv1189>
- Olivier, M., Hollstein, M., & Hainaut, P. (2010). TP53 mutations in human cancers: Origins, consequences, and clinical use. *Cold Spring Harbor Perspectives in Biology*, 2(1), a001008. <https://doi.org/10.1101/cshperspect.a001008>
- Padhi, A., Ghaly, M. M., & Ma, L. (2016). Testis-enriched heat shock protein A2 (HSPA2): Adaptive advantages of the birds with internal testes over the mammals with testicular descent. *Scientific Reports*, 6, 18770. <https://doi.org/10.1038/srep18770>
- Pan, D., & Zhang, L. (2007). Quantifying the major mechanisms of recent gene duplications in the human and mouse genomes: A novel strategy to estimate gene duplication rates. *Genome Biology*, 8(8), R158. <https://doi.org/10.1186/gb-2007-8-8-r158>
- Parker, H. G., VonHoldt, B. M., Quignon, P., Margulies, E. H., Shao, S., Mosher, D. S., Spady, T. C., Elkahoul, A., Cargill, M., Jones, P. G., Maslen, C. L., Acland, G. M., Sutter, N. B., Kuroki, K., Bustamante, C. D., Wayne, R. K., & Ostrander, E. A. (2009). An expressed *fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *Science (New York, N.Y.)*, 325(5943), 995–998. <https://doi.org/10.1126/science.1173275>
- Parmley, J. L., Urrutia, A. O., Potrzebowski, L., Kaessmann, H., & Hurst, L. D. (2007). Splicing and the Evolution of Proteins in Mammals. *PLoS Biology*, 5(2), e14. <https://doi.org/10.1371/journal.pbio.0050014>
- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14(4), 417–419. <https://doi.org/10.1038/nmeth.4197>
- Pegueroles, C., Laurie, S., & Albà, M. M. (2013). Accelerated Evolution after Gene Duplication: A Time-Dependent Process Affecting Just One Copy. *Molecular Biology and Evolution*, 30(8), 1830–1842. <https://doi.org/10.1093/molbev/mst083>
- Peng, H., Ishida, M., Li, L., Saito, A., Kamiya, A., Hamilton, J. P., Fu, R., Olaru, A. V., An, F., Popescu, I., Iacob, R., Dima, S., Alexandrescu, S. T., Grigorie, R., Nastase, A., Berindan-Neagoe, I., Tomuleasa, C., Graur, F., Zaharia, F., ... Selaru, F. M. (2015). Pseudogene INTS6P1 regulates its cognate gene INTS6 through competitive binding of miR-17-5p in hepatocellular carcinoma. *Oncotarget*, 6(8), 5666–5677. <https://doi.org/10.18632/oncotarget.3290>
- Penzo, M., Montanaro, L., Treré, D., & Derenzini, M. (2019). The Ribosome Biogenesis—Cancer Connection. *Cells*, 8(1), Article 1. <https://doi.org/10.3390/cells8010055>
- Poliseno, L., Marranci, A., & Pandolfi, P. P. (2015). Pseudogenes in Human Cancer. *Frontiers in Medicine*, 2. <https://doi.org/10.3389/fmed.2015.00068>
- Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., & Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*, 465(7301), 1033–1038. <https://doi.org/10.1038/nature09144>
- Privat, M., Cavard, A., Zekri, Y., Ponelle-Chachuat, F., Molnar, I., Sonnier, N., & Bignon, Y.-J. (2020). A high expression ratio of RhoA/RhoB is associated with the migratory and invasive properties

- of basal-like Breast Tumors. *International Journal of Medical Sciences*, 17(17), 2799–2808. <https://doi.org/10.7150/ijms.43101>
- R: The R Project for Statistical Computing. (b.d.). Pobrano 24 styczeń 2025, z <https://www.r-project.org/>
- Richardson, S. R., Salvador-Palomeque, C., & Faulkner, G. J. (2014). Diversity through duplication: Whole-genome sequencing reveals novel gene retrocopies in the human population. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, 36(5), 475–481. <https://doi.org/10.1002/bies.201300181>
- Ridley, A. J. (2013). RhoA, RhoB and RhoC have different roles in cancer cell migration. *Journal of Microscopy*, 251(3), 242–249. <https://doi.org/10.1111/jmi.12025>
- Rohozinski, J., Edwards, C. L., & Anderson, M. L. (2012). Does expression of the retrogene UTP14c in the ovary pre-dispose women to ovarian cancer? *Medical Hypotheses*, 78(4), 446–449. <https://doi.org/10.1016/j.mehy.2011.12.016>
- Rosikiewicz, W., Kabza, M., Kosiński, J. G., Ciomborowska-Basheer, J., Kubiak, M. R., & Makałowska, I. (2017). RetrogeneDB—a database of plant and animal retrocopies. *Database*, 2017(bax038). <https://doi.org/10.1093/database/bax038>
- Rosikiewicz, W., & Makałowska, I. (2016). Biological functions of natural antisense transcripts. *Acta Biochimica Polonica*, 63(4), 665–673. https://doi.org/10.18388/abp.2016_1350
- Russo, M., & Bardelli, A. (2017). Lesion-Directed Therapies and Monitoring Tumor Evolution Using Liquid Biopsies. *Cold Spring Harbor Perspectives in Medicine*, 7(2), a029587. <https://doi.org/10.1101/cshperspect.a029587>
- Rutnam, Z. J., Du, W. W., Yang, W., Yang, X., & Yang, B. B. (2014). The pseudogene TUSC2P promotes TUSC2 function by binding multiple microRNAs. *Nature Communications*, 5, 2914. <https://doi.org/10.1038/ncomms3914>
- Salgado, I., Prado Montes de Oca, E., Chairez, I., Figueroa-Yáñez, L., Pereira-Santana, A., Rivera Chávez, A., Velázquez-Fernandez, J. B., Alvarado Parra, T., & Vallejo, A. (2024). Deep Learning Techniques to Characterize the RPS28P7 Pseudogene and the Metazoa-SRP Gene as Drug Potential Targets in Pancreatic Cancer Patients. *Biomedicines*, 12(2), 0. <https://doi.org/10.3390/biomedicines12020395>
- Scieglińska, D., Sojka, D. R., Gogler-Piğłowska, A., Chumak, V., & Krawczyk, Z. (2020). Various Anti-HSPA2 Antibodies Yield Different Results in Studies on Cancer-Related Functions of Heat Shock Protein A2. *International Journal of Molecular Sciences*, 21(12), 4296. <https://doi.org/10.3390/ijms21124296>
- Seim, I., Ma, S., Zhou, X., Gerashchenko, M. V., Lee, S.-G., Suydam, R., George, J. C., Bickham, J. W., & Gladyshev, V. N. (2014). The transcriptome of the bowhead whale *Balaena mysticetus* reveals adaptations of the longest-lived mammal. *Aging*, 6(10), 879–899. <https://doi.org/10.18632/aging.100699>
- Shao, Y., Chen, C., Shen, H., He, B. Z., Yu, D., Jiang, S., Zhao, S., Gao, Z., Zhu, Z., Chen, X., Fu, Y., Chen, H., Gao, G., Long, M., & Zhang, Y. E. (2019). GenTree, an integrated resource for analyzing the evolution and function of primate-specific coding genes. *Genome Research*, 29(4), 682–696. <https://doi.org/10.1101/gr.238733.118>
- Shi, J., Yang, Y., Chen, F., Zhou, L., Wei, H., Dong, F., Wang, X., Shan, Y., & Chen, T. (2025). RPL36A activates ERK pathway and promotes colorectal cancer growth. *Translational Oncology*, 51, 102170. <https://doi.org/10.1016/j.tranon.2024.102170>
- Sojka, D. R., Gogler-Piğłowska, A., Vydra, N., Cortez, A. J., Filipczak, P. T., Krawczyk, Z., & Scieglińska, D. (2019). Functional redundancy of HSPA1, HSPA2 and other HSPA proteins in non-small cell lung carcinoma (NSCLC); an implication for NSCLC treatment. *Scientific Reports*, 9(1), 14394. <https://doi.org/10.1038/s41598-019-50840-7>

- Staszak, K., & Makałowska, I. (2021). Cancer, Retrogenes, and Evolution. *Life* (Basel, Switzerland), 11(1). <https://doi.org/10.3390/life11010072>
- Suhail, Y., Maziarz, J. D., Novin, A., Dighe, A., Afzal, J., Wagner, G., & Kshitiz. (2022). Tracing the cis-regulatory changes underlying the endometrial control of placental invasion. *Proceedings of the National Academy of Sciences*, 119(6), e2111256119. <https://doi.org/10.1073/pnas.2111256119>
- Sun, L., Lamont, S. J., Cooksey, A. M., McCarthy, F., Tudor, C. O., Vijay-Shanker, K., DeRita, R. M., Rothschild, M., Ashwell, C., Persia, M. E., & Schmidt, C. J. (2015). Transcriptome response to heat stress in a chicken hepatocellular carcinoma cell line. *Cell Stress & Chaperones*, 20(6), 939–950. <https://doi.org/10.1007/s12192-015-0621-0>
- Sunitha, J., Mahendra, J., Mahendra, L., & Devaraj, N. (2020). Molecular docking studies of a-mangostin with oral cancer targets ARRB1, FLNA, CALM3 and HTT. *Bioinformation*, 16(8), 625–630. <https://doi.org/10.6026/97320630016625>
- Szcześniak, M. W., Ciomborowska, J., Nowak, W., Rogozin, I. B., & Makałowska, I. (2011). Primate and rodent specific intron gains and the origin of retrogenes with splice variants. *Molecular Biology and Evolution*, 28(1), 33–37. <https://doi.org/10.1093/molbev/msq260>
- Tang, J., Ning, R., Zeng, B., & Li, Y. (2016). Molecular Evolution of PTEN Pseudogenes in Mammals. *PLoS ONE*, 11(12), e0167851. <https://doi.org/10.1371/journal.pone.0167851>
- Tang, Q., Gu, Y., Zhou, X., Jin, L., Guan, J., Liu, R., Li, J., Long, K., Tian, S., Che, T., Hu, S., Liang, Y., Yang, X., Tao, X., Zhong, Z., Wang, G., Chen, X., Li, D., Ma, J., ... Li, M. (2017). Comparative transcriptomics of 5 high-altitude vertebrates and their low-altitude relatives. *GigaScience*, 6(12), 1–9. <https://doi.org/10.1093/gigascience/gix105>
- The Cancer Genome Atlas Program (TCGA)—NCI (nciglobal,ncicenterprise). (2022, maj 13). [cgvMiniLanding]. <https://www.cancer.gov/ccg/research/genome-sequencing/tcga>
- The Human Protein Atlas. (b.d.). Pobrano 25 styczeń 2025, z <https://www.proteinatlas.org/>
- Tian, X., Azpurua, J., Hine, C., Vaidya, A., Myakishev-Rempel, M., Ablueva, J., Mao, Z., Nevo, E., Gorbunova, V., & Seluanov, A. (2013). High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. *Nature*, 499(7458), 346–349. <https://doi.org/10.1038/nature12234>
- Tollis, M., Schneider-Utaka, A. K., & Maley, C. C. (2020). The Evolution of Human Cancer Gene Duplications across Mammals. *Molecular Biology and Evolution*, 37(10), 2875–2886. <https://doi.org/10.1093/molbev/msaa125>
- Tutar, Y. (2012). Pseudogenes. *Comparative and Functional Genomics*, 2012, 424526. <https://doi.org/10.1155/2012/424526>
- Varley, K. E., Gertz, J., Roberts, B. S., Davis, N. S., Bowling, K. M., Kirby, M. K., Nesmith, A. S., Oliver, P. G., Grizzle, W. E., Forero, A., Buchsbaum, D. J., LoBuglio, A. F., & Myers, R. M. (2014). Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Research and Treatment*, 146(2), 287–297. <https://doi.org/10.1007/s10549-014-3019-2>
- Venkatesan, S., Swanton, C., Taylor, B. S., & Costello, J. F. (2017). Treatment-Induced Mutagenesis and Selective Pressures Sculpt Cancer Evolution. *Cold Spring Harbor Perspectives in Medicine*, 7(8), a026617. <https://doi.org/10.1101/cshperspect.a026617>
- Villalobo, A., González-Muñoz, M., & Berchtold, M. W. (2019). Proteins with calmodulin-like domains: Structures and functional roles. *Cellular and Molecular Life Sciences: CMLS*, 76(12), 2299–2328. <https://doi.org/10.1007/s00018-019-03062-z>
- Vinckenbosch, N., Dupanloup, I., & Kaessmann, H. (2006). Evolutionary fate of retroposed gene copies in the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, 103(9), 3220–3225. <https://doi.org/10.1073/pnas.0511307103>

- Vincze, O., Colchero, F., Lemaître, J.-F., Conde, D. A., Pavard, S., Bieuville, M., Urrutia, A. O., Ujvari, B., Boddy, A. M., Maley, C. C., Thomas, F., & Giraudeau, M. (2022). Cancer risk across mammals. *Nature*, 601(7892), 263–267. <https://doi.org/10.1038/s41586-021-04224-5>
- Wang, M., Hu, Y., & Stearns, M. E. (2009). RPS2: A novel therapeutic target in prostate cancer. *Journal of Experimental & Clinical Cancer Research: CR*, 28(1), 6. <https://doi.org/10.1186/1756-9966-28-6>
- Wang, M.-Y., Chen, D.-P., Qi, B., Li, M.-Y., Zhu, Y.-Y., Yin, W.-J., He, L., Yu, Y., Li, Z.-Y., Lin, L., Yang, F., Lin, Z.-R., & Liu, J.-Q. (2019). Pseudogene RACGAP1P activates RACGAP1/Rho/ERK signalling axis as a competing endogenous RNA to promote hepatocellular carcinoma early recurrence. *Cell Death & Disease*, 10(6), 426. <https://doi.org/10.1038/s41419-019-1666-2>
- Wang, Q.-S., Shi, L.-L., Sun, F., Zhang, Y.-F., Chen, R.-W., Yang, S.-L., & Hu, J.-L. (2019). High Expression of ANXA2 Pseudogene ANXA2P2 Promotes an Aggressive Phenotype in Hepatocellular Carcinoma. *Disease Markers*, 2019, 9267046. <https://doi.org/10.1155/2019/9267046>
- Wang, R., Zhang, T., Yang, Z., Jiang, C., & Seng, J. (2018). Long non-coding RNA FTH1P3 activates paclitaxel resistance in breast cancer through miR-206/ABCB1. *Journal of Cellular and Molecular Medicine*, 22(9), 4068–4075. <https://doi.org/10.1111/jcmm.13679>
- Wang, S., & Yu, J. (2019). Long non-coding RNA transcribed from pseudogene PPIAP43 is associated with radiation sensitivity of small cell lung cancer cells. *Oncology Letters*, 18(5), 4583. <https://doi.org/10.3892/ol.2019.10806>
- Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., Ellrott, K., Shmulevich, I., Sander, C., & Stuart, J. M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics*, 45(10), Article 10. <https://doi.org/10.1038/ng.2764>
- Werner, A., Kanhere, A., Wahlestedt, C., & Mattick, J. S. (2024). Natural antisense transcripts as versatile regulators of gene expression. *Nature Reviews. Genetics*, 25(10), 730–744. <https://doi.org/10.1038/s41576-024-00723-z>
- Whirl-Carrillo, M., Huddart, R., Gong, L., Sangkuhl, K., Thorn, C. F., Whaley, R., & Klein, T. E. (2021). An Evidence-Based Framework for Evaluating Pharmacogenomics Knowledge for Personalized Medicine. *Clinical Pharmacology and Therapeutics*, 110(3), 563–572. <https://doi.org/10.1002/cpt.2350>
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer. <https://doi.org/10.1007/978-0-387-98141-3>
- Wu, J., Li, X., Luo, F., Yan, J., & Yang, K. (2020). Screening key miRNAs and genes in prostate cancer by microarray analysis. *Translational Cancer Research*, 9(2), 856–868. <https://doi.org/10.21037/tcr.2019.12.30>
- Xu, H., Chen, B., Xing, J., Wei, Z., Liu, C., Qiu, Y., Lin, Y., & Ren, L. (2019). Upregulation of LGMNP1 confers radiotherapy resistance in glioblastoma. *Oncology Reports*, 41(6), 3435–3443. <https://doi.org/10.3892/or.2019.7128>
- Xu, Q., Song, Z., Zhu, C., Tao, C., Kang, L., Liu, W., He, F., Yan, J., & Sang, T. (2017). Systematic comparison of lncRNAs with protein coding mRNAs in population expression and their response to environmental change. *BMC Plant Biology*, 17(1), 42. <https://doi.org/10.1186/s12870-017-0984-8>
- Yates, A. D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Marugán, J. C., Cummins, C., Davidson, C., Dodiya, K., Fatima, R., Gall, A., ... Flicek, P. (2020). Ensembl 2020. *Nucleic Acids Research*, 48(D1), D682–D688. <https://doi.org/10.1093/nar/gkz966>

- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13, 134. <https://doi.org/10.1186/1471-2105-13-134>
- Yu, G., Yao, W., Gumireddy, K., Li, A., Wang, J., Xiao, W., Chen, K., Xiao, H., Li, H., Tang, K., Ye, Z., Huang, Q., & Xu, H. (2014). Pseudogene PTENP1 functions as a competing endogenous RNA to suppress clear-cell renal cell carcinoma progression. *Molecular Cancer Therapeutics*, 13(12), 3086–3097. <https://doi.org/10.1158/1535-7163.MCT-14-0245>
- Zhai, L.-L., Qiao, P.-P., Sun, Y.-S., Tang, Z.-G., & Ju, T.-F. (2021). Upregulated HSPA2 predicts early relapse of pancreatic cancer after surgery. *Gland Surgery*, 10(7), 2140–2149. <https://doi.org/10.21037/gs-21-262>
- Zhang, H., Chen, W., Duan, C.-J., & Zhang, C.-F. (2013). Overexpression of HSPA2 is correlated with poor prognosis in esophageal squamous cell carcinoma. *World Journal of Surgical Oncology*, 11(1), 141. <https://doi.org/10.1186/1477-7819-11-141>
- Zhang, T., Yang, P., Wei, J., Li, W., Zhong, J., Chen, H., & Cao, J. (2018). Overexpression of flavin-containing monooxygenase 5 predicts poor prognosis in patients with colorectal cancer. *Oncology Letters*, 15(3), 3923–3927. <https://doi.org/10.3892/ol.2018.7724>
- Zhang, W., & Huang, P. (2011). Cancer-stromal interactions: Role in cell survival, metabolism and drug sensitivity. *Cancer Biology & Therapy*, 11(2), 150–156. <https://doi.org/10.4161/cbt.11.2.14623>
- Zhang, Y., Liu, S., Zhang, X., Li, W.-B., Chen, Y., Huang, X., Sun, L., Luo, W., Netzer, W. J., Threadgill, R., Wiegand, G., Wang, R., Cohen, S. N., Greengard, P., Liao, F.-F., Li, L., & Xu, H. (2009). A functional mouse retroposed gene Rps23r1 reduces Alzheimer's beta-amyloid levels and tau phosphorylation. *Neuron*, 64(3), 328–340. <https://doi.org/10.1016/j.neuron.2009.08.036>
- Zhou, Y., Xu, B., Zhou, Y., Liu, J., Zheng, X., Liu, Y., Deng, H., Liu, M., Ren, X., Xia, J., Kong, X., Huang, T., & Jiang, J. (2021). Identification of Key Genes With Differential Correlations in Lung Adenocarcinoma. *Frontiers in Cell and Developmental Biology*, 9, 675438. <https://doi.org/10.3389/fcell.2021.675438>
- Zhu, X., Li, S., Xu, B., & Luo, H. (2021). Cancer evolution: A means by which tumors evade treatment. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*, 133, 111016. <https://doi.org/10.1016/j.biopha.2020.111016>

ANEKS

Tabela 21. Wyniki analizy ekspresji różnicowej retrogenów dla raka piersi z zestawu danych o numerze PRJNA251383.

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000179277	retro_hsap_1718	MEIS3P1	MEIS3	-2,14	3,19E-26	7,22E-25	24,69	24,68	5,71	3,74	-18,98
ENSG00000237550	retro_hsap_1605	retro_hsap_1605	RPL9	-1,94	4,78E-14	2,87E-13	193,31	180,69	45,30	15,32	-148,00
ENSG00000220842	retro_hsap_738	RPL21P16	RPL21	-1,43	1,09E-10	4,78E-10	452,79	444,15	153,00	112,61	-299,79
ENSG00000230953	retro_hsap_154	retro_hsap_154	TMA7	-1,42	1,09E-13	6,30E-13	25,18	19,01	9,11	5,57	-16,07
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,32	7,35E-11	3,27E-10	716,78	637,77	297,42	199,38	-419,37
ENSG00000155876	retro_hsap_4	RRAGA	RRAGB	-0,94	2,65E-10	1,11E-09	59,26	58,50	30,59	28,00	-28,67
ENSG00000182107	retro_hsap_12	TMEM30B	TMEM30A	-0,78	9,33E-05	0,00022	34,16	28,63	12,30	7,21	-21,87
zwiększona ekspresja											
ENSG00000132967	retro_hsap_2613	HMGB1P5	HMGB1	0,88	1,28E-16	9,95E-16	73,86	68,24	131,46	122,07	57,61
ENSG00000259781	retro_hsap_1589	HMGB1P6	HMGB1	0,91	1,62E-14	1,02E-13	18,88	18,08	34,28	30,86	15,40
ENSG00000235655	retro_hsap_2180	H3P6	H3-3B	1,22	9,89E-09	3,58E-08	113,93	108,16	263,31	226,69	149,38
retro_hsap_2297	retro_hsap_2297	retro_hsap_2297	PPIA	1,29	1,40E-22	2,11E-21	28,65	27,98	64,47	49,04	35,83
ENSG00000250461	retro_hsap_3179	retro_hsap_3179	FAM133B	1,53	2,08E-14	1,29E-13	14,07	13,98	32,04	20,63	17,97
ENSG00000179967	retro_hsap_3101	PPP1R14BP3	PPP1R14B	1,58	2,25E-21	2,93E-20	9,67	8,44	27,84	21,21	18,17
ENSG00000242085	retro_hsap_1383	RPS20P33	RPS20	1,82	4,95E-31	1,74E-29	7,79	6,79	25,90	19,97	18,10
ENSG00000241170	retro_hsap_794	retro_hsap_794	RPL31	1,99	4,98E-34	2,19E-32	4,69	4,17	19,00	12,68	14,31
ENSG00000250011	retro_hsap_2219	HMGB1P3	HMGB1	2,01	2,97E-23	4,86E-22	5,04	4,38	18,05	11,16	13,01
ENSG00000240429	retro_hsap_2887	LRRFIP1P1	LRRFIP1	2,36	5,57E-39	3,52E-37	3,61	3,51	16,05	10,30	12,44
ENSG00000260459	retro_hsap_1647	FTLP14	FTL	2,50	1,84E-26	4,29E-25	3,81	2,64	17,28	10,89	13,48
ENSG00000248394	retro_hsap_2981	FOSL1P1	FOSL1	2,70	4,84E-37	2,67E-35	2,15	1,59	13,55	8,61	11,41
ENSG00000178372	CALML5	CALML5	CALM3	2,87	8,60E-11	3,80E-10	45,25	19,57	328,33	12,25	283,08
ENSG00000213896	retro_hsap_3325	retro_hsap_3325	RPL31	2,91	8,98E-19	8,86E-18	2,00	1,67	13,82	4,26	11,82
ENSG00000230457	retro_hsap_2868	PA2G4P4	PA2G4	2,96	9,27E-36	4,65E-34	2,21	1,75	14,72	9,88	12,51

Tabela 22. Wyniki analizy ekspresji różnicowej retrogenów dla typów raka piersi – ER+ oraz TNBC z zestawu danych o numerze PRJNA251383.

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM ER+	Mediana TPM ER+	Średnia TPM TNBC	Mediana TPM TNBC	Różnica TPM
zmniejszona ekspresja											
ENSG00000213222	retro_hsap_2321	TOMM40P4	TOMM40	-3,08	2,59E-31	1,15E-28	14,79	8,69	1,58	1,55	-13,21
ENSG00000178363	retro_hsap_59	CALML3	CALM3	-2,24	3,14E-08	4,47E-07	14,25	8,37	2,77	1,44	-11,48
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,85	1,70E-15	9,21E-14	477,39	427,42	117,45	111,78	-359,94
ENSG00000126803	retro_hsap_68	HSPA2	HSPA8	-1,27	2,63E-06	2,40E-05	24,35	10,95	8,04	6,36	-16,31
ENSG00000248394	retro_hsap_2981	FOSL1P1	FOSL1	-1,08	5,32E-05	0,00034627	18,97	10,41	8,13	5,92	-10,84
ENSG00000182107	retro_hsap_12	TMEM30B	TMEM30A	-0,88	0,00082565	0,00372989	17,45	12,34	7,14	5,47	-10,31
ENSG00000242085	retro_hsap_1383	RPS20P33	RPS20	-0,69	0,00065792	0,00305876	33,12	22,86	18,67	15,16	-14,44
zwiększona ekspresja											
ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	0,81	0,00011059	0,00065438	140,19	118,90	213,28	167,09	73,10
ENSG00000213896	retro_hsap_3325	retro_hsap_3325	RPL31	1,72	0,0001232	0,00071804	6,75	1,74	20,89	6,30	14,14
ENSG00000178372	CALML5	CALML5	CALM3	4,16	1,46E-15	8,01E-14	38,21	2,96	618,46	322,73	580,25

Tabela 23. Wyniki analizy ekspresji różnicowej retrogenów dla gruczolakoraka płuc z zestawu danych o numerze PRJEB2784.

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-0,97	1,09E-15	1,05E-14	285,66	256,38	155,12	116,70	-130,54
ENSG00000237550	retro_hsap_1605	retro_hsap_1605	RPL9	-0,84	3,37E-12	2,32E-11	30,00	28,82	16,19	15,02	-13,81
zwiększona ekspresja											
ENSG00000213189	retro_hsap_2173	BTF3L4P2	BTF3L4	0,65	5,99E-14	4,92E-13	23,05	22,86	34,45	32,39	11,40
ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	0,68	2,40E-12	1,67E-11	531,92	491,72	826,37	716,81	294,44
ENSG00000175793	retro_hsap_8	SFN	YWHAQ	1,44	1,21E-13	9,63E-13	22,54	16,51	59,87	41,08	37,33

Tabela 24. Wyniki analizy ekspresji różnicowej retrogenów dla raka jelita grubego z zestawu danych o numerze PRJNA218851.

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM guz pierwotny	Mediana TPM guz pierwotny	Różnica TPM
zmniejszona ekspresja											
ENSG00000182107	retro_hsap_12	TMEM30B	TMEM30A	-0,87	6,47E-04	4,19E-03	31,96	31,92	18,99	17,21	-12,97
ENSG00000255112	retro_hsap_75	CHMP1B	CHMP1A	-0,63	5,97E-07	1,49E-05	39,50	37,66	28,43	26,74	-11,07
zwiększona ekspresja											
ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	0,71	3,90E-04	2,79E-03	525,63	534,61	952,97	821,72	427,33
ENSG00000259781	retro_hsap_1589	HMGB1P6	HMGB1	0,88	1,12E-06	2,51E-05	153,50	154,25	324,27	321,53	170,77
ENSG00000137124	retro_hsap_19	ALDH1B1	ALDH2	0,94	1,75E-05	2,31E-04	29,67	27,79	63,32	47,68	33,65

Tabela 25. Wyniki analizy ekspresji różnicowej retrogenów dla raka jelita grubego – przerzuty vs guzy pierwotne z zestawu danych o numerze PRJNA218851.

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM guz pierwotny	Mediana TPM guz pierwotny	średnia TPM przerzuty	mediana TPM przerzuty	Różnica TPM
zmniejszona ekspresja											
ENSG00000126803	retro_hsap_68	HSPA2	HSPA8	-1,53	2,30E-05	0,000744	15,33	11,12	5,26	3,99	-10,07

Tabela 26. Wyniki analizy ekspresji różnicowej retrogenów dla raka wątrobowokomórkowego z zestawu danych o numerze PRJNA310012.

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,82	3,83E-33	3,10E-31	221,25	181	92,96	76,75	-128,29
ENSG00000230953	retro_hsap_154	retro_hsap_154	TMA7	-1,51	8,59E-08	4,96E-07	33,00	32,34	21,78	13,62	-11,22
ENSG00000137124	retro_hsap_19	ALDH1B1	ALDH2	-1,49	4,16E-15	6,26E-14	37,32	38,15	20,50	14,94	-16,82
zwiększona ekspresja											
ENSG00000234851	retro_hsap_2874	RPL23AP42	RPL23A	0,84	6,49E-05	0,000236	121,85	109,16	342,19	301,30	220,34
ENSG00000182952	retro_hsap_100	HMGH4	HMGH2	0,84	1,30E-17	2,58E-16	17,95	15,44	46,37	37,27	28,41
ENSG00000177105	retro_hsap_28	RHOG	RAC1	0,87	4,31E-08	2,60E-07	12,16	11,32	33,31	25,11	21,15
ENSG00000240342	retro_hsap_1060	RPS2P5	RPS2	0,97	0,000291	0,000928	28,94	28,19	79,62	61,75	50,68
ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	1,32	2,21E-20	5,81E-19	181,90	164,87	708,32	591,56	526,42
ENSG00000132661	retro_hsap_80	NXT1	NXT2	1,35	2,48E-20	6,47E-19	4,21	3,65	14,97	12,84	10,76
ENSG00000237550	retro_hsap_1605	retro_hsap_1605	RPL9	1,36	9,37E-08	5,38E-07	8,24	7,12	27,63	19,36	19,39
ENSG00000175793	retro_hsap_8	SFN	YWHAQ	6,13	9,57E-60	3,47E-57	0,22	0,10	22,17	12,71	21,95

Tabela 27. Wyniki analizy ekspresji różnicowej retrogenów dla wątrobiaka zarodkowego z zestawu danych o numerze PRJNA413799.

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000137124	retro_hsap_19	ALDH1B1	ALDH2	-1,56	9,13E-13	2,87E-11	60,31	57,48	32,47	32,27	-27,85
ENSG00000226015	retro_hsap_449	CCT8P1	CCT8	-1,38	9,55E-05	0,0004948	26,08	22,35	15,36	12,72	-10,72
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,04	3,83E-07	3,53E-06	365,90	389,64	279,93	267,92	-85,97

Tabela 28. Wyniki analizy ekspresji różnicowej retrogenów dla subtypów molekularnych wątrobiaka zarodkowego – C1 oraz C2 (PRJNA413799).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM C1	Mediana TPM C1	Średnia TPM C2	Mediana TPM C2	Różnica TPM
zmniejszona ekspresja											
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,08	0,00019	0,00417	345,16	327,11	169,49	184,84	-175,67
zwiększona ekspresja											
ENSG00000240342	retro_hsap_1060	RPS2P5	RPS2	1,27	0,00008	0,00222	14,69	13,26	37,94	34,30	23,24

Tabela 29. Wyniki analizy ekspresji różnicowej retrogenów dla raka gruczołu mlekowego psa z zestawu danych o numerze PRJNA489087.

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2 FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSCAFG00000005235	retro_cfam_121	CALML5	CALM3	-5,09	5,61E-22	5,08E-20	402,61	53,34	13,95	0,22	-388,66
retro_cfam_942	retro_cfam_942	retro_cfam_942	CAPNS1	-2,87	8,06E-19	4,65E-17	18,12	6,47	3,02	0,88	-15,09
ENSCAFG00000016031	retro_cfam_140	HSPA2	HSPA8	-1,39	1,35E-18	7,49E-17	29,44	16,76	13,03	9,87	-16,41
ENSCAFG00000024041	retro_cfam_179	retro_cfam_179	RPS3A	-0,75	6,66E-11	1,07E-09	115,27	106,75	80,94	79,74	-34,33
ENSCAFG00000003850	retro_cfam_186	RHOB	RHOA	-0,64	2,18E-07	1,88E-06	221,94	178,69	160,55	134,85	-61,39
zwiększona ekspresja											
ENSCAFG00000004458	retro_cfam_1061	retro_cfam_1061	LDHA	0,60	5,28E-05	0,000263009	48,93	38,82	82,95	70,87	34,01
retro_cfam_1036	retro_cfam_1036	retro_cfam_1036	ENSCAFG00000007199	1,01	1,12E-07	1,02E-06	55,78	43,06	125,03	115,14	69,26
retro_cfam_320	retro_cfam_320	retro_cfam_320	ENO2	1,14	6,47E-18	3,24E-16	72,49	65,57	190,58	148,57	118,09

Tabela 30. Wyniki analizy ekspresji różnicowej retrogenów dla raka gruczołu mlekowego psa z zestawu danych o numerze PRJNA561580.

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalaająca	Mediana TPM tkanka okalaająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zwiększona ekspresja											
retro_cfam_320	retro_cfam_320	retro_cfam_320	<i>ENO2</i>	0,91	0,000254	0,009868	86,16	71,81	179,45	116,95	93,28

Tabela 31. Wyniki analizy ekspresji różnicowej retrogenów dla linii komórkowej raka wątrobowokomórkowego kury (PRJNA239321) oraz próbek zdrowych wątroby kury (PRJNA362606).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalaająca	Mediana TPM tkanka okalaająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSGALG00000016485	retro_ggal_5	RHOB	<i>RHOA</i>	-1,42	2,16E-06	5,00E-06	68,83	70,18	41,43	33,54	-27,40
zwiększona ekspresja											
ENSGALG00000011715	retro_ggal_8	HSPA2	<i>HSPA8</i>	1,01	1,49E-06	3,50E-06	16,65	17,16	65,51	59,13	48,86
retro_ggal_15	retro_ggal_15	retro_ggal_15	<i>C12orf45</i>	5,49	5,17E-09	1,47E-08	0	0	34,90	36,17	34,90
retro_ggal_18	retro_ggal_18	retro_ggal_18	<i>Pou5f3</i>	6,57	4,60E-14	1,82E-13	0,04	0	19,44	18,72	19,40
retro_ggal_14	retro_ggal_14	retro_ggal_14	<i>SCO1</i>	7,84	4,71E-20	2,56E-19	0	0	28,59	28,71	28,59

TCGA

Tabela 32. Wyniki analizy ekspresji różnicowej retrogenów dla raka piersi (TCGA-BRCA).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalaająca	Mediana TPM tkanka okalaająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000230953	retro_hsap_154	retro_hsap_154	<i>TMA7</i>	-1,30	2,74E-19	1,75E-18	28,57	25,39	11,57	6,86	-17,00
ENSG00000187866	retro_hsap_35	FAM122A	<i>FAM122C</i>	-0,91	6,09E-96	5,10E-94	32,61	32,13	17,43	16,88	-15,18

ENSG00000179277	retro_hsap_1718	MEIS3P1	MEIS3	-0,78	4,32E-15	2,23E-14	27,16	26,59	16,47	13,83	-10,69
zwiększona ekspresja											
ENSG00000137124	retro_hsap_19	ALDH1B1	ALDH2	0,69	3,07E-21	2,16E-20	32,64	30,97	56,31	48,99	23,66
ENSG00000198618	retro_hsap_2497	PPIAP22	PPIA	0,71	8,75E-33	1,03E-31	139,85	127,90	222,42	211,47	82,57
ENSG00000233913	retro_hsap_3265	RPL10P9	RPL10	0,75	0,000291	0,000647	43,41	11,40	62,16	9,39	18,75
ENSG00000126803	retro_hsap_68	HSPA2	HSPA8	0,78	3,02E-08	9,90E-08	17,60	15,29	34,47	18,77	16,87
ENSG00000175793	retro_hsap_8	SFN	YWHAQ	0,89	1,44E-10	5,60E-10	101,30	89,89	169,46	112,40	68,16
ENSG00000264350	retro_hsap_1943	SNRPGP2	SNRPG	1,09	5,62E-44	1,01E-42	13,57	12,31	27,97	24,72	14,40
ENSG00000186940	retro_hsap_4184	CHCHD2P9	CHCHD2	1,15	4,46E-22	3,27E-21	9,03	7,92	19,89	16,99	10,86
ENSG00000235655	retro_hsap_2180	H3P6	H3-3B	1,17	1,12E-36	1,55E-35	29,89	29,01	65,67	57,79	35,79
ENSG00000179967	retro_hsap_3101	PPP1R14BP3	PPP1R14B	1,59	8,09E-57	2,22E-55	32,47	30,46	91,96	75,75	59,49
ENSG00000228232	retro_hsap_4806	GAPDHP1	GAPDH	1,76	1,13E-21	8,12E-21	8,19	2,98	25,07	7,13	16,89
ENSG00000178372	CALML5	CALML5	CALM3	3,01	2,88E-26	2,57E-25	60,71	20,94	308,71	15,06	248,00
ENSG00000237442	retro_hsap_2096	HNRNPA1P57	HNRNPA1	6,43	5,44E-78	2,86E-76	0,18	0,05	12,69	0,24	12,51

Tabela 33. Wyniki analizy ekspresji różnicowej retrogenów dla gruczolaka jelita grubego (TCGA-COAD).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000224321	retro_hsap_117	RPL12P14	RPL12	-3,22	2,02E-75	7,86E-73	17,93	17,32	2,53	1,95	-15,40
ENSG00000126803	retro_hsap_68	HSPA2	HSPA8	-1,62	3,50E-12	2,37E-11	38,58	24,91	14,60	7,13	-23,98
ENSG00000187534	retro_hsap_2058	PRR13P5	PRR13	-1,21	2,80E-13	2,12E-12	29,35	28,60	15,95	14,58	-13,40
ENSG00000182107	retro_hsap_12	TMEM30B	TMEM30A	-1,18	1,95E-35	7,32E-34	83,96	85,01	46,44	43,92	-37,52
ENSG00000255112	retro_hsap_75	CHMP1B	CHMP1A	-0,92	1,63E-28	3,95E-27	82,18	80,94	54,03	51,39	-28,15
ENSG00000175793	retro_hsap_8	SFN	YWHAQ	-0,63	0,000153	0,000407	583,08	514,87	454,88	373,36	-128,19
zwiększona ekspresja											

ENSG00000283041	retro_hsap_3906	retro_hsap_3906	EEF1G	0,69	1,88E-05	5,77E-05	24,20	23,87	45,34	35,61	21,14
ENSG00000226479	retro_hsap_104	TMEM185B	TMEM185A	0,76	1,97E-31	5,66E-30	15,71	14,36	33,04	30,78	17,33
ENSG00000215006	retro_hsap_3186	CHCHD2P2	CHCHD2	0,80	3,90E-05	0,000114	10,33	9,62	20,33	16,47	10,00
ENSG00000240342	retro_hsap_1060	RPS2P5	RPS2	0,89	1,57E-09	8,09E-09	148,38	158,60	331,12	290,85	182,75
ENSG00000213442	retro_hsap_1053	RPL18AP3	RPL18A	0,90	9,85E-06	3,17E-05	60,30	53,49	132,10	87,13	71,80
ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	0,95	2,21E-10	1,25E-09	58,75	52,39	134,18	115,61	75,43
ENSG00000244313	retro_hsap_869	retro_hsap_869	RPS10	0,96	3,40E-07	1,33E-06	41,56	42,45	89,78	65,78	48,22
ENSG00000250182	retro_hsap_3291	EEF1A1P13	EEF1A1	0,98	1,92E-05	5,89E-05	11,29	11,12	27,01	12,86	15,73
ENSG00000218426	retro_hsap_3682	retro_hsap_3682	RPL27A	0,99	3,37E-07	1,32E-06	45,37	41,17	100,26	70,46	54,90
ENSG00000236552	retro_hsap_615	RPL13AP5	RPL13A	1,03	3,94E-07	1,53E-06	63,79	58,60	146,11	92,89	82,32
ENSG00000189343	retro_hsap_1812	RPS2P46	RPS2	1,11	1,25E-07	5,14E-07	21,11	13,42	50,60	33,16	29,49
ENSG00000230202	retro_hsap_3511	retro_hsap_3511	RPL29	1,13	1,32E-06	4,78E-06	33,25	24,69	85,18	41,01	51,92
ENSG00000233762	retro_hsap_2341	RPS15P4	RPS15	1,16	2,87E-12	1,96E-11	27,18	24,38	68,23	58,54	41,05
ENSG00000228232	retro_hsap_4806	GAPDHP1	GAPDH	1,21	1,91E-07	7,69E-07	9,52	5,24	27,25	13,75	17,73
ENSG00000229638	retro_hsap_2898	RPL4P4	RPL4	1,21	1,67E-11	1,05E-10	6,12	5,42	16,20	12,03	10,08
ENSG00000181524	retro_hsap_3583	RPL24P4	RPL24	1,25	1,24E-09	6,47E-09	14,72	12,60	40,06	26,42	25,34
ENSG00000263266	retro_hsap_1737	RPS7P1	RPS7	1,25	1,61E-12	1,13E-11	21,94	19,68	60,19	42,46	38,24
ENSG00000178715	retro_hsap_118	retro_hsap_118	RPS16	1,25	6,49E-06	2,14E-05	6,42	6,67	16,80	9,99	10,38
ENSG00000264350	retro_hsap_1943	SNRPGP2	SNRPG	1,26	4,10E-17	4,33E-16	16,78	12,48	47,98	41,16	31,20
ENSG00000230629	retro_hsap_4697	RPS23P8	RPS23	1,33	3,54E-11	2,16E-10	10,68	9,42	30,77	19,50	20,09
ENSG00000223803	retro_hsap_2739	RPS20P14	RPS20	1,38	3,79E-14	3,10E-13	5,56	5,02	16,74	12,39	11,18
ENSG00000132661	retro_hsap_80	NXT1	NXT2	1,40	4,93E-26	1,00E-24	30,31	27,26	94,28	79,92	63,97
ENSG00000213553	retro_hsap_2095	RPLP0P6	RPLP0	1,44	2,29E-11	1,42E-10	20,97	18,10	62,38	39,52	41,41
ENSG00000242299	retro_hsap_2820	retro_hsap_2820	RPS18	1,55	4,38E-12	2,93E-11	34,56	30,70	107,17	62,19	72,61
ENSG00000169564	retro_hsap_83	PCBP1	PCBP2	1,56	5,65E-10	3,06E-09	10,83	10,72	35,62	13,70	24,80
ENSG00000230897	retro_hsap_1806	RPS18P12	RPS18	1,57	1,15E-09	6,04E-09	4,55	3,55	15,91	7,25	11,36
ENSG00000240376	retro_hsap_3158	retro_hsap_3158	RPS8	1,58	1,96E-09	1,00E-08	11,37	10,74	36,19	15,25	24,82
ENSG00000228929	retro_hsap_388	RPS13P2	RPS13	1,58	1,82E-13	1,40E-12	5,33	5,12	17,65	10,62	12,33
ENSG00000233913	retro_hsap_3265	RPL10P9	RPL10	1,78	4,87E-08	2,11E-07	21,28	8,64	86,73	17,86	65,45
ENSG00000218175	retro_hsap_2348	retro_hsap_2348	RPLP1	1,79	1,75E-10	1,00E-09	28,96	19,67	101,54	36,32	72,58

ENSG00000146677	retro_hsap_3726	retro_hsap_3726	RPL32	1,79	5,89E-16	5,66E-15	11,11	9,95	41,51	23,44	30,40
ENSG00000228205	retro_hsap_2736	retro_hsap_2736	RPS3	1,87	4,78E-13	3,53E-12	4,75	4,44	18,48	8,72	13,73
ENSG00000226608	retro_hsap_2407	FTLP3	FTL	1,89	1,01E-12	7,21E-12	11,69	9,40	50,20	18,88	38,51
ENSG00000226396	retro_hsap_119	retro_hsap_119	RPS14	1,99	9,11E-25	1,70E-23	3,68	3,45	16,77	11,98	13,09
ENSG00000215030	retro_hsap_1810	RPL13P12	RPL13	2,17	1,86E-18	2,16E-17	47,61	40,57	237,10	121,12	189,49
ENSG00000180211	retro_hsap_3581	retro_hsap_3581	RPL23	2,27	1,17E-17	1,30E-16	3,06	2,57	15,98	6,98	12,92
ENSG00000239559	retro_hsap_887	RPL37P2	RPL37	2,38	2,35E-21	3,44E-20	2,38	2,10	14,53	6,74	12,15
ENSG00000239528	retro_hsap_3232	RPS14P8	RPS14	2,40	4,97E-11	2,99E-10	3,23	1,51	18,45	6,09	15,22
ENSG00000232888	retro_hsap_1068	RPS11P5	RPS11	2,61	8,88E-21	1,24E-19	6,52	5,45	41,91	13,57	35,40
ENSG00000213058	retro_hsap_268	retro_hsap_268	RPS14	2,64	2,30E-08	1,04E-07	18,71	11,36	142,30	10,89	123,59
ENSG00000235297	retro_hsap_1952	FAUP1	FAU	2,73	1,55E-16	1,56E-15	2,82	2,24	19,12	4,54	16,30
ENSG00000213315	retro_hsap_1371	retro_hsap_1371	RPS18	3,40	8,54E-27	1,82E-25	2,89	2,55	32,13	7,47	29,23
ENSG00000240950	retro_hsap_2851	retro_hsap_2851	RPL8	3,82	1,25E-27	2,84E-26	1,02	0,89	16,14	2,75	15,12
ENSG00000240668	retro_hsap_2842	KRT8P36	KRT8	3,86	4,50E-20	5,93E-19	1,65	1,62	28,23	1,59	26,58
ENSG00000240652	retro_hsap_821	retro_hsap_821	RPLP2	3,86	3,45E-28	8,17E-27	0,75	0,62	12,21	2,10	11,46
ENSG00000214612	retro_hsap_2417	RPS19P1	RPS19	3,89	6,96E-28	1,61E-26	0,84	0,72	13,50	2,43	12,67
ENSG00000213293	retro_hsap_1964	retro_hsap_1964	RPL28	4,75	2,16E-32	6,68E-31	0,57	0,48	18,24	1,86	17,68
ENSG00000176343	retro_hsap_926	RPL37AP8	RPL37A	5,23	3,57E-27	7,85E-26	0,31	0,26	12,04	0,93	11,74
ENSG00000234985	retro_hsap_3895	retro_hsap_3895	RPL18	5,29	1,40E-34	4,97E-33	0,41	0,39	16,04	1,29	15,63
ENSG00000220583	retro_hsap_3579	RPL35P2	RPL35	7,18	2,88E-25	5,56E-24	0,10	0	13,98	0,22	13,89

Tabela 34. Wyniki analizy ekspresji różnicowej retrogenów dla raka jasnokomórkowego nerki (TCGA-KIRC).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000126803	retro_hsap_68	HSPA2	HSPA8	-3,89	8,46E-154	6,69E-151	166,89	87,94	10,91	7,15	-155,98
ENSG00000182107	retro_hsap_12	TMEM30B	TMEM30A	-2,85	3,68E-34	4,63E-33	53,20	39,23	6,44	2,22	-46,76

ENSG00000258626	retro_hsap_1434	COX7A2P1	COX7A2	-2,09	1,28E-19	7,61E-19	13,19	7,29	3,06	1,05	-10,13
ENSG00000198743	retro_hsap_105	SLC5A3	SLC5A1	-1,93	1,13E-38	1,74E-37	116,33	75,75	30,66	19,82	-85,67
ENSG00000174599	retro_hsap_51	TRAM1L1	TRAM1	-1,51	1,55E-51	4,16E-50	22,02	17,24	8,15	7,10	-13,87
ENSG00000137124	retro_hsap_19	ALDH1B1	ALDH2	-1,28	8,67E-22	5,82E-21	95,46	88,82	46,22	33,85	-49,24
ENSG00000179277	retro_hsap_1718	MEIS3P1	MEIS3	-1,08	6,94E-20	4,21E-19	23,33	22,93	12,38	9,91	-10,95
zwiększona ekspresja											
ENSG00000263266	retro_hsap_1737	RPS7P1	RPS7	0,61	5,57E-09	1,61E-08	28,67	28,68	46,35	41,13	17,68
ENSG00000226084	retro_hsap_172	retro_hsap_172	RPL17	0,70	7,24E-15	3,20E-14	86,54	82,89	151,87	136,14	65,33
ENSG00000218426	retro_hsap_3682	retro_hsap_3682	RPL27A	0,81	1,54E-11	5,42E-11	13,56	10,59	23,57	18,48	10,01
ENSG00000237550	retro_hsap_1605	retro_hsap_1605	RPL9	0,82	8,86E-09	2,52E-08	287,57	273,67	511,26	468,59	223,69
ENSG00000143878	retro_hsap_108	RHOB	RHOA	0,87	2,00E-13	8,06E-13	583,79	524,88	1191,50	955,57	607,70
ENSG00000234518	retro_hsap_417	PTGES3P1	PTGES3	0,90	1,55E-23	1,15E-22	14,92	13,72	29,53	25,33	14,61
ENSG00000242299	retro_hsap_2820	retro_hsap_2820	RPS18	0,99	6,64E-22	4,49E-21	13,36	11,49	27,61	24,15	14,25
ENSG00000179967	retro_hsap_3101	PPP1R14BP3	PPP1R14B	1,00	2,60E-16	1,26E-15	27,47	26,18	50,71	41,52	23,24
ENSG00000177105	retro_hsap_28	RHOG	RAC1	1,01	2,96E-47	6,58E-46	48,47	43,75	99,27	96,26	50,80
ENSG00000237039	retro_hsap_2220	RPS28P4	RPS28	1,01	1,60E-17	8,40E-17	34,19	31,36	73,41	59,28	39,22
ENSG00000227097	retro_hsap_799	RPS28P7	RPS28	1,07	4,08E-08	1,10E-07	437,71	401,63	942,82	865,30	505,11
ENSG00000213442	retro_hsap_1053	RPL18AP3	RPL18A	1,08	2,10E-19	1,24E-18	19,67	18,29	42,54	35,03	22,87
ENSG00000240342	retro_hsap_1060	RPS2P5	RPS2	1,17	7,42E-25	5,89E-24	54,52	48,22	128,36	113,37	73,84
ENSG00000233762	retro_hsap_2341	RPS15P4	RPS15	1,18	1,47E-18	8,21E-18	14,19	13,23	31,47	25,03	17,27
ENSG00000185641	retro_hsap_3371	retro_hsap_3371	RPS25	1,22	2,25E-20	1,40E-19	16,73	15,47	38,54	29,32	21,81
ENSG00000136149	retro_hsap_1268	RPL13AP25	RPL13A	1,23	2,97E-19	1,73E-18	7,81	7,29	19,93	13,82	12,12
ENSG00000220842	retro_hsap_738	RPL21P16	RPL21	1,35	2,80E-15	1,27E-14	6,74	5,86	16,82	10,64	10,09
ENSG00000189343	retro_hsap_1812	RPS2P46	RPS2	1,36	2,00E-29	2,01E-28	8,37	7,99	22,21	17,85	13,84
ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	1,39	1,13E-29	1,15E-28	25,24	21,25	65,56	54,05	40,31
ENSG00000215030	retro_hsap_1810	RPL13P12	RPL13	1,45	1,50E-20	9,46E-20	22,46	19,82	60,26	39,75	37,80
ENSG00000230953	retro_hsap_154	retro_hsap_154	TMA7	1,56	4,71E-20	2,88E-19	3,91	3,57	14,32	8,25	10,42
ENSG00000228232	retro_hsap_4806	GAPDHP1	GAPDH	1,70	7,24E-15	3,21E-14	11,69	5,94	38,12	11,66	26,44
ENSG00000233913	retro_hsap_3265	RPL10P9	RPL10	1,78	7,55E-12	2,72E-11	13,10	3,60	48,73	7,73	35,63
ENSG00000213144	retro_hsap_1061	retro_hsap_1061	PDAP1	2,36	1,50E-58	5,18E-57	4,35	2,94	24,29	18,72	19,93

Tabela 35. Wyniki analizy ekspresji różnicowej retrogenów dla raka wątrobowokomórkowego (TCGA-LIHC).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,49	3,76E-22	9,48E-21	969,38	876,57	567,80	444,25	-401,58
ENSG00000137124	retro_hsap_19	ALDH1B1	ALDH2	-1,36	1,63E-17	2,56E-16	173,35	153,92	114,32	83,61	-59,03
zwiększona ekspresja											
ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	0,59	0,000852	0,002117	22,68	20,25	55,63	37,94	32,96
ENSG00000198618	retro_hsap_2497	PPIAP22	PPIA	0,62	7,97E-15	9,30E-14	47,36	46,01	127,13	112,09	79,77
ENSG00000182952	retro_hsap_100	HMGNA4	HMGNA2	0,66	3,82E-16	5,15E-15	15,91	14,21	45,90	39,14	29,99
ENSG00000242299	retro_hsap_2820	retro_hsap_2820	RPS18	0,67	2,81E-05	9,05E-05	7,71	7,41	21,04	13,34	13,33
ENSG00000218426	retro_hsap_3682	retro_hsap_3682	RPL27A	0,72	2,43E-06	9,28E-06	9,38	8,06	26,31	20,06	16,93
ENSG00000137970	retro_hsap_426	RPL7P9	RPL7	0,72	0,000858	0,002131	13,51	12,44	33,60	20,72	20,09
ENSG00000179967	retro_hsap_3101	PPP1R14BP3	PPP1R14B	0,83	8,65E-09	4,71E-08	19,05	17,81	56,78	43,25	37,73
ENSG00000218175	retro_hsap_2348	retro_hsap_2348	RPLP1	0,97	8,17E-08	3,89E-07	5,17	4,52	17,91	11,04	12,74
ENSG00000228232	retro_hsap_4806	GAPDHP1	GAPDH	1,10	5,60E-05	0,000171	7,61	7,46	25,66	12,69	18,05
ENSG00000175793	retro_hsap_8	SFN	YWHAQ	3,86	9,01E-34	6,64E-32	3,07	0,79	72,65	25,37	69,58
ENSG00000170516	retro_hsap_16	COX7B2	COX7B	9,23	2,94E-39	3,21E-37	0,02	0	20,55	0,14	20,53

Tabela 36. Wyniki analizy ekspresji różnicowej retrogenów dla raka płaskonabłonkowy płuc (TCGA-LUSC).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000231322	retro_hsap_3752	RPL13AP17	RPL13A	-4,51	1,71E-35	2,26E-34	15,30	13,96	0,74	0,14	-14,56
ENSG00000213144	retro_hsap_1061	retro_hsap_1061	PDAP1	-2,49	2,21E-61	7,59E-60	13,06	10,59	2,55	1,85	-10,51
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,79	2,86E-42	4,94E-41	809,52	762,43	250,39	199,41	-559,14
ENSG00000196205	retro_hsap_4220	EEF1A1P5	EEF1A1	-1,67	5,17E-22	3,63E-21	257,64	224,71	84,72	58,26	-172,92

ENSG00000233476	retro_hsap_3825	EEF1A1P6	EEF1A1	-1,67	1,24E-24	9,86E-24	101,44	87,59	34,31	22,73	-67,13
ENSG00000179277	retro_hsap_1718	MEIS3P1	MEIS3	-1,35	8,41E-20	5,24E-19	20,06	17,93	8,66	7,02	-11,40
ENSG00000177105	retro_hsap_28	RHOG	RAC1	-0,63	3,33E-19	2,01E-18	159,85	152,32	109,66	107,02	-50,19
zwiększona ekspresja											
ENSG00000254612	retro_hsap_939	DNAJB6P1	DNAJB6	0,58	1,73E-12	7,04E-12	17,82	17,14	28,27	27,21	10,46
ENSG00000235552	retro_hsap_1905	RPL6P27	RPL6	0,59	1,04E-07	2,97E-07	18,88	18,20	30,41	26,71	11,53
ENSG00000215030	retro_hsap_1810	RPL13P12	RPL13	0,65	8,28E-06	1,99E-05	33,50	26,17	55,06	42,66	21,56
ENSG00000226479	retro_hsap_104	TMEM185B	TMEM185A	0,66	2,67E-18	1,54E-17	27,32	26,01	46,66	44,88	19,34
ENSG00000283041	retro_hsap_3906	retro_hsap_3906	EEF1G	0,70	2,53E-08	7,62E-08	20,98	21,67	36,07	30,36	15,08
ENSG00000223612	retro_hsap_222	retro_hsap_222	VOPP1	0,75	1,08E-07	3,08E-07	13,20	13,23	24,03	17,51	10,83
ENSG00000137124	retro_hsap_19	ALDH1B1	ALDH2	0,77	1,36E-11	5,20E-11	23,73	22,25	43,74	37,83	20,02
ENSG00000218426	retro_hsap_3682	retro_hsap_3682	RPL27A	0,77	2,26E-09	7,39E-09	22,56	20,51	39,98	33,70	17,42
ENSG00000132661	retro_hsap_80	NXT1	NXT2	0,80	1,68E-16	8,71E-16	50,27	43,98	93,67	83,88	43,40
ENSG00000198618	retro_hsap_2497	PPIAP22	PPIA	0,81	3,22E-16	1,65E-15	118,84	114,64	215,75	200,77	96,91
ENSG00000233762	retro_hsap_2341	RPS15P4	RPS15	0,81	9,39E-10	3,16E-09	25,02	21,42	47,14	40,34	22,12
ENSG00000137970	retro_hsap_426	RPL7P9	RPL7	0,81	2,04E-12	8,24E-12	18,86	16,95	34,46	30,20	15,60
ENSG00000240342	retro_hsap_1060	RPS2P5	RPS2	0,82	1,72E-12	6,99E-12	87,06	86,92	164,73	154,12	77,68
ENSG00000224631	retro_hsap_1694	RPS27AP16	RPS27A	0,89	5,65E-14	2,53E-13	19,06	17,05	37,41	31,64	18,36
ENSG00000244313	retro_hsap_869	retro_hsap_869	RPS10	0,89	2,09E-08	6,35E-08	36,00	30,78	69,23	51,40	33,23
ENSG00000189343	retro_hsap_1812	RPS2P46	RPS2	0,90	2,72E-12	1,09E-11	12,91	10,71	25,90	21,22	12,99
ENSG00000181524	retro_hsap_3583	RPL24P4	RPL24	0,91	2,25E-10	7,93E-10	11,68	11,51	22,66	17,95	10,98
ENSG00000235655	retro_hsap_2180	H3P6	H3-3B	0,92	5,33E-11	1,96E-10	21,30	16,54	43,04	35,79	21,74
ENSG00000242299	retro_hsap_2820	retro_hsap_2820	RPS18	1,05	2,75E-13	1,18E-12	22,11	19,41	46,87	35,75	24,76
ENSG00000213442	retro_hsap_1053	RPL18AP3	RPL18A	1,09	1,48E-15	7,29E-15	36,69	31,99	81,24	65,20	44,55
ENSG00000186940	retro_hsap_4184	CHCHD2P9	CHCHD2	1,17	4,05E-12	1,60E-11	13,01	11,57	30,77	24,11	17,76
ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	1,19	3,82E-21	2,56E-20	43,31	41,70	105,72	89,16	62,41
ENSG00000218175	retro_hsap_2348	retro_hsap_2348	RPLP1	1,29	5,70E-14	2,55E-13	16,34	14,47	42,05	26,47	25,70
ENSG00000215006	retro_hsap_3186	CHCHD2P2	CHCHD2	1,31	2,90E-14	1,32E-13	6,60	6,20	17,28	13,07	10,68
ENSG00000213326	retro_hsap_1837	RPS7P11	RPS7	1,37	1,50E-14	6,94E-14	6,00	4,73	16,11	10,29	10,11
ENSG00000213553	retro_hsap_2095	RPLP0P6	RPLP0	1,40	1,12E-17	6,22E-17	14,31	11,52	39,37	26,53	25,06
ENSG00000174028	retro_hsap_4657	FAM3C2P	FAM3C	1,42	2,23E-29	2,24E-28	10,37	10,12	29,33	23,80	18,96

ENSG00000263266	retro_hsap_1737	RPS7P1	RPS7	1,51	4,18E-18	2,37E-17	22,16	17,43	65,36	44,63	43,19
ENSG00000264350	retro_hsap_1943	SNRPGP2	SNRPG	1,71	6,49E-44	1,19E-42	16,60	14,18	57,01	47,67	40,41
ENSG00000179967	retro_hsap_3101	PPP1R14BP3	PPP1R14B	1,98	1,97E-72	9,35E-71	28,82	27,01	119,02	107,39	90,20
ENSG00000234964	retro_hsap_880	FABP5P7	FABP5	2,03	4,82E-13	2,03E-12	3,96	3,35	15,11	5,25	11,15
ENSG00000226415	retro_hsap_171	TPI1P1	TPI1	2,07	4,01E-48	8,64E-47	3,18	2,64	14,32	11,41	11,14
ENSG00000228232	retro_hsap_4806	GAPDHP1	GAPDH	2,34	1,71E-22	1,23E-21	8,42	8,96	43,29	16,04	34,87
ENSG00000175793	retro_hsap_8	SFN	YWHAQ	2,74	1,22E-51	2,95E-50	299,69	203,59	2123,46	1767,16	1823,77
ENSG00000232383	retro_hsap_3710	retro_hsap_3710	RPL23	5,51	8,44E-43	1,48E-41	0,28	0,21	13,57	0,85	13,29
ENSG00000267056	retro_hsap_2032	retro_hsap_2032	IPMK	6,42	1,67E-59	5,33E-58	0,43	0,25	41,21	3,78	40,78
ENSG00000256812	retro_hsap_98	CAPNS2	CAPNS1	6,55	5,61E-108	6,91E-106	0,65	0,52	65,81	31,48	65,16
ENSG00000178372	CALML5	CALML5	CALM3	8,01	1,43E-77	7,93E-76	0,46	0,16	132,51	7,81	132,05
ENSG00000178363	retro_hsap_59	CALML3	CALM3	8,52	2,45E-146	5,81E-144	0,04	0,02	15,43	10,91	15,39

Tabela 37. Wyniki analizy ekspresji różnicowej retrogenów dla gruczolakoraka prostaty (TCGA-PRAD).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000213144	retro_hsap_1061	retro_hsap_1061	PDAP1	-1,54	1,41E-18	2,95E-17	17,69	12,98	5,80	4,31	-11,89
ENSG00000231991	retro_hsap_4150	ANXA2P2	ANXA2	-1,46	3,70E-28	2,65E-26	17,17	16,21	6,43	4,93	-10,74
ENSG00000179277	retro_hsap_1718	MEIS3P1	MEIS3	-1,34	5,49E-20	1,40E-18	21,42	18,55	8,67	8,05	-12,75
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,10	9,67E-18	1,81E-16	1161,98	1165,76	559,62	451,81	-602,36
ENSG00000255112	retro_hsap_75	CHMP1B	CHMP1A	-0,64	9,35E-29	7,13E-27	71,13	66,92	49,90	49,55	-21,23
ENSG00000137124	retro_hsap_19	ALDH1B1	ALDH2	-0,60	4,32E-12	3,97E-11	111,00	91,08	72,91	67,95	-38,09
ENSG00000182952	retro_hsap_100	HMGNA4	HMGNA2	-0,60	2,74E-35	3,93E-33	146,77	141,81	95,91	96,49	-50,86
zwiększona ekspresja											
ENSG00000218426	retro_hsap_3682	retro_hsap_3682	RPL27A	0,60	2,68E-06	1,08E-05	38,66	33,02	57,96	52,23	19,31
ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	0,68	3,34E-07	1,55E-06	73,42	65,69	112,56	97,03	39,14

ENSG00000137970	retro_hsap_426	RPL7P9	RPL7	0,69	1,74E-06	7,22E-06	34,65	31,98	52,23	41,37	17,58
ENSG00000237039	retro_hsap_2220	RPS28P4	RPS28	0,72	1,59E-05	5,62E-05	100,61	78,60	150,66	115,17	50,05
ENSG00000215030	retro_hsap_1810	RPL13P12	RPL13	0,75	1,69E-06	7,01E-06	52,92	46,69	83,28	62,61	30,36
ENSG00000218175	retro_hsap_2348	retro_hsap_2348	RPLP1	0,77	4,35E-06	1,69E-05	21,28	18,69	37,84	23,32	16,56
ENSG00000136149	retro_hsap_1268	RPL13AP25	RPL13A	0,90	1,55E-09	1,01E-08	32,64	29,61	56,03	44,97	23,39
ENSG00000179362	retro_hsap_1492	HMG2N2P46	HMG2N2	1,01	3,98E-06	1,56E-05	102,06	71,58	182,53	111,28	80,47
ENSG00000197149	retro_hsap_754	retro_hsap_754	RPS2	1,05	1,50E-13	1,65E-12	11,94	13,19	24,33	22,63	12,39
ENSG00000189343	retro_hsap_1812	RPS2P46	RPS2	1,08	5,10E-13	5,26E-12	15,73	14,50	31,78	22,47	16,04
ENSG00000240342	retro_hsap_1060	RPS2P5	RPS2	1,08	2,60E-19	6,00E-18	110,78	114,25	228,54	218,73	117,76
ENSG00000179967	retro_hsap_3101	PPP1R14BP3	PPP1R14B	1,21	5,72E-20	1,46E-18	31,90	26,76	74,77	67,33	42,86
ENSG00000217275	retro_hsap_3436	retro_hsap_3436	RPS10	1,69	1,28E-14	1,60E-13	7,42	5,56	21,50	11,72	14,07
ENSG00000242899	retro_hsap_2837	RPL7P16	RPL7	1,99	3,37E-25	1,65E-23	90,57	41,32	389,55	338,78	298,97

Tabela 38. Wyniki analizy ekspresji różnicowej retrogenów dla gruczolakoraka odbytnicy (TCGA-READ).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000213144	retro_hsap_1061	retro_hsap_1061	PDAP1	-2,72	2,55E-15	2,17E-13	14,39	12,94	2,26	1,67	-12,14
ENSG00000126803	retro_hsap_68	HSPA2	HSPA8	-1,80	0,000166	0,001216	54,90	45,54	15,85	8,24	-39,05
ENSG00000255112	retro_hsap_75	CHMP1B	CHMP1A	-1,12	7,30E-11	2,50E-09	100,91	96,46	50,89	47,28	-50,02
ENSG00000182107	retro_hsap_12	TMEM30B	TMEM30A	-1,06	6,30E-09	1,42E-07	85,14	80,06	45,66	42,00	-39,48
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,01	0,000238	0,00167	597,90	587,43	304,07	264,33	-293,84
ENSG00000165502	retro_hsap_2	RPL36AL	RPL36A	-0,71	9,51E-05	0,000746	905,32	835,65	609,62	584,93	-295,70
ENSG00000165527	retro_hsap_36	ARF6	ARF3	-0,65	2,90E-06	3,39E-05	186,43	173,28	133,89	123,26	-52,53
zwiększona ekspresja											
ENSG00000226479	retro_hsap_104	TMEM185B	TMEM185A	0,72	1,45E-08	3,01E-07	19,10	18,67	34,48	31,90	15,38
ENSG00000264350	retro_hsap_1943	SNRPGP2	SNRPG	0,95	0,000583	0,003638	21,87	17,68	46,93	39,97	25,06
ENSG00000240342	retro_hsap_1060	RPS2P5	RPS2	1,11	0,000204	0,001454	130,09	144,59	305,89	274,98	175,80

ENSG00000244398	retro_hsap_852	retro_hsap_852	RPL36A	1,20	0,000592	0,003683	62,79	50,52	154,36	123,05	91,57
ENSG00000132661	retro_hsap_80	NXT1	NXT2	1,32	1,31E-07	2,13E-06	38,66	36,17	105,30	95,59	66,64
ENSG00000233762	retro_hsap_2341	RPS15P4	RPS15	1,34	1,50E-05	0,000146	25,53	21,76	67,83	55,44	42,30
ENSG00000242299	retro_hsap_2820	retro_hsap_2820	RPS18	1,37	0,000812	0,004824	41,37	26,32	113,14	70,03	71,77
ENSG00000223803	retro_hsap_2739	RPS20P14	RPS20	1,42	6,14E-05	0,000508	5,77	6,84	16,05	12,06	10,28
ENSG00000215006	retro_hsap_3186	CHCHD2P2	CHCHD2	1,47	0,000957	0,005576	8,86	8,23	24,70	16,03	15,84
ENSG00000228929	retro_hsap_388	RPS13P2	RPS13	1,57	0,000132	0,000994	5,35	4,49	16,77	9,66	11,42
ENSG00000230629	retro_hsap_4697	RPS23P8	RPS23	1,57	3,53E-05	0,000314	11,13	8,32	35,19	21,83	24,06
ENSG00000230202	retro_hsap_3511	retro_hsap_3511	RPL29	1,63	0,000106	0,000825	45,98	19,71	71,40	41,58	25,42
ENSG00000234782	retro_hsap_4307	TPT1P9	TPT1	1,76	0,000155	0,001147	3,88	3,46	13,93	7,08	10,05
ENSG00000146677	retro_hsap_3726	retro_hsap_3726	RPL32	1,95	1,30E-05	0,000129	11,57	9,80	45,82	23,73	34,25
ENSG00000228205	retro_hsap_2736	retro_hsap_2736	RPS3	2,05	4,35E-06	4,86E-05	4,35	3,76	18,76	9,58	14,40
ENSG00000169564	retro_hsap_83	PCBP1	PCBP2	2,09	8,96E-05	0,000707	10,23	7,91	42,94	12,67	32,71
ENSG00000218175	retro_hsap_2348	retro_hsap_2348	RPLP1	2,36	5,10E-06	5,60E-05	21,30	20,05	112,19	42,59	90,89
ENSG00000226396	retro_hsap_119	retro_hsap_119	RPS14	2,54	8,32E-11	2,81E-09	2,84	2,72	17,37	12,00	14,53
ENSG00000233913	retro_hsap_3265	RPL10P9	RPL10	2,54	2,19E-05	0,000205	33,38	8,38	76,39	16,14	43,01
ENSG00000239559	retro_hsap_887	RPL37P2	RPL37	2,54	5,81E-08	1,02E-06	2,48	2,08	14,59	8,02	12,12
ENSG00000232888	retro_hsap_1068	RPS11P5	RPS11	2,56	3,34E-06	3,85E-05	7,00	6,05	41,84	12,61	34,84
ENSG00000215030	retro_hsap_1810	RPL13P12	RPL13	2,65	1,12E-07	1,86E-06	42,92	38,53	265,61	127,95	222,69
ENSG00000226608	retro_hsap_2407	FTLP3	FTL	2,72	5,23E-08	9,39E-07	8,00	7,14	55,11	20,65	47,11
ENSG00000180211	retro_hsap_3581	retro_hsap_3581	RPL23	2,73	1,68E-07	2,65E-06	2,89	2,59	18,91	7,81	16,02
ENSG00000235297	retro_hsap_1952	FAUP1	FAU	2,84	3,90E-06	4,41E-05	2,59	2,10	19,06	4,74	16,47
ENSG00000213315	retro_hsap_1371	retro_hsap_1371	RPS18	3,34	8,13E-08	1,39E-06	3,46	2,43	36,00	8,79	32,54
ENSG00000239528	retro_hsap_3232	RPS14P8	RPS14	3,81	1,63E-07	2,59E-06	1,27	0,79	18,15	5,43	16,88
ENSG00000240950	retro_hsap_2851	retro_hsap_2851	RPL8	3,82	1,91E-09	4,81E-08	1,16	1,06	16,42	3,40	15,27
ENSG00000213293	retro_hsap_1964	retro_hsap_1964	RPL28	3,83	1,42E-07	2,28E-06	0,96	0,90	13,75	2,07	12,79
ENSG00000240652	retro_hsap_821	retro_hsap_821	RPLP2	3,91	3,48E-09	8,30E-08	0,82	0,88	12,30	2,28	11,48
ENSG00000214612	retro_hsap_2417	RPS19P1	RPS19	3,97	4,49E-09	1,05E-07	0,85	0,75	13,90	2,58	13,04
ENSG00000213058	retro_hsap_268	retro_hsap_268	RPS14	4,71	3,52E-07	5,15E-06	5,39	0,94	137,91	11,03	132,52
ENSG00000240668	retro_hsap_2842	KRT8P36	KRT8	4,91	2,77E-09	6,74E-08	0,93	0,51	27,65	1,60	26,72
ENSG00000234985	retro_hsap_3895	retro_hsap_3895	RPL18	5,25	1,69E-10	5,36E-09	0,44	0,46	16,81	1,29	16,37

ENSG00000176343	retro_hsap_926	RPL37AP8	RPL37A	5,75	1,19E-09	3,16E-08	0,24	0,19	12,40	1,04	12,16
ENSG00000220583	retro_hsap_3579	RPL35P2	RPL35	7,01	8,91E-08	1,51E-06	0,10	0,07	12,85	0,26	12,74

Tabela 39. Wyniki analizy ekspresji różnicowej retrogenów dla raka endometrium trzonu macicy (TCGA-UCEC).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM tkanka okalająca	Mediana TPM tkanka okalająca	Średnia TPM nowotwór	Mediana TPM nowotwór	Różnica TPM
zmniejszona ekspresja											
ENSG00000197406	retro_hsap_65	DIO3	DIO1	-2,76	7,10E-14	6,74E-13	12,20	9,33	1,44	0,48	-10,77
ENSG00000253520	retro_hsap_3408	retro_hsap_3408	PTS	-1,75	8,91E-14	8,37E-13	16,60	16,39	4,35	2,62	-12,25
ENSG00000137124	retro_hsap_19	ALDH1B1	ALDH2	-1,57	6,91E-34	2,93E-32	132,94	73,46	42,18	36,99	-90,76
ENSG00000143878	retro_hsap_108	RHOB	RHOA	-1,53	4,23E-20	6,98E-19	973,94	763,35	303,43	238,18	-670,51
ENSG00000196205	retro_hsap_4220	EEF1A1P5	EEF1A1	-1,52	1,13E-08	6,11E-08	235,36	135,51	71,97	39,02	-163,39
ENSG00000233476	retro_hsap_3825	EEF1A1P6	EEF1A1	-1,47	8,83E-07	3,72E-06	59,31	45,41	18,75	6,10	-40,56
ENSG00000226084	retro_hsap_172	retro_hsap_172	RPL17	-1,40	0,000109	0,000334	106,15	114,58	38,11	5,17	-68,05
ENSG00000179277	retro_hsap_1718	MEIS3P1	MEIS3	-1,33	1,28E-10	8,65E-10	17,61	16,28	6,71	4,74	-10,90
ENSG00000178700	retro_hsap_6	DHFR2	DHFR	-1,24	8,15E-35	3,64E-33	16,14	16,10	6,12	5,86	-10,02
ENSG00000214389	retro_hsap_3879	RPS3AP26	RPS3A	-1,21	9,94E-09	5,40E-08	18,58	16,66	7,28	4,98	-11,30
ENSG00000187866	retro_hsap_35	FAM122A	FAM122C	-1,16	1,72E-54	2,51E-52	25,12	24,20	10,01	9,85	-15,10
ENSG00000237550	retro_hsap_1605	retro_hsap_1605	RPL9	-1,11	1,47E-06	6,02E-06	904,43	806,86	350,93	276,22	-553,50
ENSG00000145945	retro_hsap_94	FAM50B	FAM50A	-1,07	2,13E-06	8,50E-06	36,57	36,19	15,41	12,48	-21,16
ENSG00000218283	retro_hsap_309	MORF4L1P1	MORF4L1	-1,04	1,99E-11	1,47E-10	28,98	29,32	13,58	12,02	-15,40
ENSG00000233426	retro_hsap_2106	EIF3FP3	EIF3F	-0,97	3,27E-07	1,46E-06	19,02	16,31	8,76	6,01	-10,25
ENSG00000249353	retro_hsap_3355	NPM1P27	NPM1	-0,81	6,88E-07	2,94E-06	22,23	19,19	11,70	9,09	-10,54
ENSG00000182952	retro_hsap_100	HMGNA4	HMGNA2	-0,76	5,06E-16	5,88E-15	162,69	157,95	89,10	79,53	-73,59
ENSG00000139239	retro_hsap_1015	RPL14P1	RPL14	-0,70	1,19E-05	4,25E-05	50,45	42,54	26,57	21,34	-23,88
ENSG00000155876	retro_hsap_4	RRAGA	RRAGB	-0,70	2,64E-09	1,54E-08	285,35	287,08	155,24	151,64	-130,11
ENSG00000224631	retro_hsap_1694	RPS27AP16	RPS27A	-0,70	0,000218	0,000634	49,90	41,90	25,87	18,90	-24,03
ENSG00000188021	retro_hsap_58	UBQLN2	UBQLN1	-0,69	5,68E-16	6,56E-15	60,20	60,48	34,33	32,79	-25,86

zwiększona ekspresja											
ENSG00000228232	retro_hsap_4806	GAPDHP1	<i>GAPDH</i>	0,95	0,000491	0,001338	18,08	6,15	31,00	14,15	12,92
ENSG00000179967	retro_hsap_3101	PPP1R14BP3	<i>PPP1R14B</i>	0,97	8,44E-05	0,000264	45,26	32,69	76,45	36,16	31,19
ENSG00000169564	retro_hsap_83	PCBP1	<i>PCBP2</i>	1,06	4,78E-05	0,000156	30,64	13,19	53,87	48,97	23,23
ENSG00000239559	retro_hsap_887	RPL37P2	<i>RPL37</i>	1,08	5,56E-05	0,000179	21,50	8,52	36,10	24,06	14,59
ENSG00000213315	retro_hsap_1371	retro_hsap_1371	<i>RPS18</i>	1,10	1,36E-05	4,85E-05	20,29	9,34	34,46	27,19	14,17
ENSG00000226608	retro_hsap_2407	FTLP3	<i>FTL</i>	1,20	5,07E-08	2,52E-07	17,74	10,95	33,66	23,24	15,92
ENSG00000235655	retro_hsap_2180	H3P6	<i>H3-3B</i>	1,35	9,80E-14	9,17E-13	26,02	24,72	59,12	49,58	33,10
ENSG00000264350	retro_hsap_1943	SNRPGP2	<i>SNRPG</i>	1,35	4,16E-32	1,58E-30	19,00	17,85	43,02	41,58	24,03
ENSG00000226976	retro_hsap_25	COX6A1P2	<i>COX6A1</i>	1,46	1,74E-13	1,60E-12	10,48	7,18	23,56	16,79	13,08
ENSG00000213293	retro_hsap_1964	retro_hsap_1964	<i>RPL28</i>	2,21	1,86E-12	1,53E-11	4,76	1,15	16,25	11,15	11,49
ENSG00000234685	retro_hsap_1174	NUS1P2	<i>NUS1</i>	2,26	2,70E-08	1,39E-07	5,65	1,35	21,77	5,93	16,12
ENSG00000240668	retro_hsap_2842	KRT8P36	<i>KRT8</i>	4,02	3,26E-24	7,45E-23	1,18	0,17	15,03	4,65	13,86
ENSG00000175793	retro_hsap_8	SFN	<i>YWHAQ</i>	4,81	1,00E-61	2,02E-59	22,95	9,78	530,32	269,98	507,37
ENSG00000178372	CALML5	CALML5	<i>CALM3</i>	5,98	4,02E-21	7,23E-20	0,14	0	19,03	0,36	18,89

Tabela 40. Wyniki analizy ekspresji różnicowej retrogenów podczas rozwoju guza (stadium IV vs stadium I) (TCGA-KIRC, TCGA-LIHC).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Stadium I średnia TPM	Stadium I mediana TPM	Stadium IV średnia TPM	Stadium IV mediana TPM	Różnica TPM
KIRC											
ENSG00000143878	retro_hsap_108	RHOB	<i>RHOA</i>	-0,72	1,56E-09	7,39E-08	1402,17	1269,00	751,92	556,27	-650,25
ENSG00000179967	retro_hsap_3101	PPP1R14BP3	<i>PPP1R14B</i>	0,77	6,95E-11	5,00E-09	42,96	37,59	70,36	58,06	27,40
LIHC											
ENSG00000242299	retro_hsap_2820	retro_hsap_2820	<i>RPS18</i>	1,89	2,40E-05	0,018909434	16,18	11,93	72,14	21,13	55,97

Tabela 41. Wyniki analizy ekspresji różnicowej retrogenów pomiędzy próbkami o odmiennej odpowiedzi na zastosowane leczenie (postępujący nowotwór vs remisja) (TCGA-LUSC, TCGA-OV, TCGA-UCEC).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Remisja średnia PM	Remisja mediana TPM	Postępujący nowotwór średnia TPM	Postępujący nowotwór mediana TPM	Różnica TPM
LUSC											
ENSG00000233913	retro_hsap_3265	RPL10P9	<i>RPL10</i>	-1,70	3,82E-06	0,001447	34,19	9,78	18,13	6,59	-16,06
OV											
ENSG00000233913	retro_hsap_3265	RPL10P9	<i>RPL10</i>	1,19	0,000206	0,036024	119,78	45,43	256,23	63,51	136,45
UCEC											
ENSG00000179967	retro_hsap_3101	PPP1R14BP3	<i>PPP1R14B</i>	1,30	0,000218	0,015556	67,77	34,60	163,51	118,93	95,74
ENSG00000227097	retro_hsap_799	RPS28P7	<i>RPS28</i>	2,07	0,000166	0,0131	631,21	99,49	2236,66	1112,69	1605,45

Tabela 42. Korelacja ekspresji retrogenów i genów rodzicielskich w badanych nowotworach oraz tkankach okalających.

rho – współczynnik korelacji ekspresji, pvalue – wartość p, BRCA – rak piersi, COAD – gruczolakorak jelita grubego, GBM – glejak wielopostaciowy, KIRC – rak jasnokomórkowy nerki, LIHC – rak wątrobowokomórkowy, LUSC- rak płaskonabłonkowy płuc, PRAD – gruczolakorak prostaty, READ – gruczolakorak odbytnicy, UCEC – rak endometrium trzonu macicy

Retrogen	Gen rodzicielski	LIHC próby zdrowe		LIHC próby nowotworowe		COAD próby zdrowe		COAD próby nowotworowe	
		rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue
HMGNA4	HMGNA2	0,75	3,22E-10	0,61	3,21E-40	NA	NA	0,41	6,87E-21
RPL14P1	RPL14	0,72	4,25E-09	0,69	9,60E-55	NA	NA	0,5	8,12E-31
RHOB	RHOA	0,38	6,05E-03	NA	NA	NA	NA	NA	NA
retro_hsap_119	RPS14	0,4	4,45E-03	0,48	8,87E-23	NA	NA	0,53	1,06E-35
TMEM30B	TMEM30A	0,53	7,13E-05	NA	NA	NA	NA	0,27	1,69E-09
CAPNS2	CAPNS1	NA	NA	NA	NA	NA	NA	NA	NA
HMGNA2P46	HMGNA2	NA	NA	0,39	1,13E-14	NA	NA	NA	NA
DNAJB6P1	DNAJB6	0,67	1,06E-07	0,69	9,28E-54	NA	NA	0,43	9,78E-23
retro_hsap_172	RPL17	0,42	2,52E-03	0,63	4,75E-43	NA	NA	NA	NA
RPS7P1	RPS7	NA	NA	0,72	3,86E-62	0,6	4,04E-05	0,45	1,25E-24
RPL13P12	RPL13	0,41	3,40E-03	0,73	1,83E-62	0,78	1,72E-09	0,73	8,32E-81
RPL6P27	RPL6	0,63	7,78E-07	0,73	1,12E-62	0,76	9,99E-09	0,78	1,96E-96
HNRNPA1P57	HNRNPA1	NA	NA	NA	NA	NA	NA	NA	NA
EIF3FP3	EIF3F	0,32	2,48E-02	0,55	4,07E-31	NA	NA	NA	NA
H3P6	H3-3B	NA	NA	0,25	1,08E-06	NA	NA	NA	NA
retro_hsap_222	VOPP1	0,58	1,04E-05	0,65	3,99E-46	0,8	2,26E-10	0,66	2,97E-61
RPS15P4	RPS15	0,5	2,45E-04	0,77	8,11E-74	0,86	5,56E-13	0,82	1,27E-117
PPIAP22	PPIA	0,69	2,66E-08	0,81	2,70E-89	NA	NA	0,28	8,85E-10
retro_hsap_268	RPS14	0,28	5,07E-02	NA	NA	NA	NA	NA	NA
KRT8P36	KRT8	0,69	2,19E-08	0,56	1,10E-31	NA	NA	0,31	8,00E-12
RPS14P8	RPS14	NA	NA	NA	NA	NA	NA	0,34	3,31E-14
RPL10P9	RPL10	0,31	2,86E-02	0,35	2,31E-12	NA	NA	0,25	2,16E-08
EEF1A1P13	EEF1A1	0,69	3,35E-08	0,79	1,67E-82	0,63	8,40E-06	0,41	1,24E-20
NPM1P27	NPM1	0,67	7,79E-08	0,81	1,07E-86	0,8	2,81E-10	0,74	2,30E-82
retro_hsap_3408	PTS	NA	NA	NA	NA	NA	NA	NA	NA
retro_hsap_3581	RPL23	0,68	5,63E-08	0,63	7,99E-43	NA	NA	NA	NA
retro_hsap_3710	RPL23	0,4	4,03E-03	0,42	2,85E-17	NA	NA	NA	NA
EEF1A1P6	EEF1A1	0,58	9,07E-06	0,72	6,82E-62	0,55	0,0002	0,45	1,24E-24
TPT1P9	TPT1	0,38	6,16E-03	0,46	3,12E-21	NA	NA	0,47	5,52E-27
GAPDHP1	GAPDH	NA	NA	0,26	2,73E-07	NA	NA	0,37	3,28E-17
UBQLN2	UBQLN1	0,72	3,57E-09	0,64	1,08E-43	0,5	0,00098	0,5	9,57E-32
DHFR2	DHFR	NA	NA	0,4	1,07E-15	NA	NA	0,35	3,48E-15
CHMP1B	CHMP1A	0,57	1,83E-05	0,31	9,61E-10	NA	NA	NA	NA
SFN	YWHAQ	NA	NA	NA	NA	NA	NA	NA	NA
PCBP1	PCBP2	NA	NA	0,32	1,93E-10	NA	NA	NA	NA
FABP5P7	FABP5	0,33	1,87E-02	0,63	3,96E-42	NA	NA	0,26	1,21E-08

Retrogen	Gen rodzicielski	KIRC próby zdrowe		KIRC próby nowotworowe		LUSC próby zdrowe		LUSC próby nowotworowe	
		rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue
HMGN4	HMGN2	0,78	6,95E-16	0,68	5,33E-74	0,6	2,66E-06	0,33	7,81E-14
RPL14P1	RPL14	0,66	1,92E-10	0,61	4,17E-57	0,53	7,13E-05	0,68	2,00E-70
RHOB	RHOA	0,66	2,55E-10	0,45	3,18E-28	NA	NA	NA	NA
retro_hsap_119	RPS14	NA	NA	NA	NA	NA	NA	NA	NA
TMEM30B	TMEM30A	0,77	2,65E-15	0,33	8,79E-15	0,52	0,0001019	NA	NA
CAPNS2	CAPNS1	NA	NA	NA	NA	NA	NA	NA	NA
HMGN2P46	HMGN2	NA	NA	0,34	1,17E-15	NA	NA	NA	NA
DNAJB6P1	DNAJB6	0,85	1,57E-21	0,71	1,86E-83	0,7	1,46E-08	0,8	1,67E-112
retro_hsap_172	RPL17	0,54	7,76E-07	NA	NA	NA	NA	NA	NA
RPS7P1	RPS7	0,69	2,78E-11	0,7	9,83E-81	NA	NA	0,69	1,39E-72
RPL13P12	RPL13	NA	NA	0,63	2,56E-60	NA	NA	0,64	2,08E-58
RPL6P27	RPL6	0,39	0,0007049	0,59	7,48E-52	NA	NA	0,73	1,19E-85
HNRNPA1P57	HNRNPA1	NA	NA	NA	NA	NA	NA	0,26	4,21E-09
EIF3FP3	EIF3F	NA	NA	0,35	2,73E-17	NA	NA	0,42	2,48E-22
H3P6	H3-3B	0,54	8,19E-07	0,34	1,06E-15	NA	NA	NA	NA
retro_hsap_222	VOPP1	0,89	5,48E-26	0,72	7,36E-86	0,53	6,62E-05	0,77	4,10E-100
RPS15P4	RPS15	0,52	2,58E-06	0,72	9,91E-88	NA	NA	0,73	5,36E-84
PPIAP22	PPIA	0,91	2,22E-28	0,81	7,09E-125	0,79	3,68E-12	0,91	1,24E-189
retro_hsap_268	RPS14	NA	NA	NA	NA	NA	NA	NA	NA
KRT8P36	KRT8	0,57	1,64E-07	0,36	5,98E-18	0,53	7,50E-05	0,38	1,71E-18
RPS14P8	RPS14	NA	NA	NA	NA	NA	NA	NA	NA
RPL10P9	RPL10	NA	NA	0,32	1,22E-14	NA	NA	0,31	3,01E-12
EEF1A1P13	EEF1A1	0,73	4,11E-13	0,8	3,12E-121	NA	NA	0,66	8,10E-63
NPM1P27	NPM1	0,62	7,99E-09	0,76	2,28E-103	NA	NA	0,75	9,07E-92
retro_hsap_3408	PTS	0,58	8,62E-08	NA	NA	NA	NA	-0,3	6,54E-12
retro_hsap_3581	RPL23	0,38	0,0008663	0,51	7,10E-37	NA	NA	0,63	1,56E-57
retro_hsap_3710	RPL23	NA	NA	0,29	1,20E-11	NA	NA	0,3	1,17E-11
EEF1A1P6	EEF1A1	0,78	8,90E-16	0,75	9,77E-99	NA	NA	0,69	9,70E-71
TPT1P9	TPT1	0,55	6,83E-07	0,46	1,17E-29	0,48	0,000381	0,53	1,55E-37
GAPDHP1	GAPDH	NA	NA	0,31	2,18E-13	NA	NA	0,42	4,99E-23
UBQLN2	UBQLN1	0,88	1,03E-24	0,8	1,00E-118	0,74	3,95E-10	0,41	4,41E-22
DHFR2	DHFR	NA	NA	0,54	6,50E-42	NA	NA	0,3	1,26E-11
CHMP1B	CHMP1A	0,85	1,68E-21	0,36	4,37E-18	NA	NA	NA	NA
SFN	YWHAQ	0,52	2,49E-06	NA	NA	NA	NA	NA	NA
PCBP1	PCBP2	NA	NA	0,39	3,71E-21	NA	NA	NA	NA
FABP5P7	FABP5	0,6	2,04E-08	0,6	3,80E-54	0,7	1,03E-08	0,93	1,92E-218

Retrogen	Gen rodzicielski	PRAD próby zdrowe		PRAD próby nowotworowe		READ próby zdrowe		READ próby nowotworowe	
		rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue
HMGN4	HMGN2	0,74	2,78E-10	0,4	2,09E-20	NA	NA	0,6	1,23E-17
RPL14P1	RPL14	0,62	1,20E-06	0,59	3,09E-49	0,89	0,000542144	0,61	3,12E-18
RHOB	RHOA	NA	NA	NA	NA	NA	NA	NA	NA
retro_hsap_119	RPS14	NA	NA	0,28	2,82E-10	NA	NA	0,5	5,42E-12
TMEM30B	TMEM30A	0,58	5,66E-06	0,69	1,64E-72	NA	NA	0,33	1,50E-05
CAPNS2	CAPNS1	NA	NA	NA	NA	NA	NA	NA	NA
HMGN2P46	HMGN2	0,65	1,82E-07	NA	NA	NA	NA	NA	NA
DNAJB6P1	DNAJB6	0,79	2,02E-12	0,66	5,35E-63	NA	NA	0,55	2,03E-14
retro_hsap_172	RPL17	0,57	1,01E-05	0,46	3,12E-27	NA	NA	NA	NA
RPS7P1	RPS7	0,72	1,60E-09	0,7	1,07E-75	NA	NA	0,42	1,46E-08
RPL13P12	RPL13	0,64	3,07E-07	0,6	1,80E-49	NA	NA	0,79	2,44E-36
RPL6P27	RPL6	0,76	8,92E-11	0,63	7,81E-58	NA	NA	0,78	8,42E-36
HNRNPA1P57	HNRNPA1	NA	NA	NA	NA	NA	NA	NA	NA
EIF3FP3	EIF3F	NA	NA	NA	NA	NA	NA	NA	NA
H3P6	H3-3B	NA	NA	0,3	5,20E-12	NA	NA	0,31	4,46E-05
retro_hsap_222	VOPP1	0,74	5,28E-10	0,51	2,42E-34	NA	NA	0,79	3,90E-37
RPS15P4	RPS15	0,81	3,89E-13	0,46	2,97E-28	NA	NA	0,81	1,12E-40
PPIAP22	PPIA	0,85	9,54E-16	0,75	7,35E-91	NA	NA	0,26	0,000737495
retro_hsap_268	RPS14	NA	NA	NA	NA	NA	NA	0,28	0,000194682
KRT8P36	KRT8	NA	NA	0,33	7,18E-14	NA	NA	0,34	7,80E-06
RPS14P8	RPS14	NA	NA	NA	NA	NA	NA	0,46	3,52E-10
RPL10P9	RPL10	NA	NA	NA	NA	0,92	0,000204472	0,35	4,48E-06
EEF1A1P13	EEF1A1	0,57	1,15E-05	0,7	7,05E-74	0,93	0,000112035	0,54	7,75E-14
NPM1P27	NPM1	0,74	4,63E-10	0,72	2,92E-80	0,9	0,000343612	0,82	1,55E-41
retro_hsap_3408	PTS	NA	NA	-0,28	2,56E-10	NA	NA	NA	NA
retro_hsap_3581	RPL23	NA	NA	0,46	4,23E-27	0,88	0,000813862	NA	NA
retro_hsap_3710	RPL23	NA	NA	0,33	4,93E-14	NA	NA	NA	NA
EEF1A1P6	EEF1A1	0,52	7,30E-05	0,65	9,62E-62	0,96	7,32E-06	0,42	1,40E-08
TPT1P9	TPT1	NA	NA	0,64	6,98E-59	0,94	5,48E-05	0,36	1,85E-06
GAPDHP1	GAPDH	NA	NA	0,27	5,46E-10	NA	NA	0,37	6,00E-07
UBQLN2	UBQLN1	0,86	1,70E-16	0,87	5,42E-152	NA	NA	0,33	1,13E-05
DHFR2	DHFR	0,58	6,87E-06	0,44	1,02E-24	NA	NA	0,41	2,66E-08
CHMP1B	CHMP1A	NA	NA	NA	NA	NA	NA	NA	NA
SFN	YWHAQ	NA	NA	NA	NA	NA	NA	NA	NA
PCBP1	PCBP2	NA	NA	NA	NA	NA	NA	NA	NA
FABP5P7	FABP5	0,66	1,21E-07	0,87	8,29E-153	NA	NA	NA	NA

Retrogen	Gen rodzicielski	UCEC próby zdrowe		UCEC próby nowotworowe		BRCA próby zdrowe		GBM próby nowotworowe	
		rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue
HMGN4	HMGN2	NA	NA	0,61	3,15E-57	0,66	1,27E-15	0,56	4,46E-15
RPL14P1	RPL14	0,66	1,38E-05	0,63	8,81E-62	0,79	1,57E-25	0,72	9,46E-29
RHOB	RHOA	NA	NA	0,26	1,02E-09	NA	NA	NA	NA
retro_hsap_119	RPS14	NA	NA	0,29	6,92E-12	0,42	3,23E-06	0,57	1,01E-15
TMEM30B	TMEM30A	NA	NA	0,46	4,22E-30	0,66	2,95E-15	0,29	0,000131964
CAPNS2	CAPNS1	NA	NA	NA	NA	NA	NA	NA	NA
HMGN2P46	HMGN2	NA	NA	NA	NA	0,81	3,60E-27	NA	NA
DNAJB6P1	DNAJB6	0,61	0,00010457	0,64	8,08E-65	0,75	2,40E-21	0,72	3,29E-28
retro_hsap_172	RPL17	NA	NA	NA	NA	0,6	2,02E-12	0,4	6,15E-08
RPS7P1	RPS7	NA	NA	0,62	7,05E-59	0,77	1,19E-23	0,9	1,59E-63
RPL13P12	RPL13	NA	NA	0,62	5,51E-60	0,71	1,31E-18	0,48	4,67E-11
RPL6P27	RPL6	0,72	1,20E-06	0,75	5,80E-99	0,6	2,45E-12	0,85	7,15E-48
HNRNPA1P57	HNRNPA1	NA	NA	NA	NA	NA	NA	0,51	1,32E-12
EIF3FP3	EIF3F	NA	NA	0,5	2,94E-36	0,36	7,14E-05	0,76	1,12E-33
H3P6	H3-3B	NA	NA	0,27	6,97E-11	0,56	8,25E-11	NA	NA
retro_hsap_222	VOPP1	0,74	2,93E-07	0,62	4,22E-60	0,6	1,74E-12	0,89	6,39E-60
RPS15P4	RPS15	NA	NA	0,73	2,76E-94	0,74	3,30E-21	0,9	4,00E-63
PPIAP22	PPIA	NA	NA	0,53	7,67E-41	0,86	2,94E-34	0,9	1,52E-63
retro_hsap_268	RPS14	NA	NA	NA	NA	NA	NA	NA	NA
KRT8P36	KRT8	NA	NA	0,27	1,99E-10	0,62	2,68E-13	NA	NA
RPS14P8	RPS14	NA	NA	NA	NA	NA	NA	0,25	0,000973645
RPL10P9	RPL10	NA	NA	NA	NA	0,33	0,0003406	0,47	9,76E-11
EEF1A1P13	EEF1A1	0,65	2,03E-05	0,71	2,36E-84	0,73	6,95E-20	0,8	9,61E-39
NPM1P27	NPM1	0,78	3,18E-08	0,8	6,88E-122	0,71	8,98E-19	0,78	2,44E-36
retro_hsap_3408	PTS	NA	NA	NA	NA	0,43	2,47E-06	NA	NA
retro_hsap_3581	RPL23	NA	NA	0,39	4,10E-21	0,48	8,34E-08	0,77	2,28E-34
retro_hsap_3710	RPL23	NA	NA	0,34	1,24E-16	0,38	2,96E-05	0,53	1,41E-13
EEF1A1P6	EEF1A1	NA	NA	0,48	3,00E-33	0,69	3,86E-17	0,78	4,41E-36
TPT1P9	TPT1	NA	NA	0,43	3,15E-26	0,49	2,89E-08	0,82	5,71E-42
GAPDHP1	GAPDH	NA	NA	0,46	1,01E-29	NA	NA	0,83	2,42E-43
UBQLN2	UBQLN1	0,63	4,62E-05	0,8	1,12E-121	0,81	5,54E-28	0,58	2,62E-16
DHFR2	DHFR	NA	NA	0,49	3,61E-34	NA	NA	0,31	3,13E-05
CHMP1B	CHMP1A	NA	NA	0,27	7,74E-11	0,48	9,07E-08	NA	NA
SFN	YWHAQ	NA	NA	NA	NA	0,53	2,35E-09	-0,26	0,000704769
PCBP1	PCBP2	NA	NA	NA	NA	NA	NA	NA	NA
FABP5P7	FABP5	0,8	9,95E-09	0,45	1,83E-28	0,86	3,90E-34	0,94	6,20E-83

Retrogen	Gen rodzicielski	OV próby nowotworowe		BRCA próby nowotworowe	
		rho	pvalue	rho	pvalue
HMGN4	HMGN2	NA	NA	NA	NA
RPL14P1	RPL14	0,81	6,24E-100	0,72	2,41E-179
RHOB	RHOA	NA	NA	0,29	3,73E-22
retro_hsap_119	RPS14	0,54	4,52E-34	0,28	1,48E-21
TMEM30B	TMEM30A	0,4	5,37E-18	0,4	4,95E-44
CAPNS2	CAPNS1	NA	NA	NA	NA
HMGN2P46	HMGN2	0,48	3,88E-26	NA	NA
DNAJB6P1	DNAJB6	0,52	1,91E-30	0,72	2,31E-178
retro_hsap_172	RPL17	0,34	3,41E-13	0,43	6,87E-52
RPS7P1	RPS7	0,87	6,84E-131	0,6	8,66E-110
RPL13P12	RPL13	0,65	7,14E-54	0,69	5,57E-157
RPL6P27	RPL6	0,87	1,96E-130	0,62	6,52E-120
HNRNPA1P57	HNRNPA1	NA	NA	NA	NA
EIF3FP3	EIF3F	0,78	4,12E-88	0,38	5,49E-39
H3P6	H3-3B	NA	NA	NA	NA
retro_hsap_222	VOPP1	0,72	6,94E-71	0,66	8,23E-140
RPS15P4	RPS15	0,85	2,05E-120	0,78	1,02E-225
PPIAP22	PPIA	0,86	1,37E-126	0,84	1,31E-292
retro_hsap_268	RPS14	0,26	5,39E-08	NA	NA
KRT8P36	KRT8	0,35	3,81E-14	0,53	1,12E-80
RPS14P8	RPS14	0,31	2,70E-11	NA	NA
RPL10P9	RPL10	0,48	1,53E-25	NA	NA
EEF1A1P13	EEF1A1	0,83	1,10E-112	0,74	3,63E-190
NPM1P27	NPM1	0,79	6,04E-91	0,67	3,72E-145
retro_hsap_3408	PTS	NA	NA	NA	NA
retro_hsap_3581	RPL23	0,81	4,04E-103	0,62	1,37E-118
retro_hsap_3710	RPL23	0,59	3,09E-42	0,39	1,21E-41
EEF1A1P6	EEF1A1	0,78	4,05E-88	0,71	1,15E-167
TPT1P9	TPT1	0,8	1,39E-97	0,52	5,44E-77
GAPDHP1	GAPDH	0,74	7,12E-75	0,41	2,15E-46
UBQLN2	UBQLN1	0,57	1,36E-37	0,62	1,35E-120
DHFR2	DHFR	NA	NA	0,31	1,47E-25
CHMP1B	CHMP1A	0,25	1,72E-07	NA	NA
SFN	YWHAQ	NA	NA	NA	NA
PCBP1	PCBP2	NA	NA	NA	NA
FABP5P7	FABP5	0,85	1,00E-118	0,71	1,81E-172

Tabela 43. Korelacja ekspresji retrogenów i genów gospodarzy w badanych nowotworach oraz tkankach okalających.

rho – współczynnik korelacji ekspresji, pvalue – wartość p, BRCA – rak piersi, COAD – gruczolakorak jelita grubego, GBM – glejak wielopostaciowy, KIRC – rak jasnokomórkowy nerki, LIHC – rak wątrobowokomórkowy, LUSC- rak płaskonabłonkowy płuc, PRAD – gruczolakorak prostaty, READ – gruczolakorak odbytnicy, UCEC – rak endometrium trzonu macicy

Retrogen	Gen gospodarza	LIHC próby zdrowe		LIHC próby nowotworowe		COAD próby zdrowe		COAD próby nowotworowe		KIRC próby zdrowe		KIRC próby nowotworowe		LUSC próby zdrowe		LUSC próby nowotworowe	
		rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue
<i>retro_hsap_1061</i>	<i>HSPB8</i>	0,54	4,60E-05	0,77	7,97E-75	NA	NA	0,44	1,99E-23	0,94	1,37E-33	0,87	7,54E-163	0,74	4,07E-10	0,63	2,18E-56
<i>RPS11P5</i>	<i>CHFR</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>CAPNS2</i>	<i>LPCAT2</i>	0,38	6,20E-03	0,35	6,64E-12	NA	NA	0,36	2,37E-16	0,53	2,13E-06	0,3	1,13E-12	NA	NA	0,26	3,63E-09
<i>retro_hsap_268</i>	<i>RASAL2</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>RPL4P4</i>	<i>MAP3K13</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>RPS14P8</i>	<i>SEMA6A</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>NPM1P27</i>	<i>FAM172A</i>	0,4	3,87E-03	0,41	1,76E-16	NA	NA	0,29	1,39E-10	0,46	5,03E-05	0,55	7,46E-44	NA	NA	NA	NA
<i>RPL13AP17</i>	<i>MAGI2</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>retro_hsap_869</i>	<i>AMBRA1</i>	NA	NA	0,02	0,6682514	NA	NA	-0,28	5,83E-10	NA	NA	NA	NA	NA	NA	NA	NA

Retrogen	Gen gospodarza	PRAD próby zdrowe		PRAD próby nowotworowe		READ próby zdrowe		READ próby nowotworowe		UCEC próby zdrowe		UCEC próby nowotworowe		BRCA próby zdrowe		BRCA próby nowotworowe	
		rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue	rho	pvalue
<i>retro_hsap_1061</i>	<i>HSPB8</i>	0,87	3,40E-17	0,93	7,02E-213	0,92	0,0002	0,38	4,86E-07	0,88	3,40E-12	0,42	1,40E-24	0,73	2,65E-20	0,84	1,68E-290
<i>RPS11P5</i>	<i>CHFR</i>	NA	NA	NA	NA	-0,93	0,0001	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>CAPNS2</i>	<i>LPCAT2</i>	NA	NA	0,39	4,14E-20	NA	NA	0,33	1,78E-05	NA	NA	0,47	5,70E-31	NA	NA	0,4	3,25E-44
<i>retro_hsap_268</i>	<i>RASAL2</i>	NA	NA	NA	NA	NA	NA	-0,29	0,0001316	NA	NA	NA	NA	NA	NA	NA	NA
<i>RPL4P4</i>	<i>MAP3K13</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0,36	0,0001	NA	NA
<i>RPS14P8</i>	<i>SEMA6A</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>NPM1P27</i>	<i>FAM172A</i>	0,63	6,01E-07	0,43	2,40E-24	NA	NA	0,55	1,77E-14	NA	NA	0,5	1,17E-35	0,39	2,39E-05	NA	NA
<i>RPL13AP17</i>	<i>MAGI2</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0,31	1,41E-25
<i>retro_hsap_869</i>	<i>AMBRA1</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0,63	1,07E-13	NA	NA

Retrogen	Gen gospodarza	GBM próby nowotworowe		OV próby nowotworowe	
		rho	pvalue	rho	pvalue
<i>retro_hsap_1061</i>	<i>HSPB8</i>	0,78	1,36E-35	0,4	3,73E-18
<i>RPS11P5</i>	<i>CHFR</i>	NA	NA	NA	NA
<i>CAPNS2</i>	<i>LPCAT2</i>	0,38	4,46E-07	NA	NA
<i>retro_hsap_268</i>	<i>RASAL2</i>	NA	NA	NA	NA
<i>RPL4P4</i>	<i>MAP3K13</i>	NA	NA	NA	NA
<i>RPS14P8</i>	<i>SEMA6A</i>	NA	NA	NA	NA
<i>NPM1P27</i>	<i>FAM172A</i>	NA	NA	0,3	1,43E-10
<i>RPL13AP17</i>	<i>MAGI2</i>	NA	NA	NA	NA
<i>retro_hsap_869</i>	<i>AMBRA1</i>	-0,4	1,01E-07	NA	NA

Tabela 44. Wyniki analizy ekspresji różnicowej retrogenów dla eksperymentu inwazji łożyska (PRJNA562073).

Ensembl ID	RetrogeneDB2 ID	Retrogen	Gen rodzicielski	Log2FC	pvalue	padj	Średnia TPM fibroblasty	Mediana TPM fibroblasty	Średnia TPM trofoblast	Mediana TPM trofoblast	Różnica TPM
zmniejszona ekspresja											
retro_hsap_1849	retro_hsap_1849	<i>retro_hsap_1849</i>	<i>AP1S2</i>	-1,9099	5,55E-67	5,05E-66	106,70	104,71	24,76	26,10	-81,94
ENSG00000215030	retro_hsap_1810	<i>RPL13P12</i>	<i>RPL13</i>	-1,8816	2,85E-70	2,72E-69	97,44	107,79	24,92	26,45	-72,53
ENSG00000233913	retro_hsap_3265	<i>RPL10P9</i>	<i>RPL10</i>	-1,5555	4,45E-57	3,54E-56	54,46	52,70	17,56	16,29	-36,90
ENSG00000236698	retro_hsap_350	<i>EIF1AXP1</i>	<i>EIF1AX</i>	-0,8699	0,000435	0,000898	56,87	49,27	29,51	31,55	-27,36
ENSG00000253683	retro_hsap_3402	<i>retro_hsap_3402</i>	<i>TMA7</i>	-0,8615	4,20E-07	1,05E-06	452,04	461,41	239,86	248,56	-212,18
ENSG00000177105	retro_hsap_28	<i>RHOG</i>	<i>RAC1</i>	-0,6364	3,83E-25	1,73E-24	37,46	38,39	22,34	22,40	-15,12
zwiększona ekspresja											
ENSG00000224543	retro_hsap_1968	<i>SNRPGP15</i>	<i>SNRPG</i>	0,6382	5,36E-05	0,000119	432,35	422,29	655,19	686,01	222,84
ENSG00000256210	retro_hsap_1969	<i>retro_hsap_1969</i>	<i>TMEM167A</i>	0,6759	0,000145	0,000312	167,49	170,68	249,80	251,72	82,30
ENSG00000244398	retro_hsap_852	<i>retro_hsap_852</i>	<i>RPL36A</i>	0,7473	1,72E-65	1,54E-64	1108,72	1075,76	1746,66	1676,71	637,94
ENSG00000169564	retro_hsap_83	<i>PCBP1</i>	<i>PCBP2</i>	0,7777	0	0	153,66	154,83	246,27	242,90	92,62
ENSG00000240342	retro_hsap_1060	<i>RPS2P5</i>	<i>RPS2</i>	0,8568	2,77E-44	1,79E-43	121,30	122,27	203,55	200,07	82,25

ENSG00000220842	retro_hsap_738	RPL21P16	<i>RPL21</i>	0,8785	8,64E-75	8,76E-74	620,04	568,25	1072,39	1047,98	452,35
ENSG00000213442	retro_hsap_1053	RPL18AP3	<i>RPL18A</i>	1,0576	2,41E-20	9,57E-20	186,82	195,88	372,02	394,34	185,20
ENSG00000196205	retro_hsap_4220	EEF1A1P5	<i>EEF1A1</i>	1,2607	6,19E-05	0,000137	27,39	22,45	61,63	49,62	34,24
retro_hsap_2297	retro_hsap_2297	retro_hsap_2297	<i>PPIA</i>	1,3123	7,16E-06	1,68E-05	9,29	9,63	22,82	22,71	13,53
ENSG00000235655	retro_hsap_2180	H3P6	<i>H3-3B</i>	1,3221	5,92E-24	2,59E-23	124,00	122,66	288,31	265,62	164,31
ENSG00000230629	retro_hsap_4697	RPS23P8	<i>RPS23</i>	1,425	3,87E-49	2,71E-48	603,53	527,11	1515,87	1673,13	912,34
ENSG00000233270	retro_hsap_2022	SNRPEP4	<i>SNRPE</i>	1,9498	1,50E-35	8,32E-35	159,73	156,66	592,85	650,68	433,12
ENSG00000224831	retro_hsap_2863	TMEM183B	<i>TMEM183A</i>	2,1359	3,47E-50	2,47E-49	8,41	8,22	33,59	34,42	25,18
ENSG00000237550	retro_hsap_1605	retro_hsap_1605	<i>RPL9</i>	2,5321	1,16E-42	7,28E-42	164,39	116,16	905,28	912,73	740,89
ENSG00000224837	retro_hsap_465	GCSHP5	<i>GCSH</i>	3,2863	7,69E-18	2,87E-17	6,20	5,19	57,84	55,55	51,64
ENSG00000234964	retro_hsap_880	FABP5P7	<i>FABP5</i>	3,8521	2,60E-28	1,25E-27	4,75	3,34	67,30	69,31	62,55
ENSG00000179059	retro_hsap_106	ZFP42	<i>YY1</i>	5,0373	8,49E-59	6,93E-58	0,90	0,29	26,11	26,15	25,21
ENSG00000260459	retro_hsap_1647	FTLP14	<i>FTL</i>	5,3451	1,67E-33	8,93E-33	0,31	0	13,89	14,32	13,58