Uniwersytet Warmińsko-Mazurski w Olsztynie
Wydział Biologii i Biotechnologii

mgr Karol G. Makowczenko

Wpływ chemeryny na transkryptom komórek lutealnych jajnika świni domowej (*Sus scrofa domestica L.*)

Praca doktorska wykonana
w Katedrze Anatomii i Fizjologii Zwierząt

Promotor: prof. dr hab. Tadeusz Kamiński
Promotor pomocniczy: dr Jan Paweł Jastrzębski

Olsztyn 2022

Praca finansowana przez Narodowe Centrum Nauki w ramach projektu badawczego: 2015/17/B/NZ9/03595

Spis treści

SPIS TREŚCI 1
STRESZCZENIE 2
SUMMARY 5
1. SPIS PUBLIKACJI SKŁADAJĄCYCH SIĘ NA ROZPRAWĘ DOKTORSKĄ 8
2. WSTĘP 9
3. CEL I ZAKRES BADAŃ 15
4. MATERIAŁY I METODY 16
4.1. Zwierzęta doświadczalne i materiał badawczy 16
4.2. Izolacja i hodowle *in vitro* komórek lutealnych jajnika 16
4.3. Izolacja RNA oraz wysokoprzepustowe sekwencjonowanie transkryptomu 18
4.4. Analizy in silico 18
4.4.1. Kontrola jakości sekwencjonowania oraz wstępne przetwarzanie odczytów 19
4.4.2. Proces mapowania odczytów oraz analiza różnic ekspresji genów 19
4.4.3. Identyfikacja oraz analiza różnic ekspresji długich niekodujących RNA 20
4.4.4. Analiza zmienności miejsc alternatywnego składania transkryptów 22
4.4.5. Analiza allelicznej specyficzności ekspresji 22
4.4.6. Analiza funkcjonalna zidentyfikowanych przypadków DEG, DEL, DAS i ASE 23
4.5. Walidacja ekspresji wybranych genów oraz długich niekodujących RNA metodą qPCR 24
5. WYNIKI 26
5.1. Wyniki wysokoprzepustowego sekwencjonowania transkryptomów 26
5.2. Wpływ chemeryny na ekspresję genów kodujących białka 26
5.3. Wpływ chemeryny na ekspresję długich niekodujących RNA 27
5.4. Modyfikacje procesu alternatywnego składania transkryptów pod wpływem chemeryny 28
5.5. Oddziaływanie chemeryny na zmienność allelicznej specyficzności ekspresji 29
5.6. Analizy funkcjonalne uzyskanych wyników *in silico* 30
5.6.1. Wpływ chemeryny na ekspresję genów kodujących funkcjonalne białka. Adnotacje funkcjonalne DEG 30
5.6.2. Wpływ chemeryny na ekspresję genów kodujących funkcjonalne białka. Adnotacje funkcjonalne DEL, DAS i ASE 31
5.7. Walidacja wybranych wyników *in silico* metodą qPCR 32
6. PODSUMOWANIE I WNIOSKI 34
7. BIBLIOGRAFIA 35

STRESZCZENIE

WPŁYW CHEMERYNY NA TRANSKRYPTOM KOMÓREK LUTEALNYCH JAJNIKA ŚWINI DOMOWEJ (*SUS SCROFA DOMESTICA* L.)

mgr Karol G. Makowczenko

Praca doktorska została wykonana w Katedrze Anatomii i Fizjologii Zwierząt Wydziału Biologii i Biotechnologii Uniwersytetu Warmińsko-Mazurskiego w Olsztynie pod kierunkiem prof. dr. hab. Tadeusza Kamińskiego (promotor) i dr. Jana Pawła Jastrzębskiego (promotor pomocniczy).

Rozprawa doktorska składa się z dwóch prac opublikowanych w 2020 i 2022 roku (1 – *Genes* 2020; 11:651. doi: 10.3390/genes11060651; 2 – *Cells* 2022; 11:715. doi: 10.3390/cells11040715), których tematyka obejmuje wpływ chemeryny na ekspresję genów i długich niekodujących RNA oraz modyfikacje procesów alternatywnego składania transkryptów, a także allelicznej specyficzności ekspresji w komórkach lutealnych jajnika świni domowej (*Sus scrofa domestica* L.) w fazie środkowolutealnej cyklu rujowego.

Szereg badań wskazuje na istnienie w organizmach ssaków systemu endokrynnego, kontrolującego jednocześnie procesy metaboliczne oraz funkcjonowanie układu rozrodczego. Zgodnie z doniesieniami z ostatnich lat, jednym z kluczowych elementów tego systemu, wpływającym także na układ immunologiczny, wydaje się być niedawno odkryty hormon – chemeryna. Ze względu na funkcje pełnione przez ten peptyd, związane z inicjowaniem lub wyciszaniem lokalnej reakcji zapalnej oraz chemotaksją wybranych grup komórek odpornościowych, zaliczany jest on do chemokin. Ponadto, ze względu na produkcję chemeryny głównie przez adipocyty oraz jej udział w regulacji procesów metabolicznych, hormon ten zaliczany jest także do grupy adipokin. Dostępne dane wskazują również na udział chemeryny w kontroli układu rozrodczego ssaków, m.in. hamujący wpływ na steroidogenezę jajnikową. Stwierdzono także obecność transkryptów oraz białek chemeryny i jej receptorów w strukturach endokrynnej osi regulacyjnej podwzgórze–przysadka–gonady (HPG) u ludzi, szczurów, świń, bydła i owiec. Chemeryna oddziałuje na komórki docelowe poprzez trzy receptory: CMKLR1, GPR1 oraz CCRL2. Dwa pierwsze, poprzez sprzężenie z białkami G, generują odpowiedź komórki, m.in. w postaci transkrypcji wybranych genów. Ostatni z receptorów nie powoduje transdukcji sygnału, lecz odpowiada

za lokalną agregację cząsteczek chemeryny i udostępnianie ich pozostałym receptorom. Fragmentaryczne informacje dotyczące aktywności chemeryny w organizmach ssaków oraz wspomniane wyżej obserwacje były przesłanką do przeprowadzenia badań opisanych w rozprawie doktorskiej.

Postawiona w rozprawie hipoteza zakładała wpływ chemeryny na profile ekspresji genów, długich niekodujących RNA oraz przebieg procesów alternatywnego składania transkryptów i allelicznej specyficzności ekspresji w komórkach lutealnych jajnika świni domowej (*S. scrofa domestica* L.) w fazie środkowolutealnej cyklu rujowego (10–12 dzień cyklu, gdy aktywność wydzielnicza komórek jest najwyższa). Do realizacji badań zastosowano metodę wysokoprzepustowego sekwencjonowania transkryptomów (RNA-Seq), która w połączeniu z zaawansowanymi analizami *in silico* pozyskanych danych umożliwiła wgląd w mechanizmy molekularne inicjowane przez badany hormon.

Badania przeprowadzono z wykorzystaniem komórek lutelanych hodowanych w warunkach *in vitro,* pobranych od dojrzałych płciowo samic świni domowej. RNA, wyizolowane z komórek hodowanych w obecności chemeryny oraz z prób kontrolnych, poddano sekwencjonowaniu RNA-Seq, uzyskując zestaw danych, które analizowano przy użyciu narzędzi bioinformatycznych. W trakcie analiz zaobserwowano istotne zmiany ekspresji 509 genów oraz 24 transkryptów zakwalifikowanych jako długie niekodujące RNA, a także 137 przypadków modyfikacji procesu alternatywnego składania transkryptów i 18 przypadków zmienności allelicznej specyficzności ekspresji. Na podstawie uzyskanych wyników przeprowadzono szczegółowe analizy funkcjonalne, których efektem było przypisanie genów (powiązanych ze wspomnianymi procesami transkryptomicznymi) do takich procesów wewnątrzkomórkowych, jak: interakcje cytokin z receptorami cytokin, apoptoza, metabolizm kwasu arachidonowego oraz ścieżki transdukcji sygnałów chemokin, TNF, MAPK i in.. Bazując na pozyskanych danych funkcjonalnych sporządzono dokładne mapy prezentujące ujednolicony wpływ chemeryny na szlaki sygnałowe oraz metaboliczne w komórkach lutealnych świni domowej. Wyniki uzyskane metodami *in silico* walidowano na bazie materiału użytego do RNA-Seq, z zastosowaniem ilościowej reakcji łańcuchowej polimerazy w czasie rzeczywistym (qPCR).

Uzyskane wyniki poszerzyły wiedzę dotyczącą wpływu chemeryny na proces transkrypcji oraz towarzyszące mechanizmy molekularne w komórkach lutealnych.

Zaobserwowane zmiany transkryptomiczne mogą prowadzić do aktywacji szlaków transdukcji sygnałów MAPK oraz JAK/STAT, indukcji zarówno mechanizmów pro-apoptycznych i pro-przeżyciowych w komórkach lutealnych. Ponadto chemeryna, poprzez modyfikację transkryptomu komórek lutealnych, może kontrolować funkcjonowanie ciałka żółtego za pośrednictwem regulacji procesów steroidogenezy i angiogenezy oraz produkcji i wydzielania czynników prozapalnych oraz prostaglandyn. Przeprowadzone badania sugerują, że chemeryna jest hormonem łączącym regulację homeostazy energetycznej i układu rozrodczego w organizmach ssaków.

Praca finansowana przez Narodowe Centrum Nauki w ramach projektu badawczego: 2015/17/B/NZ9/03595

SUMMARY

THE INFLUENCE OF CHEMERIN ON THE TRANSCRIPTOME OF OVARIAN LUTEAL CELLS OF THE DOMESTIC PIG (*SUS SCROFA DOMESTICA* L.)

Karol G. Makowczenko, M.Sc.

Doctoral thesis was carried out at the Department of Animal Anatomy and Physiology, Faculty of Biology and Biotechnology, University of Warmia and Mazury in Olsztyn under the guidance of prof. dr hab. Tadeusz Kaminski (supervisor) and Jan Pawel Jastrzebski, PhD (ancillary supervisor).

The doctoral thesis comprises two papers published in 2020 and 2022 (1 – *Genes* 2020; 11:651. doi: 10.3390/genes11060651; 2 – *Cells* 2022; 11:715. doi: 10.3390/cells11040715), focused on the influence of chemerin on genes and long non-coding RNAs expression and modifications of alternative splicing and allel-specific expression processes in ovarian luteal cells of the domestic pig (*Sus scrofa domestica* L.) during the mid-luteal phase of the estrous cycle.

Many studies indicate the existence of an endocrine system that simultaneously controls metabolic processes and the functioning of the reproductive system in mammalian organisms. According to recently published reports, one of the key elements of this system, which also interacts with the immune system, seems to be a lately discovered hormone – chemerin. Due to the functions performed by this peptide, related to the initiation or silencing of the local inflammatory response and chemotaxis of selected immune cell groups, it is categorized as a chemokine. In addition, due to the chemerin's production mainly by adipocytes and its involvement in the regulation of metabolic processes, this hormone is also included in the adipokine group. Available data indicate that chemerin is also involved in the control of the mammalian reproductive system, including an inhibitory effect on the ovarian steroidogenesis. Transcripts and proteins of chemerin and its receptors have also been found in structures of the endocrine regulatory hypothalamic-pituitary-gonadal axis (HPG) in humans, rats, pigs, cattle and sheep. Chemerin interacts with target cells through three receptors: CMKLR1, GPR1 and CCRL2. The first two, by binding to G proteins, generate a cellular response, including transcription of selected genes. The last receptor does not cause signal transduction, but it is responsible for the local aggregation of chemerin molecules and

sharing them with other receptors. Fragmentary information on the activity of chemerin in mammalian organisms and the aforementioned observations were the premis for the research described in the doctoral thesis.

The hypothesis formulated in this dissertation assumed, that chemerin influences expression profiles of genes and long non-coding RNAs, and modifies the alternative splicing and allel-specific expression processes in the ovarian luteal cells of the domestic pig (*S. scrofa domestica* L.) during the mid-luteal phase of the estrous cycle (days 10–12 of the cycle, when the secretory activity of the cells is the highest). High-throughput transcriptome sequencing (RNA-Seq) methods were used to conduct the study, which, combined with advanced *in silico* analyses of the acquired data, provided detailed insight into the molecular mechanisms initiated by the studied hormone.

The research was conducted using *in vitro* cultured luteal cells obtained from sexually mature female domestic pigs. RNA, isolated from cells treated with chemerin and from control samples, was subjected to RNA-Seq sequencing, generating a bioinformatically processed data set. During the analyses, significant changes in the expression of 509 genes and 24 transcripts classified as long non-coding RNAs were observed, as well as 137 cases of modifications in the alternative splicing process and 18 cases of allel-specific expression variation. Based on the results, detailed functional analyses were performed, resulting in the assignment of genes (affected by the observed transcriptomic phenomena) to such intracellular processes as cytokine interactions with cytokine receptors, apoptosis, arachidonic acid metabolism and signal transduction pathways of chemokines, TNF, MAPK and others. Based on the acquired functional data, detailed maps showing the unified effects of chemokine on signaling and metabolic pathways in domestic pig luteal cells were prepared. The results obtained by *in silico* methods were validated on the basis of the material used for RNA-Seq, using quantitative real-time polymerase chain reaction (qPCR).

The obtained results expanded the knowledge of the chemerin effect on the transcription process and accompanying molecular mechanisms in luteal cells. The observed transcriptomic changes may lead to activation of MAPK and JAK/STAT signal transduction pathways, and stimulation of both pro-apoptotic and pro-survival mechanisms in luteal cells. In addition, chemerin, by modifying the luteal cells' transcriptome, may control the corpus luteum functions through the regulation of steroidogenesis and angiogenesis, as well as the

production and secretion of pro-inflammatory factors and prostaglandins. The conducted studies suggest that chemerin is the hormone linking the regulation of energy homeostasis and the reproductive system in mammalian organisms.

This research was supported by National Science Centre (Project no.: 2015/17/B/NZ9/03595)

1. SPIS PUBLIKACJI SKŁADAJĄCYCH SIĘ NA ROZPRAWĘ DOKTORSKĄ

Rozprawa doktorska składa się z dwóch prac opublikowanych w 2020 i 2022 roku, których tematyka obejmuje wpływ chemeryny na profile ekspresji genów oraz na mechanizmy towarzyszące transkrypcji, tj. zmiany ekspresji długich niekodujących RNA i ich interakcje z genami o znacząco zmodyfikowanej ekspresji, a także zmienność procesów alternatywnego składania transkryptów oraz allelicznej specyficzności ekspresji związanej z frekwencją wariantów pojedynczych nukleotydów, w komórkach lutealnych jajnika świni domowej (*Sus scrofa domestica* L.) w fazie środkowolutealnej cyklu rujowego:

Publikacja 1: MAKOWCZENKO K.G., JASTRZĘBSKI J.P., SZESZKO K., SMOLIŃSKA N., PAUKSZTO Ł., DOBRZYŃ K., KIEŻUN M., RYTELEWSKA E., KAMIŃSKA B., KAMIŃSKI T. 2020 Transcription analysis of the chemerin impact on gene expression profile in the luteal cells of gilts. Genes, 11:651. doi: 10.3390/genes11060651
(IF[1] = 3,759; pięcioletni IF[2] = 3,822; pkt.[3] = 100)

Publikacja 2: MAKOWCZENKO K.G., JASTRZĘBSKI J.P., PAUKSZTO Ł., DOBRZYŃ K., KIEŻUN M., SMOLIŃSKA N., KAMIŃSKI T. 2022 Chemerin impact on alternative mRNA transcription in the porcine luteal cells. Cells, 11:715. doi: 10.3390/cells11040715
(IF[1] = 7,666; pięcioletni IF[2] = 7,677; pkt.[3] = 140)

Razem:

Sumaryczny IF: 11,425

Sumaryczny pięcioletni IF: 11,499

Liczba punktów MNiSW: 240 pkt.

[1]IF – wskaźnik Impact Factor czasopism z roku wydania wg bazy Web of Science

[2]pięcioletni IF – pięcioletni wskaźnik Impact Factor czasopism wg bazy Web of Science

[3]Liczba punktów wg Komunikatu Ministra Edukacji i Nauki w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych z dnia 21 grudnia 2021 roku.

2. WSTĘP

Chemerynę, jako substancję ulegającą zwiększonej produkcji pod wpływem syntetycznej pochodnej kwasu retinowego – tazarotenu, odkryto ćwierć wieku temu w trakcie badań nad patogenezą łuszczycy (NAGPAL i in. 1997). Hormon ten powiązano z ekspresją *TIG2* (ang. *tazarotene-induced gene 2*), z czasem przemianowanego na *RARRES2* (ang. *retinoic acid receptor responder protein 2*). Sekwencję białkową produktu tego genu zidentyfikowały dwa niezależne zespoły badawcze w 2003 roku (MEDER i in. 2003; WITTAMER i in. 2003).

Chemeryna, nazwana tak przez Valerie Wittamer, syntetyzowana jest przez komórki ssaków w formie nieaktywnych cząsteczek pre-prochemeryny o długości 163 aminokwasów i masie molekularnej 18,62 kDa (KENNEDY i DAVENPORT 2018). Dwadzieścia aminokwasów od strony końca aminowego tworzy peptyd sygnalny. Ten hydrofobowy fragment wycinany jest w trakcie wydzielania makromolekuł do krwioobiegu, które w dalszym ciągu tworzą nieaktywną formę propeptydu (16,53 kDa). Skomplikowany i często wieloetapowy mechanizm aktywacji białka wymaga poddania końca karboksylowego reakcjom hydrolizy przez endogenne proteinazy serynowe i/lub cysteinowe (ERNST i SINAL 2010). W zależności od lokalnie przeważającej klasy enzymów, przeprowadzających selektywną proteolizę różnych wiązań między aminokwasami C-końca peptydu (wynikającej z mikrośrodowiska tkanki), generowane mogą być różne izoformy chemeryny o zróżnicowanej aktywności biologicznej (DU i LEUNG 2009). Izoforma powstała w wyniku oddzielenia 6 ostatnich aminokwasów – chemeryna-157 (15,88 kDa), charakteryzuje się największą obserwowaną bioaktywnością (około 100-krotnie wyższą od prochemeryny) (BONDUE i in. 2011; DU i in. 2009). Dalsze wieloetapowe procesy hydrolizy końca karboksylowego, powodujące skracanie łańcucha białkowego, prowadzą do inaktywacji hormonu – izoformy chemeryny-154 i -152 uznaje się za relatywnie nieaktywne (DU i LEUNG, 2009; ROURKE i in. 2013; ZABEL i in. 2005). Mimo braku dostępnych danych krystalograficznych dotyczących struktury przestrzennej cząsteczek omawianego hormonu, przewidywania *in silico* przeprowadzone metodą rozpoznawania zwojów (ang. *fold recognition*) wskazały na bliskie podobieństwo cząsteczek chemeryny do rodzin katelicydyn

(białek antybakteryjnych), kininogenów (kofaktorów procesów krzepnięcia i kaskady zapalnej) oraz cystatyn (inhibitrów peptydaz cystynowych) (LEI i in. 2019; ZABEL i in. 2014).

Dotychczas w organizmach ssaków zidentyfikowano trzy transbłonowe receptory metabotropowe zdolne do wiązania cząsteczek chemeryny: CMKLR1 (ang. *chemerin chemokine-like receptor 1*), GPR1 (ang. *G protein-coupled receptor 1*) oraz CCRL2 (ang. *C–C motif chemokine receptor like 2*) (KENNEDY i DAVENPORT 2018). Dwa pierwsze z nich wiążą N-koniec cząsteczek chemeryny, a poprzez tworzenie wiązań z białkami G powodują transdukcję sygnału za pośrednictwem szlaków PI3K/Akt i MAPK-ERK1/2 (CMKLR1) lub AMPK (CMKLR1 i GPR1) (DE HENAU i in. 2016; KENNEDY i DAVENPORT 2018; MATTERN i in. 2014). Ostatni z wymienionych receptorów – CCRL2 nazywany jest „receptorem atypowym", ze względu na wiązanie C-końca cząsteczek chemeryny ze znacznie niższym powinowactwem względem pozostałych białek receptorowych oraz brak indukcji wewnątrzkomórkowych kaskad sygnałowych. Zidentyfikowano natomiast udział cząsteczek CCRL2 w lokalnej agregacji chemeryny oraz ekspozycji tego białka dwóm pozostałym receptorom (MONNIER i in. 2012; ZABEL i in. 2008).

W 2007 roku zaklasyfikowano chemerynę do grupy adipokin ze względu na wytwarzanie przez tkankę tłuszczową oraz udowodnione funkcje metaboliczne, tj. regulację adipogenezy oraz metabolizmu białej tkanki tłuszczowej i zwiększanie wychwytu glukozy (GÓRALSKI i in. 2007; MATTERN i in. 2014). Ponadto, z uwagi na przejawianie aktywności związanych z procesami odpornościowymi, takimi jak tkankowo-specyficzne właściwości pro- lub przeciwzapalne, właściwości bakteriostatyczne oraz wywoływanie chemotaksji makrofagów, komórek dendrytycznych i komórek NK (ang. *natural killer*), chemerynę przypisano również do grupy cytokin chemotaktycznych (in. chemokin) (MATTERN i in. 2014; PAROLINI i in. 2007; WITTAMER i in. 2003). Szczególnie ciekawą i wartą podkreślenia obserwacją jest wzrost koncentracji chemeryny oraz jej aktywacji proteolitycznej w regionach tkanki objętych reakcją zapalną, w sposób niezwiązany ze stężeniem tej cytokiny w surowicy (PAROLINI i in. 2007; ROURKE i in. 2013; WITTAMER i in. 2003; ZABEL i in. 2005).

W 2012 roku zauważono bezpośredni wpływ chemeryny na funkcjonowanie układów rozrodczych ssaków, szczególnie ludzi oraz czołowych laboratoryjnych organizmów modelowych – myszy i szczurów. Początkowo skupiono się na wpływie

omawianej adipokiny na steroidogenezę w komórkach ziarnistych budujących ściany pęcherzyków jajnikowych (REVERCHON i in. 2012; WANG i in. 2012, 2013b), a także na roli chemeryny w ciąży (BARKER i in. 2012; GARCES i in. 2012). Z czasem badania rozszerzono na pozostałe struktury osi regulacyjnej podwzgórze–przysadka–gonady (ang. *hypothalamus–pituitary–gonadal axis*, HPG) oraz układy rozrodcze innych, istotnych dla gospodarki, zwierząt, tj. bydła, trzody chlewnej, kóz i owiec, w różnych fazach cyklu rujowego oraz na wczesnych etapach ciąży (KUROWSKA i in. 2018, 2021; REVERCHON i in. 2014a, 2014b). U dojrzałych płciowo loszek obecność elementów systemu chemerynowego (chemeryny i jej receptorów) wykryto w wybranych jądrach podwzgórza (SMOLIŃSKA i in. 2019), przednim i tylnym płacie przysadki (KISIELEWSKA i in. 2020a, 2020b), strukturach budujących jajnik, takich jak pęcherzyki jajnikowe oraz ciałka żółte (RYTELEWSKA i in. 2020) oraz w macicy (GUDELSKA i in. 2020). Niestety, mimo szeroko zakrojonych badań tłumaczących wpływ chemeryny na ekspresję pojedynczych genów lub produkcję wąskiej gamy metabolitów, w dalszym ciągu brakuje eksperymentów prezentujących pełny obraz efektów oddziaływania chemeryny na transkryptom, proteom i metabolom, umożliwiających szczegółowe wnioskowanie dotyczące uruchamianych i/lub wyłączanych mechanizmów molekularnych zachodzących w grupach komórek, tkankach oraz całych narządach ssaków.

Uzupełniając wspomniane luki w wiedzy na temat wpływu chemeryny na funkcjonowanie układu rozrodczego ssaków, w przeprowadzonych badaniach składających się na niniejszą rozprawę doktorską posłużono się metodą wysokoprzepustowego sekwencjonowania transkryptomów następnej generacji (RNA-Seq). Metoda ta, dostarczając informacji jakościowych – dotyczących budowy poszczególnych cząsteczek RNA oraz ilościowych – związanych z liczbami wykrytych kopii konkretnych transkryptów, umożliwia dokładne zobrazowanie profili tanskryptomicznych badanych komórek/tkanek w postaci kolejnych bibliotek poddanych sekwencjonowaniu. Standardowym i najczęściej przeprowadzanym pomiarem porównawczym dla danych RNA-Seq jest analiza różnic ekspresji genów lub transkryptów. Polega ona na porównaniu uśrednionych wartości ekspresji (z powtórzeń biologicznych), oszacowanych dla genów lub pojedynczych transkryptów, między badanymi grupami. Na podstawie danych dotyczących genów o zróżnicowanej ekspresji (ang. *differentially expressed genes*, DEG), śledzi się zmiany zachodzące w transkryptomie. Następnie na podstawie zmian w kodujących sekwencjach

DNA (ang. *coding DNA sequence,* CDS) przewiduje się zmiany w proteomie badanych komórek lub tkanek, a w dalszej kolejności wysnuwa się prognozy dotyczące modyfikacji funkcjonowania wyższych struktur organizacyjnych organizmu (narządów, układów wielonarządowych). Należy odnotować, że na podstawie prac SCHWANHÄUSSERA i in. (2011) oraz DE SOUSA ABREU i in. (2009), obecnie często powtarzana jest informacja, iż różnice w poziomach ekspresji genów wpływają tylko w ok. 40 procentach na koncentrację białek wytwarzanych w komórkach ssaków (współczynnik determinacji R^2 między zawartością matrycowego RNA – mRNA i białka wynosi 0,41). Jednakże, uwzględniając wprowadzone z czasem korekty przytoczonych badań, zmiany koncentracji mRNA wyjaśniają w 68–84% różnice w zawartości białek w komórkach, dając znacznie wyższy wskaźnik predykcyjny badań transkryptomowych (BATTLE i in. 2015; JOVANOVIC i in. 2015; LI i in. 2014b). Warto zaznaczyć, że oprócz analizy różnic ekspresji genów, równie istotne jest uwzględnienie zjawisk towarzyszących modyfikacjom transkrypcyjnym i posttranskrypcyjnym. Takie podejście skutkuje bardziej kompletnym obrazem złożonych procesów zachodzących w komórkach, a w szerszej perspektywie – pełniejszym opisem molekularnych procesów fizjologicznych zachodzących w tkankach i narządach.

Długie niekodujące RNA (lncRNA) stanowią najliczniejszą grupę transkryptów w komórkach ssaków (KORNIENKO i in. 2013). Definiuje się je jako cząsteczki RNA pozbawione potencjału do kodowania białek, dłuższe niż 200 nukleotydów (nt) i posiadające strukturę egzonowo-intronową. Zgodnie z coraz większą liczbą doniesień, lncRNA są istotnymi regulatorami ekspresji genów, tzn. poprzez kontrolę procesu transkrypcji wpływają na stabilność mRNA oraz przebieg translacji i obróbki posttranslacyjnej białek (YAO i in. 2019). Uogólniając, wpływ cząsteczek lncRNA na kontrolę transkrypcji może mieć charakter *cis*-interakcji, gdy regulacja zachodzi w obrębie jednego chromosomu w odległości umożliwiającej wpływ na ekspresję, lub *trans*-interakcji, gdy gen docelowy znajduje się na innym chromosomie lub w znacznym oddaleniu na tym samym chromosomie (ALESSIO i in. 2020). Długie niekodujące RNA mogą wpływać na ekspresję genów docelowych zarówno bezpośrednio, poprzez wiązanie sekwencji położonej na komplementarnej nici DNA, jak i w sposób pośredni, m.in. poprzez oddziaływanie z drugim chromosomem z pary lub wpływając na ekspresję polimeraz RNA (KORNIENKO i in. 2013). Ponadto lncRNA mogą zmieniać aktywność białek poprzez wiązanie ich specyficznych domen (SUN i in. 2013) oraz

pełnią funkcje pomocnicze w organizacji kompleksów białkowych oraz nukleoproteinowych (RAO 2017). Długie niekodujące RNA często wykazują ekspresję specyficzną dla konkretnej tkanki lub komórki (FRASER i in. 2020). Zauważalny jest wzrost liczby doniesień dotyczących udziału lncRNA w regulacji funkcji rozrodczych ssaków (LIU i in. 2018; TAYLOR i in. 2015), w tym ich zaangażowanie w prawidłowe funkcjonowanie ciałka żółtego (NAKAGAWA i in. 2014). W opisanych poniżej badaniach wykorzystano analizy identyfikacji oraz selekcji lncRNA o zróżnicowanej ekspresji (ang. *differentially expressed lncRNA*, DEL) pod wpływem chemeryny. Wykonano także analizę *cis*-interakcji z DEG, na podstawie ko-lokalizacji w genomie oraz *trans*-interakcji, uwzględniając korelację ekspresji i podobieństwo struktur przestrzennych między cząsteczkami lncRNA a mRNA/białkami.

Alternatywne składanie transkryptów (in. alternatywny splicing) jest mechanizmem polegającym na wytwarzaniu wielu wariantów transkryptów na podstawie pojedynczego wzorca w postaci genu (EARLY i in. 1980; SALDI i in. 2016). Zjawisko to, wykorzystując proces wycinania intronów w trakcie dojrzewania prekursorowego mRNA, powoduje modyfikację struktury egzonowo-intronowej dojrzałych transkryptów, m.in. poprzez pomijanie wybranych egzonów, zmianę długości egzonów lub zatrzymanie sekwencji intronowych w mRNA – są to miejsca alternatywnego splicingu. Alternatywne składanie matryc, bezpośrednio kodujących białka, przyczynia się do zwiększenia różnorodności proteomu, bez zwiększania rozmiaru genomu organizmów eukariotycznych (NOBLE i in. 2020). Warto zaznaczyć, że wynikiem opisanego procesu może być polimorfizm białek kodowanych przez jeden gen, które w konsekwencji posiadania zróżnicowanych domen, mogą pełnić różne funkcje biologiczne (THATCHER i in. 2014). Warianty mRNA powstałe w wyniku alternatywnego splicingu mogą zawierać kodony STOP przedwcześnie kończące transkrypcję białka. Produkty transkryptów zawierających taki nonsensowny kodon mogą być dysfunkcyjne, skutkując skierowaniem nieprawidłowych makromolekuł do degradacji. Mechanizm ten służy do kontrolowania ilości produktów białkowych danego genu bez zmiany jego ekspresji w komórkach ssaków (SCHWEINGRUBER i in. 2013). Niektóre powstałe w ten sposób białka o skróconym łańcuchu aminokwasowym mogą bezpośrednio blokować aktywność białek o pełnej długości, dzięki zachowaniu domen zdolnych do oddziaływania z kompleksami białkowymi lub substratami (STAUDT i WENKEL 2011). Co szczególnie istotne, ekspresja wariantów transkryptów ulegających alternatywnemu składaniu jest często

specyficzna dla konkretnych tkanek (WANG i in. 2008). W ciałkach żółtych zaobserwowano przypadki alternatywnego splicingu w obrębie genów, których produkty są zaangażowane w luteolizę, matczyne rozpoznawanie ciąży oraz wzrost i regresję naczyń krwionośnych (MADHRA i in. 2004; RIBEIRO i in. 2007). Poniżej opisane badania komórek lutealnych świni domowej objęły identyfikację przypadków alternatywnego składnia transkryptów, wykazujących różnice frekwencji (ang. *differential alternative splicing events*, DAS) pod wpływem chemeryny, z zamiarem wykrycia zmian w obrębie transkryptomu niewidocznych podczas eksperymentów bazujących na modyfikacji ekspresji genów/transkryptów.

Zdecydowana większość genów wykazuje jednakową ekspresję obu alleli, jednak ekspresja niektórych genów może być modyfikowana w sposób zależny od macierzystych alleli. Pod wpływem czynników regulacyjnych, takich jak metylacja DNA, modyfikacje białek histonowych przyłączonych do DNA oraz aktywności niekodujących cząsteczek RNA, produkcja białka kodowanego przez jeden z alleli tego samego genu może być promowana u heterozygot (GAUR i in. 2013). Zjawisko to jest istotnym genetycznym procesem prowadzącym do zwiększonego zróżnicowania fenotypowego, wywołanego m.in. zmiennymi warunkami środowiskowymi (KHANSEFID i in. 2018). Przykładem może być wzrost angiogenezy w organizmach ludzi spowodowany występowaniem specyficznych polimorfizmów pojedynczych nukleotydów, bezpośrednio związanych ze stężeniem chemeryny w osoczu (BOZAOGLU i in. 2010). Obserwacja ta doprowadziła do wniosku, że chemeryna może być zdolna do promowania ekspresji jednego z dwóch alleli dostępnych w DNA komórek. W poniżej przedstawionym eksperymencie zastosowano procedurę umożliwiającą wykrycie przypadków allelicznej specyficzności ekspresji (ang. *allele-specific expression*, ASE) pośród wariantów pojedynczych nukleotydów (ang. *single nucleotide variants*, SNV) wykazujących dysproporcje frekwencji w obrębie DEG.

3. CEL I ZAKRES BADAŃ

Celem niniejszej pracy była szczegółowa analiza wpływu chemeryny na profil strukturalny oraz ekspresyjny transkryptomu komórek lutealnych świni domowej (*Sus scrofa domestica* L.) w trakcie fazy środkowolutealnej cyklu rujowego. Przesłanką do przeprowadzenia poniższych badań był jak dotąd niepoznany wpływ wspomnianego hormonu na ekspresję genów oraz towarzyszące mechanizmy molekularne w tych komórkach. Dotychczas przeprowadzone eksperymenty dostarczyły fragmentarycznych informacji dotyczących wpływu chemeryny na ekspresję jedynie pojedynczych genów lub wydzielanie wybranych substancji przez komórki lutealne świni. Do tej pory brakowało badań szczegółowo analizujących działanie chemeryny na procesy towarzyszące ekspresji i dojrzewaniu makromolekuł budujących transkryptom komórek lutealnych. Hipotezy badawcze pracy zakładały, że chemeryna w fizjologicznych stężeniach (1) wpływa na globalną ekspresję genów, (2) jest odpowiedzialna za modyfikacje ekspresji długich niekodujących RNA oraz ich interakcje z produktami genów kodujących białka, (3) reguluje proces allelicznej specyficzności ekspresji oraz (4) moduluje procesy posttranskrypcyjne, takie jak alternatywne składanie transkryptów. Postawione hipotezy zweryfikowano stosując metody wysokoprzepustowego sekwencjonowania transkryptomów hodowanych *in vitro* komórek lutealnych świni domowej, będących w środkowolutealnej fazie cyklu rujowego (dzień 10–12). Wyniki analiz *in silico* zweryfikowano poprzez pomiar przyrostu ilości produktu reakcji łańcuchowej polimerazy w czasie rzeczywistym, z wykorzystaniem RNA użytego do sekwencjonowania.

4. MATERIAŁY I METODY

4.1. Zwierzęta doświadczalne i materiał badawczy

Eksperyment przeprowadzono z wykorzystaniem materiału biologicznego pozyskanego z pięciu dojrzałych samic świni domowej (*S. scrofa domestica* L.), mieszańców towarowych. Zwierzęta w wieku 7–8 miesięcy i o wadze 120–130 kg pochodziły z prywatnej farmy hodowlanej. W doświadczeniu wykorzystano samice będące w okresie najwyższej aktywności sekrecyjnej ciałka żółtego w trakcie cyklu rujowego – fazie środkowolutealnej (10–12 dzień cyklu). Dzień, w którym wystąpiły pierwsze objawy rui, oznaczono jako „dzień 0" cyklu rujowego. Dodatkowo sprawdzono morfologię jajników w celu potwierdzenia fazy cyklu (AKINS i MORRISSETTE 1968).

Materiał badawczy został zebrany od zwierząt przeznaczonych do uboju gospodarczego i przetwórstwa mięsa. Jajniki pobrano natychmiast po uboju zwierząt, a następnie umieszczono w sterylnym i schłodzonym do 4 ºC roztworze PBS (buforowanej fosforanem soli fizjologicznej), uzupełnionym antybiotykami: 100 IU/ml penicyliny (Sigma-Aldrich, St. Louis, Missouri, USA) i 100 µg/ml streptomycyny (Sigma-Aldrich, St. Louis, Missouri, USA). Materiał badawczy niezwłocznie przetransportowano do laboratorium w celu izolacji komórek oraz założenia hodowli *in vitro*.

Pobór materiału zwierzęcego oraz kolejne eksperymenty badawcze przeprowadzono zgodnie z wytycznymi zaktualizowanej ustawy z dnia 15 stycznia 2015 roku o ochronie zwierząt wykorzystywanych do celów naukowych lub edukacyjnych (Dz.U. z 2021 r., poz. 1331 oraz z 2022 r., poz. 1117), która wdraża dyrektywę Parlamentu Europejskiego i Rady 2010/63/UE z dnia 22 września 2010 r. w sprawie ochrony zwierząt wykorzystywanych do celów naukowych (Dz. Urz. UE L 276 z 20.10.2010, str. 33).

4.2. Izolacja i hodowle *in vitro* komórek lutealnych jajnika

Komórki lutealne wyizolowano z pobranych narządów z zastosowaniem procedury zaproponowanej przez KAMIŃSKIEGO i in. (1999). Wypreparowane z jajników ciałka żółte rozdrobniono na fragmenty wielkości 1–2 mm i przeniesiono do płynnej pożywki hodowlanej F-12 (Sigma-Aldrich, St. Louis, Missouri, USA), zawierającej 1 procentowy

dodatek albuminy surowicy bydlęcej (BSA; MP Biomedicals, Irvine, Kalifornia, USA), 50 μg/ml gentamycyny (Sigma-Aldrich, St. Louis, Missouri, USA) oraz 240 U/ml nystatyny (Sigma-Aldrich, St. Louis, Missouri, USA). Fragmenty ciałka żółtego poddano 4–6-krotnemu enzymatycznemu trawieniu w 0,125 procentowym roztworze trypsyny (Biomed Lublin, Polska), w temperaturze 38 ºC, do całkowitego strawienia tkanki. Zawiesinę wyizolowanych komórek przenoszono do sterylnych probówek, wirowano przez 10 minut przy 300 × g w temperaturze 21 ºC oraz trzykrotnie przemywano medium F-12. Wyizolowane komórki lutealne odfiltrowano od pozostałości trypsynizowanej tkanki z użyciem filtra nylonowego o średnicy porów 75 μm, a następnie ponownie zawieszono w świeżej pożywce F-12. Liczbę uzyskanych komórek oszacowano z pomocą komory Bürkera, a ich żywotność, na poziomie ≈90%, określono z wykorzystaniem 0,4 procentowego roztworu błękitu trypanu (Sigma-Aldrich, St. Louis, MO, USA). Komórki lutealne zawieszono w ilości 2×10^6/2 ml pożywki F-12 wzbogaconej o 20% płodowej surowicy bydlęcej (FCS; Sigma-Aldrich, St. Louis, MO, USA), 1% BSA oraz antybiotyki i następnie przez 48 godzin preinkubowano w 6-studzienkowych płytkach hodowlanych w inkubatorze, w atmosferze 95% powietrza i 5% CO_2 oraz temperaturze 37 °C. Po zakończeniu tego etapu podłoże zawierające FCS zostało usunięte, a komórki lutealne przepłukano medium F-12 bez surowicy. Komórki inkubowano w tych samych warunkach przez kolejne 24 godziny w pożywce F-12 z 1% BSA, antybiotykami oraz z dodatkiem rekombinowanej ludzkiej chemeryny (RayBiotech, Peachtree Corners, Georgia, USA) w stężeniu 200 ng/ml (n = 5). Grupę kontrolną stanowiła hodowla bez dodatku chemeryny (n = 5).

W trakcie badań zastosowano ludzką rekombinowaną chemerynę ze względu na niedostępność na rynku homologa pochodzącego od świni domowej. Zgodnie z badaniami DU i LEUNG (2009) cząsteczki chemeryny ludzi i świń wykazują identyczność sekwencji rzędu 84%. Ponadto LUANGSAY i in. (2009) wykazali, iż fragment C-końca dojrzałej chemeryny (YFPGQFAFS) jest wysoce konserwatywny u wszystkich gatunków ssaków i ma istotne znaczenie dla aktywności biologicznej tego białka. Zgodnie z doniesieniem RYTELEWSKIEJ i in. (2021) poziom identyczności wspomnianego fragmentu chemeryny wynosi 89%, a interakcja cząsteczek chemeryny pochodzących od różnych gatunków ssaków z jej receptorami pozostaje zachowana. Stężenie chemeryny w hodowlach *in vitro* dobrano

na podstawie jej koncentracji w osoczu krwi kobiet (BOZAOGLU i in. 2010) oraz loszek (SMOLIŃSKA i in. 2019).

4.3. Izolacja RNA oraz wysokoprzepustowe sekwencjonowanie transkryptomu

Całkowite RNA wyizolowano z hodowanych *in vitro* komórek lutealnych, poddanych działaniu chemeryny oraz tworzących grupę kontrolną, z użyciem zestawu RNeasy Mini Kit (Qiagen, Germantown, Maryland, USA), zgodnie z procedurą producenta. Pozostałości DNA usunięto stosując DNazę I zawartą w zestawie RNase-free DNase Set (Qiagen, Germantown, Maryland, USA). Metodą spektrofotometryczną zmierzono koncentrację (długość fali 260 nm, A_{260}) i stopień czystości (A_{260}/A_{280}) wyizolowanego RNA przy użyciu czytnika Infinite M200 Pro (Tecan, Männedorf, Szwajcaria). Integralność RNA zbadano za pomocą aparatu Bioanalyzer 2100 (Agilent Technology, St. Clara, Kalifornia, USA). Współczynnik integralności RNA (RIN) w przedziale 8–10 kwalifikował wyizolowane RNA do dalszych etapów eksperymentu, czyli sekwencjonowania powiązanego z analizami bioinformatycznymi oraz walidacji uzyskanych *in silico* wyników za pomocą pomiaru przyrostu ilości produktu reakcji łańcuchowej polimerazy w czasie rzeczywistym (qPCR).

Po wcześniejszym oczyszczeniu i fragmentacji cząsteczek, osobno dla każdej próby RNA, zsyntetyzowano biblioteki komplementarnego DNA (cDNA) z użyciem zestawu TruSeq RNA Sample Prep Kit (Illumina, San Diego, Kalifornia, USA). Utworzone biblioteki cDNA przechowywano w temperaturze –80 ºC. Na podstawie zamplifikowanych matryc cDNA przeprowadzono proces sekwencjonowania RNA-Seq z zastosowaniem platformy NovaSeq 6000 (Illumina, San Diego, Kalifornia, USA). Zakładana minimalna głębokość sekwencjonowania wyniosła 100 milionów odczytów na próbę. W efekcie uzyskano fragmenty sekwencji transkryptów – odczyty o sparowanych końcach (ang. *paired-end reads*) i długości 2 × 150 nt.

4.4. Analizy *in silico*

Analizy bioinformatyczne wymagające większej mocy obliczeniowej wykonywano na serwerze, posiadającym 40 rdzeni obliczeniowych, 120 GB pamięci RAM oraz macierz dyskową o pojemności 70 TB. Serwer został udostępniony przez Regionalne Centrum

Informatyczne (należące do Uniwersytetu Warmińsko-Mazurskiego w Olsztynie). Graficzne zestawienie kolejnych etapów przeprowadzonych analiz bioinformatycznych przedstawiono na Rycinach 1 zawartych w obu publikacjach.

Pliki zawierające surowe odczyty uzyskane w wyniku sekwencjonowania zdeponowano w Europejskim Archiwum Nukleotydów (ang. *European Nucleotide Archive*, ENA) pod numerem akcesyjnym **PRJEB35892** (https://www.ebi.ac.uk/ena/browser/view/PRJEB35892; dostęp: 14 czerwca 2022 r.). Szczegółowe procedury eksperymentu oraz dane dotyczące różnic ekspresji genów i transkryptów, uzyskane dla wszystkich prób uwzględnionych w badaniu, udostępniono w bazie danych ArrayExpress pod numerem akcesyjnym **E-MTAB-11434** (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-11434/; dostęp: 14 czerwca 2022 r.).

4.4.1. Kontrola jakości sekwencjonowania oraz wstępne przetwarzanie odczytów

Jakość surowych odczytów, wygenerowanych w wyniku sekwencjonowania, sprawdzono z użyciem programu FastQC wer. 0.11.8 (ANDREWS 2010). Sekwencje adaptorowe oraz regiony surowych odczytów o niskiej jakości (współczynnik jakości sekwencji QPhred na końcach 5' i/lub 3' < 20 oraz średnia wartość QPhred < 30) wycięto za pomocą programu Trimmomatic wer. 0.38 (BOLGER i in. 2014). Odczyty przycięto do równych segmentów długości 90 nt. Przetworzone odczyty ponownie skontrolowano pod względem jakości i obecności sekwencji adaptorowych za pomocą FastQC.

4.4.2. Proces mapowania odczytów oraz analiza różnic ekspresji genów

Przetworzone odczyty przyrównano (zmapowano) do genomu referencyjnego świni domowej, w oparciu o program STAR wer. 2.6.1d (publikacja 1) i 2.7.3a (publikacja 2) (DOBIN i in. 2013). Genom referencyjny, w dwóch wersjach, pozyskano z ogólnodostępnej bazy danych Ensembl (YATES i in. 2020) – *S. scrofa* wer. 11.1.91, do analizy różnic ekspresji genów (publikacja 1), i *S. scrofa* wer. 11.1.99, na potrzeby analizy mechanizmów towarzyszących transkrypcji RNA (publikacja 2). Zmiana wersji genomu referencyjnego została uwarunkowana uzupełnieniem baz danych o informacje dotyczące sekwencji nowopoznanych cząsteczek lncRNA. Liczby odczytów przyrównanych, w trakcie procesu mapowania, do sekwencji CDS, intronów, regionów mRNA niepodlegających translacji

(ang. *untranslated regions*, UTR) oraz regionów międzygenowych oszacowano z użyciem programu CollectRnaSeqMetrics, będącego elementem pakietu Picard wer. 2.21.1 (BROAD INSTITUTE 2021). Wykorzystując oprogramowanie StringTie wer. 1.3.5 (PERTEA i in. 2015), wyznaczono sekwencje transkryptów oraz profile ekspresji poznanych genów i uprzednio nieopisanych regionów aktywnych transkrypcyjnie (ang. *transcriptionally active regions*, TAR). W celu oszacowania homogenności i powtarzalności danych wewnątrz grup badanej i kontrolnej przeprowadzono analizę składowych głównych (ang. *principal component analysis*, PCA) i analizę skupień (ang. *cluster analysis*), bazującą na pomiarze odległości euklidesowych, z użyciem biblioteki ggplot2 (WICKHAM 2021) w środowisku R wer. 3.5.2 (R CORE TEAM 2021). Między grupą poddaną działaniu chemeryny oraz kontrolną przeprowadzono analizę różnic ekspresji genów, w środowisku R, z użyciem pakietu Ballgown wer. 2.18.0 (FRAZEE i in. 2015). W ocenie zmian ekspresji genów zastosowano test dwumianowy i przyjęto graniczne wartości istotności $q < 0{,}05$ oraz $|\log_2(\text{krotności zmian})| \geq 0{,}5$. Geny spełniające przytoczone kryteria zostały zakwalifikowane jako DEG. Graficzną wizualizację znormalizowanych poziomów ekspresji genów we wszystkich próbach wygenerowano z użyciem pakietu gplots (WARNES i in. 2021) w środowisku R.

Analiza różnic ekspresji genów została powtórzona dla puli danych uzyskanych po ponowieniu procesu mapowania odczytów, względem nowszej wersji genomu referencyjnego (*S. scrofa* wer. 11.1.99), z zastosowaniem tych samych programów i identycznych parametrów konfiguracji. Uzyskana pula genów wykazujących statystycznie różnice ekspresji była identyczna względem pierwotnej puli z pierwszego mapowania (*S. scrofa* wer. 11.1.91). Powtórzoną analizę różnic ekspresji genów przeprowadzono w zaktualizowanej wersji środowiska statystycznego R wer. 4.0.3.

4.4.3. Identyfikacja oraz analiza różnic ekspresji długich niekodujących RNA

Identyfikację lncRNA w komórkach lutealnych świni domowej przeprowadzono przy użyciu wieloetapowej procedury. Na początku ze zbioru danych usunięto transkrypty o niskim poziomie ekspresji (wartość FPKM – ang. *fragments per kilobase milion* < 1) oraz kodujące białka. Następnie z puli transkryptów wykluczono sekwencje jednoegzonowe i o długości < 200 nt. Potencjał kodowania pozostałych transkryptów oszacowano za pomocą

czterech programów: CPAT (ang. *coding potential assessment tool*) wer. 1.2.4 (WANG i in. 2013a), CPC2 (ang. *coding potential calculator 2*) wer. 2.0.1 (KANG i in. 2017), Feelnc (ang. *flexible extraction of lncRNAs*) (WUCHER i in. 2017), PLEK (ang. *predictor of lncRNAs and mRNAs based on an improved k-mer scheme*) wer. 0.2 (LI i in. 2014a). Równocześnie sekwencje transkryptów dopasowano do bazy danych rodzin białkowych Pfam (ang. *protein families*) wer. 33.1 (EL-GEBALI i in. 2019) przy użyciu programu HMMER wer. 3.3.2 (EDDY 2011), opartego na algorytmach „ukrytych modeli Markowa" (ang. *hidden Markov model*). Na dalszych etapach analizy uwzględniono transkrypty pozbawione potencjału kodowania, zgodnie z wynikami trzech z wymienionych programów (wartości progowe potencjału kodowania CPAT < 0,364, CPC2 < 0, Feelnc < 0,558; oznaczenie PLEK = "noncoding"), oraz o niskim stopniu podobieństwa do rekordów Pfam (*e*-value > 10^{-3}). Różnice ekspresji zidentyfikowanych lncRNA pomiędzy próbami badanymi i kontrolnymi obliczono za pomocą programu Ballgown, z parametrami odcięcia wartości $q < 0{,}05$ i $|\log_2(\text{krotności zmian})| \geq 0{,}5$. W dalszej kolejności sekwencje uzyskanych DEL przyrównano do rekordów modeli małych RNA, w bazie danych Rfam (ang. *RNA families*) wer. 14.4 (KALVARI i in. 2021), z wykorzystaniem programu Infernal cmscan wer. 1.1.3 (NAWROCKI i EDDY 2013).

Kolejne etapy analizy przeprowadzono w środowisku R wer. 4.0.3 (R CORE TEAM 2021) z wykorzystaniem zintegrowanego zestawu narzędzi do wysokoprzepustowej analizy danych genomowych Bioconductor wer. 3.12 (RAMOS i in. 2017). Celem dalszych postępowań było odkrycie i scharakteryzowanie zależności między uzyskanymi DEL oraz cząsteczkami mRNA kodowanymi przez zidentyfikowane DEG. Bliskie sąsiedztwo par DEL–DEG na tym samym chromosomie opisano jako *cis*-interakcje i podzielano na dwie podkategorie: pokrywające się (ang. *overlapping*) oraz odległe (ang. *distant*), w przypadkach gdy pary genów znajdowały się w odległości do 10.000 par zasad (pz). Pary DEL–DEG zlokalizowane na różnych chromosomach, lecz wykazujące podobieństwo profili transkrypcyjnych, w oparciu o współczynnik korelacji Pearsona scharakteryzowano jako *trans*-interakcje (wartość $p < 0{,}05$ i $|r| > 0{,}9$). Dodatkowo, w oparciu o obliczenia znormalizowanej energii swobodnej (ndG), oszacowano siłę wiązania między przewidywanymi parami cząsteczek lncRNA–mRNA przy użyciu programu LncTar wer. 1.0 (LI i in. 2015). Ponadto, wykorzystując program lncPro wer. 1.0 (LU i in. 2013), obliczono

prawdopodobieństwo tworzenia wiązań wodorowych oraz oddziaływań van der Waalsa pomiędzy strukturami drugorzędowymi wykrytych lncRNA i białek kodowanych przez zidentyfikowane DEG. W analizach funkcjonalnych uwzględniono wyłącznie *trans*-interakcje, dla których LncTar wskazał ndG < –0,1 lub lncPro wykazał prawdopodobieństwo > 0,9. Zidentyfikowane relacje pomiędzy parami DEL–DEG zwizualizowano przy użyciu oprogramowania Cytoscape v. 3.8.3 (SHANNON i in. 2003).

4.4.4. Analiza zmienności miejsc alternatywnego składania transkryptów

Identyfikację miejsc składania RNA, wykazujących zmienność pod wpływem chemeryny, przeprowadzono w oparciu o analizę tzw. wykorzystania egzonów (ang. *exon usage*), stosując oprogramowanie SUPPA2 (ang. *super-fast pipeline for alternative splicing analysis 2*) wer. 2.3 (TRINCADO i in. 2018). Odczyty, uprzednio przycięte do równej długości (90 nt) i zmapowane względem genomu referencyjnego (*S. scrofa* wer. 11.1.99) przy pomocy programu STAR, przyrównano względem referencyjnego transkryptomu *S. scrofa* przy użyciu programu Salmon wer. 1.1.0 (PATRO i in. 2017). Analizę przypadków DAS w próbach poddanych działaniu chemeryny i kontrolnych przeprowadzono poprzez badanie ekspresji miejsc splicingowych w regionach łączenia egzon–egzon oraz w egzonach flankujących. Na podstawie bezwzględnych poziomów ekspresji kolejnych egzonów wyliczono procentowy udział wtrąceń splicingowych (ang. *the percentage of splicing inclusions*, PSI). Za istotne statystycznie uznano przypadki DAS spełniające następujące progi odcięcia: wartość $p < 0,05$ i $|\Delta PSI| > 0,1$. Na podstawie programu SUPPA2 wszystkie wykryte przypadki DAS podzielono na siedem kategorii: alternatywne miejsce akceptorowe (A3), alternatywne miejsce donorowe (A5), wzajemnie wykluczające się egzony (MX), zachowanie intronu (RI), pominięcie egzonu (SE), splicing pierwszego egzonu (AF) oraz splicing ostatniego egzonu (AL). Przypadki DAS biorące udział w procesach fizjologicznych, istotnych dla funkcjonowania komórek lutealnych, zwizualizowano przy użyciu programu ggsashimi wer. 0.5.0 (GARRIDO-MARTÍN i in. 2018).

4.4.5. Analiza allelicznej specyficzności ekspresji

Badanie przypadków ASE oparto na ocenie zmienności alternatywnych wersji SNV, poprzez zmierzenie frekwencji alleli oraz ich zróżnicowania między komórkami

hodowanymi w obecności chemeryny, a grupą kontrolną. W procesie wywoływania SNV (ang. *SNV calling*) zastosowano oprogramowania Picard wer. 2.6.0 (BROAD INSTITUTE 2021) oraz rMATS-DVR (ang. *replicate multivariate analysis of transcript splicing and discovery of differential variants in RNA*) wer. 1.0.0 (WANG i in. 2017), wykorzystujące pakiet GATK (ang. *genome analysis toolkit*) wer. 3.6.0 (DEPRISTO i in. 2011). Wcześniej utworzone pliki, zawierające przyrównania przetworzonych odczytów do genomu referencyjnego wer. 11.1.99, zrekalibrowano za pomocą programu Picard. Następnie zidentyfikowano potencjalne przypadki SNV we wszystkich replikatach prób i wykryto różnice we frekwencjach alleli między grupami badaną i kontrolną. Przypadki SNV o niskiej jakości odfiltrowano stosując standardowe parametry GATK. W dalszych analizach uwzględniono wyłącznie alternatywne warianty wykazujące ekspresję, czyli posiadające niezerową wartość alternatywnej frakcji alleli (ang. *alternative allele fraction*, AAF), w co najmniej połowie prób. Istotny wpływ chemeryny na proces ASE stwierdzono w przypadku SNV, dla których |ΔAAF| był większy od 0,1 oraz spodziewany odsetek wyników fałszywie dodatnich (ang. *false discovery rate*, FDR) był mniejszy od 0,05. Stosunek nierównowagi allelicznej kandydatów do ASE potwierdzono za pomocą testu zgodności dopasowania χ^2 (wartość $p < 0{,}05$). Lokalizację zidentyfikowanych przypadków ASE w obrębie elementów genów oraz dokładny wpływ SNV na proces transkrypcji kodowanych białek określono przy użyciu programu Ensembl VEP (ang. *variant effect predictor*) (MCLAREN i in. 2016).

4.4.6. Analiza funkcjonalna zidentyfikowanych przypadków DEG, DEL, DAS i ASE

Poszerzone metaanalizy genetyczne (ang. *gene enrichment analyses*) przeprowadzono oddzielnie dla wyników wchodzących w skład każdej z publikacji, tworzących niniejszą rozprawę doktorską. Podczas opracowywania wyników do publikacji 1, analizę aktywności produktów DEG, ich występowania w przedziałach komórkowych oraz udziału w szlakach metabolicznych i sygnałowych przeprowadzono przy użyciu oprogramowania g:Profiler (REIMAND i in. 2016) oraz KOBAS (ang. *ko-based annotation system*) wer. 3.0 (WU i in. 2006), przeszukujących internetowe bazy danych Encyklopedii Kyoto Genów i Genomów (ang. *Kyoto Encyclopedia of Genes and Genomes*, KEGG) (KANEHISA i in. 2017; KANEHISA i Sato 2020), Konsorcjum Ontologii Genów (ang. *Gene Ontology Consortium*, GO) (ASHBURNER i in. 2000; Carbon i in. 2017) i the Reactome

Knowledgebase (FABREGAT i in. 2018). W trakcie analizy wyników uzyskanych podczas prac nad publikacją 2, w celu kompleksowego przeanalizowania biologicznych funkcji zidentyfikowanych przypadków DEL, DAS i ASE, wykorzystano dane dotyczące powiązanych z nimi DEG. Udział wyselekcjonowanych genów w procesach biologicznych (dla każdego ze zjawisk osobno) sprawdzono programem KOBAS przeszukując trzy wymienione bazy danych. Wszystkie analizy funkcjonalne przeprowadzono stosując test Fishera (wykorzystujący rozkład hipergeometryczny) z poprawką Benjaminiego i Hochberga (FDR < 0,05). Mapy sieci molekularnych z bazy KEGG wygenerowano przy pomocy pakietu R Pathview wer. 1.30.1 (LUO i BROUWER 2013). Podczas wizualizowania wyników analiz DEL, DAS i ASE, dane dotyczące genów docelowych połączono w klastry oraz ponownie obliczono wartości istotności statystycznej dla przegrupowanych danych (FDR < 0,05).

4.5. Walidacja ekspresji wybranych genów oraz długich niekodujących RNA metodą qPCR

W celu weryfikacji wyników uzyskanych w trakcie analiz *in silico* opisanych w obu publikacjach, przeprowadzono ilościowe reakcje qPCR w maksymalnie zbliżonych warunkach. Do syntezy matryc cDNA wykorzystano po 500 ng każdej z próbek użytej w RNA-Seq. Reakcje odwrotnej transkrypcji przeprowadzono z użyciem zestawu Omniscript RT Kit (Qiagen, Germantown, Maryland, USA), starterów Oligo(dT) (Roche, Penzberg, Niemcy) oraz inhibitora RNaz (Invitrogen, Waltham, Massachusetts, USA), uzyskując całkowitą objętość 10 µL każdej z mieszanin. Reakcje prowadzono w termocyklerze w temperaturze 37 °C przez 1 godzinę i kończono inkubacją w temperaturze 93 °C przez 5 min. Uzyskane matryce cDNA poddano amplifikacji w obecności starterów specyficznych dla wybranych genów (publikacja 1): *CASP3* (ang. *caspase 3*), *HSD3B7* (ang. *hydroxy-δ-5-steroid dehydrogenase, 3 β- and steroid δ-isomerase 7*), *IL1B* (ang. *interleukin 1β*) i *PTGS2* (ang. *prostaglandin-endoperoxide synthase 2*); oraz transkryptów (publikacja 2): *CL.9638.3*, *CL.12742.3*, *ENSSSCT00000075362* i *ENSSSCT00000078829*. Jako geny referencyjne o konstytutywnej ekspresji (tzn. stałej ekspresji niezależnie od zmiennych czynników działających na komórkę) zastosowano geny *ACTB* (ang. *β-actin*) oraz *GADPH* (ang. *glyceraldehyde-3-phosphate dehydrogenase*). Sekwencje starterów dla

badanych genów i transkryptów opracowano z użyciem programu Primer Express 3 (Applied Biosystems, Waltham, Massachusetts, USA) oraz zestawiono w Tabelach 1 w obu publikacjach. Reakcje qPCR każdej próby przeprowadzono w technicznych duplikatach przy użyciu termocyklerów: Applied Biosystems 7300 Real-Time PCR System (Applied Biosystems, Foster, California, USA) w przypadku walidacji w publikacji 1 oraz AriaMx Real-Time PCR (Agilent Technology, St. Clara, California, USA) dla walidacji w publikacji 2. Do qPCR przygotowano mieszaniny o objętości 20 µL, składające się z 40 µg (publikacja 1) lub 30 µg (publikacja 2) cDNA, 400 nM starterów, 12,5 µL Power SYBR Green PCR Master Mix (Applied Biosystems, Foster, California, USA) i wody wolnej od RNaz. Reakcje qPCR w publikacji 1 przeprowadzono zgodnie z następującą procedurą: wstępna denaturacja cDNA i aktywacja enzymów w temp. 95 °C przez 10 min, 40 cykli denaturacji w temp. 95 °C przez 15 s, hybrydyzacji w temp. 60 °C przez 1 min oraz elongacji w temp. 72 °C przez 1 min. Dla starterów *ACTB* i *HSD3B7* temperaturę hybrydyzacji zmieniono na 61 °C. Reakcje qPCR w publikacji 2 przeprowadzono w następujących warunkach: denaturacja cDNA i aktywacja enzymów w temp. 95 °C przez 3 min, 40 cykli denaturacji w temp. 95 °C przez 30 s, hybrydyzacji w temp. 66 °C przez 1 min i elongacji w temp. 72 °C przez 45 s. Dla starterów *ENSSSCT00000075362* wstępną denaturację przedłużono do 10 min, z obniżeniem temp. hybrydyzacji do 60 °C. Kontrole negatywne przygotowano poprzez zastąpienie cząsteczek cDNA ekwiwalentem wody. Specyficzność procesu amplifikacji zbadano pod koniec qPCR za pomocą analizy krzywej topnienia. Względne poziomy ekspresji walidowanych genów i transkryptów obliczano metodą ΔΔCT (LIVAK i SCHMITTGEN 2001) i normalizowano za pomocą średnich geometrycznych ekspresji genów referencyjnych. Wyniki qPCR opracowano statystycznie w środowisku R przy użyciu testu t-Studenta z założonym progiem istotności statystycznej $p < 0,05$ i przedstawiono jako wartości średnie ± błąd standardowy średniej (SEM) w publikacji 1 oraz $\log_2$(krotności zmian) ± przedział ufności w publikacji 2.

5. WYNIKI

5.1. Wyniki wysokoprzepustowego sekwencjonowania transkryptomów

W efekcie przeprowadzonej procedury sekwencjonowania RNA-Seq uzyskano 1.154.494.264 surowych i sparowanych odczytów. Po wstępnym przetworzeniu, polegającym na wycięciu z odczytów sekwencji adaptorowych, usunięciu regionów o niskiej jakości oraz wyrównaniu długości wszystkich sekwencji do 90 nt, uzyskano 1.024.385.352 przetworzonych odczytów. W trakcie mapowania przetworzonych sekwencji względem genomu referencyjnego *S. scrofa* wer. 11.1.91 przyrównano 1.018.891.762 odczytów, w tym średnio 96,52% sekwencji przyrównano unikatowo, a 3,48% zmapowano do wielu loci (publikacja 1, Tabela 2). Podczas kolejnych badań i po zaktualizowaniu genomu referencyjnego *S. scrofa* do wer. 11.1.99 zmapowano 1.021.940.610 przetworzonych odczytów, z czego 96,24% unikatowo oraz 3,76% do wielu loci (publikacja 2, Tabela 2).

Wśród wszystkich sekwencji uzyskanych z bibliotek RNA-Seq i dwukrotnie przyrównanych do dwóch wersji genomu referencyjnego świni domowej, średnio 56,63% (wer. 11.1.91) i 59,52% (wer. 11.1.99) nukleotydów dopasowano do regionów CDS, 22,08% i 23,32% do regionów UTR, 2,59% i 3,35% do sekwencji intronowych oraz 18,69% i 13,81% zlokalizowano poza genami (publikacja 1, Rycina 2; publikacja 2, Tabela 2). Analizy PCA oraz skupień wykazały dużą homogenność profili ekspresji wewnątrz grup badawczej i kontrolnej oraz wysoki stopień zróżnicowania pomiędzy grupami (publikacja 1, Rycina 3). Sumarycznie zidentyfikowano 16.612 regionów TAR w co najmniej połowie przeanalizowanych prób.

5.2. Wpływ chemeryny na ekspresję genów kodujących białka

Analiza ekspresji genów, ujawniła 509 DEG kodujących 721 transkryptów (publikacja 1, Tabela Suplementowa 1). Wśród wszystkich zidentyfikowanych DEG, 301 wykazało wzrost ekspresji w próbach inkubowanych z chemeryną, natomiast 208 obniżenie ekspresji. Różnice pomiędzy pomiarami wykonanymi dla grup eksperymentalnej oraz kontrolnej zwizualizowano w postaci wykresu MA (ang. *MA plot*; implementacja wykresu Blanda–Altmana dla biologicznych analiz wysokoprzepustowych) oraz wykresu typu

wulkan (ang. *volcano plot*) w publikacji 1 na Rycinie 4. Zlogarytmowane wartości krotności zmiany ekspresji mieściły się w przedziale od –2,17 (*PANK1*, ang. *pantothenate kinase 1*) do 5,45 (*IL1B*). 98 DEG zaangażowanych w szlaki sygnałowe i metaboliczne, istotne dla funkcjonowania ciałka żółtego, przestawiono w postaci mapy ciepła (ang. *heat map*) w publikacji 1 na Rycinie 5.

5.3. Wpływ chemeryny na ekspresję długich niekodujących RNA

Identyfikację nowych cząsteczek lncRNA przeprowadzono na podstawie szczegółowej analizy wstępnie przetworzonych danych RNA-Seq. Początkowo, z puli analizowanych sekwencji transkryptów usunięto te, które oznaczono w bazie danych Ensembl jako „biotyp kodujący białko" (ang. *protein-coding biotype*), a także RNA jednoegzonowe i krótsze od 200 nt. W efekcie pierwszego etapu filtracji w puli pozostało 25.252 potencjalnych kandydatów na lncRNA, w tym 9.340 cząsteczek oznaczonych w bazie Ensembl jako „biotyp lncRNA". Nieopisane sekwencje przeanalizowano pod kątem potencjału kodowania przy pomocy czterech programów – CPAT, CPC2, Feelnc i PLEK, co umożliwiło zawężenie puli nowych lncRNA do 1.759 cząsteczek. Spośród nich, 107 sekwencji pozostało po procesie przyrównania do rekordów rodzin białkowych w bazie danych Pfam, przy użyciu programu HMMER. Efekty działania wszystkich wspomnianych programów zestawiono w postaci diagramu Venna w publikacji 2 na Rycinie 2. Żaden z wykrytych transkryptów nie został dopasowany do modeli małych RNA w bazie danych Rfam przez program Infernal cmscan.

Analiza różnic ekspresji przeprowadzona na łączonej puli nowopoznanych lncRNA oraz makromolekuł uprzednio opisanych w bazie Ensembl wykazała, iż ekspresja 24 cząsteczek (kodowanych przez 12 genów) uległa modyfikacji pod wpływem chemeryny. Spośród nich, 20 lncRNA poznano wcześniej, a pozostałe 4 zidentyfikowano w niniejszym badaniu. Profile ekspresji czterech wybranych cząsteczek DEL (*CL.9638.3*, *CL.12742.3*, *ENSSSCT00000028661* i *ENSSSCT00000087136*) zwizualizowano na tle ekspresji DEG w publikacji 2 na Rycinie 3. Pod wpływem badanej chemokiny ekspresja 16 DEL uległa obniżeniu, a 8 podwyższeniu. Parametry wykrytych DEL zestawiono w publikacji 2 w Tabeli 3.

Analiza *cis*-interakcji ujawniła 11 przypadków DEL oddziaływujących z 6 DEG na podstawie ich ko-lokalizacji w genomie świni domowej. Wśród nich 3 interakcje zaklasyfikowano jako pokrywające się z DEG. W 5 interakcjach gen DEL wystąpił w regionie do 10.000 pz poprzedzającym (ang. *downstream*) DEG, a w 3 interakcjach w regionie następującym (ang. *upstream*) po DEG (publikacja 2, Tabela 4).

Analiza *trans*-interakcji wykazała 9.295 przypadków ko-ekspresji DEL z 469 DEG (przy założeniu silnej zależności wynikającej z korelacji Pearsona $|r| > 0,9$). Dodatkowe analizy wiązań mRNA–lncRNA oraz dopasowania przestrzennego białko–lncRNA ujawniły, odpowiednio, 1.393 oraz 115 relacji, ograniczając liczbę wykrytych *trans*-interakcji do 1.440 przypadków powiązanych z 300 DEG (publikacja 2; Tabela Suplementowa 1). Spośród nich, 676 przypadków wykazało pozytywną, natomiast 764 negatywną korelację. Wszystkie opisane interakcje bazujące zarówno na ko-lokalizacji oraz ko-ekspresji par DEL–DEG, z uwzględnieniem zmian ekspresji pod wpływem chemeryny, zwizualizowano w publikacji 2 na Rycinie 4.

5.4. Modyfikacje procesu alternatywnego składania transkryptów pod wpływem chemeryny

Zastosowana procedura, wykorzystująca program SUPPA2, umożliwiła wykrycie 43.918 zdarzeń splicingowych, w tym 137 przypadków DAS, zaobserwowanych na podstawie porównania frekwencji przypadków alternatywnego splicingu między grupą badaną oraz kontrolną. Wszystkie zidentyfikowane przypadki DAS zlokalizowano w obrębie 92 DEG (publikacja 2, Tabela Suplementowa 2) oraz przypisano do 7 grup: 13 sklasyfikowano jako A5, 12 jako A3, 2 jako MX, 13 jako RI, 21 jako SE, 60 jako AF, a 16 jako AL (publikacja 2, Rycina 5). Wybrane przypadki alternatywnego składania transkryptów, występujące w obrębie istotnych dla funkcjonowania ciałka żółtego świni domowej genów: *CXCL12* (ang. *C–X–C motif chemokine ligand 12*), *MCL1* (ang. *MCL1 apoptosis regulator, BCL2 family member*), *FHL1* (ang. *four and a half LIM domains 1*) i *SLA-DQB1* (ang. *porcine tissue compatibility antigen type DQB1*), przedstawiono w publikacji 2 na Rycinie 6.

5.5. Oddziaływanie chemeryny na zmienność allelicznej specyficzności ekspresji

Podczas analiz *in silico* dziesięciu bibliotek RNA-Seq stwierdzono 151.603 przypadki SNV w transkryptomie komórek lutealnych świni domowej. W wyniku przeprowadzenia wieloetapowej procedury filtrowania, opartej na standardowych parametrach GATK oraz analizie wartości AAF w próbach, uzyskano 37 SNV będących kandydatami do ASE. Pośród zawężonej puli wariantów, 18 wykazało istotną statystycznie zmienność ASE pomiędzy grupą badaną oraz kontrolną. Zgodnie z adnotacjami programu VEP: 5 przypadków ASE opisano uprzednio w bazie Ensembl, tj. *rs1110933145*, *rs322079801*, *rs332564681*, *rs346401196* i *rs55619031*, natomiast 13 przypadków było nowo wykrytych w trakcie niniejszych badań. Wszystkie zidentyfikowane ASE wystąpiły w regionach 16 genów, kodujących 38 transkryptów (publikacja 2, Tabela Suplementowa 3).

Zgodnie z danymi wygenerowanymi przez program VEP, 11 przypadków ASE powiązano z niekodującymi regionami transkryptów kodowanych przez 11 genów, co powoduje trudne do przewidzenia konsekwencje biologiczne. Cztery ASE, powiązane z niskim prawdopodobieństwem modyfikacji kodowanego białka, wystąpiły w regionach CDS transkryptów kodowanych przez geny *EIF3F* (ang. *eukaryotic translation initiation factor 3 subunit F*), *OSBP* (ang. *oxysterol binding protein*), *PIK3C2A* (ang. *phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2*α) i *SLC25A24* (ang. *solute carrier family 25 member 24*). Cztery kolejne ASE zidentyfikowano w sekwencjach transkryptów kodowanych przez geny *EIF3F*, *HPS5* (ang. *HPS5 biogenesis of lysosomal organelles complex 2 subunit 2*) oraz *OSBP*, jednocześnie przewidując ich wpływ na zmianę aktywności kodowanych białek. Wpływ jednego ASE na 3 transkrypty kodowane przez gen *CXCL2* (ang. *C–X–C motif chemokine ligand 2*) opisano jako destrukcyjny i powodujący utratę funkcji przez kodowane białko, poprzez modyfikację miejsca akceptorowego drugiego intronu. Zgodnie z wynikami uzyskanymi w oparciu o jeden z modułów VEP – SIFT (NG i HENIKOFF 2003), spośród 4 mutacji nonsensownych mogących wpływać na konformację produktu białkowego, 2 mają działanie uszkadzające (geny *HPS5* i *OSBP*), a 2 mogą być tolerowane przez komórki (geny *EIF3F* i *OSBP*). Konsekwencje molekularne wystąpienia poszczególnych ASE w strukturach genów i transkryptów podsumowano w publikacji 2 na Rycinie 7.

5.6. Analizy funkcjonalne uzyskanych wyników *in silico*

5.6.1. Wpływ chemeryny na ekspresję genów kodujących funkcjonalne białka. Adnotacje funkcjonalne DEG

Analizy funkcjonalne, przeprowadzone w oparciu o dane z bazy GO, wykazały, iż z puli wszystkich DEG, 457 powiązano z co najmniej jednym terminem (in. podkategorią) w obrębie trzech kategorii: proces biologiczny (ang. *biological process*, BP), funkcja molekularna (ang. *molecular function*, MF) oraz komponent komórkowy (ang. *cellular component*, CC). W obrębie kategorii BP, DEG przypisano do 407 podkategorii, takich jak: „odpowiedź komórkowa na bodziec" (ang. *cellular response to stimulus*), „regulacja transdukcji sygnału" (ang. *regulation of signal transduction*), „funkcjonowanie układu odpornościowego" (ang. *immune system process*), „programowana śmierć komórki" (ang. *programmed cell death*), „odpowiedź na cytokiny" (ang. *response to cytokine*), „produkcja cytokin" (ang. *cytokine production*) i „sygnalizacja kinazy I-kappaB/NF-kappa B" (ang. *I-kappaB kinase/NF-kappa B signaling*). W zakresie kategorii MF, największą ilość DEG przypisano następującym podkategoriom: „wiązanie receptora cytokinowego" (ang. *cytokine receptor binding*), „aktywność regulatora receptora" (ang. *receptor regulator activity*), „wiązanie receptora czynnika wzrostu" (ang. *growth factor receptor binding*) i „wiązanie receptora chemokinowego CCR" (ang. *CCR chemokine receptor binding*). W kategorii CC, spośród 15 podkategorii, do których przypisano DEG, istotnymi biologicznie wydają się być: „region zewnątrzkomórkowy" (ang. *extracellular region*), „powierzchnia komórki" (ang. *cell surface*), „cytoplazma" (ang. *cytoplasm*) i „połączenie komórka-komórka" (ang. *cell-cell junction*). Szczegółowe dane zestawiono w publikacji 1 na Rycinie 6.

Analiza ścieżek metabolicznych i sygnałowych KEGG ujawniła, iż produkty 286 DEG biorą udział w 51 ścieżkach biologicznych. Wśród nich, 32 szlaki były bezpośrednio związane z procesami patologicznymi, takimi jak kancerogeneza, choroby autoimmunologiczne, infekcje wywoływane przez różne patogeny (bakterie, wirusy, protisty, priony). Wspomniane szlaki nie były analizowane ze względu na brak związku z badaną tkanką oraz z celem przeprowadzanych badań, czyli zrozumieniem fizjologicznego wpływu chemeryny na komórki lutealne świni domowej. Skupiono się na sprawdzeniu następujących szlaków metabolicznych i sygnałowych spełniających założony próg

istotności statystycznej: „ścieżka sygnałowa TNF" (ang. *TNF signaling pathway*), „interakcje cytokina-receptor cytokinowy" (ang. *cytokine-cytokine receptor interaction*), „ścieżka sygnałowa NF-kappa B" (ang. *NF-kappa B signaling pathway*), „ścieżka sygnałowa chemokin" (ang. *chemokine signaling pathway*), „ścieżka sygnałowa MAPK" (ang. *MAPK signaling pathway*), „cząsteczki adhezji komórek" (ang. *cell adhesion molecules – CAMs*) oraz „apoptoza" (ang. *apoptosis*). Dodatkowo przeanalizowano szlaki KEGG ważne dla funkcjonowania ciałka żółtego, dla których analiza funkcjonalna nie osiągnęła zakładanego progu istotności: „biosynteza hormonów steroidowych" (ang. *steroid hormone biosynthesis*), „steroidogeneza jajnikowa" (ang. *ovarian steroidogenesis*), „metabolizm kwasu arachidonowego" (ang. *arachidonic acid metabolism*) i „szlak sygnałowy JAK-STAT" (ang. *JAK-STAT signaling pathway*). Uprzednio wspomnianą mapę ciepła, prezentującą profile ekspresji DEG (publikacja 1, Rycina 5), wzbogacono o wykres strumieniowy Sankeya przypisujący DEG do szlaków sygnałowych i metabolicznych KEGG, w których uczestniczą ich produkty. Wizualizacje udziału produktów DEG w wymienionych szlakach, uwzględniające zmiany ich ekspresji pod wpływem chemeryny, przedstawiono w publikacji 1 na Rycinach Suplementowych 1–11.

Analiza rekordów bazy Reactome wykazała znaczący udział 261 DEG w 19 procesach molekularnych. Z tej puli warto wyróżnić: „sygnalizację cytokin w układzie odpornościowym" (ang. *cytokine signaling in immune system*), „programowaną śmierć komórki" (ang. *programmed cell death*), „TAK1 aktywuje NFkB przez fosforylację i aktywację kompleksu IKK" (ang. *TAK1 activates NFkB by phosphorylation and activation of IKKs complex*), „sygnalizację interleukiny 1" (ang. *interleukin-1 signaling*) i „receptory chemokin wiążą chemokiny" (ang. *chemokine receptors bind chemokines*).

5.6.2. Wpływ chemeryny na ekspresję genów kodujących funkcjonalne białka. Adnotacje funkcjonalne DEL, DAS i ASE

Przebieg analizy funkcjonalnej, przeprowadzonej w trakcie badań wykonanych w ramach publikacji 2, zmodyfikowano względem analogicznej procedury zastosowanej w ramach publikacji 1, w celu podsumowania i połączenia wyników uzyskanych dla trzech różnych zjawisk towarzyszących transkrypcji. Proces przypisania DEL, DAS i ASE do kategorii GO i Reactome oraz szlaków w bazie KEGG z użyciem programu KOBAS

przeprowadzono oddzielnie dla każdego ze zjawisk: *cis*-interakcji DEL z DEG, *trans*-interakcji DEL z DEG, obecności DAS lub ASE w obrębie DEG. Analiza ontologii genów wykazała istotny udział 5, 273, 39 oraz 10 genów (w kolejności wymienionej powyżej), które weszły w skład 30, 159, 2 i 39 podkategorii w obrębie BP; 5, 34, 1 i 17 podkategorii MF oraz 4, 24, 3 i 12 podkategorii CC (publikacja 2, Tabela Suplementowa 4).

Wykryte DEL, DAS i ASE zgrupowano w klastry odpowiadające przynależności do ścieżek KEGG oraz ponownie przeprowadzono analizę statystyczną przegrupowanych danych. Założony poziom istotności statystycznej spełniły 84 szlaki metaboliczne lub transdukcji sygnałów. Ponownie pominięto ścieżki związane z innymi tkankami lub procesami patologicznymi, skupiając się na przeanalizowaniu szlaków wpływających na funkcjonowanie jajnika, ciałka żółtego lub samych komórek lutealnych. Szczegółowe informacje dotyczące zawężonej puli 35 procesów metabolicznych lub transdukcji sygnałów wpływających na fizjologię ciałka żółtego zestawiono w publikacji 2 w Tabeli 5 oraz na Rycinach Suplementowych 1–18. Dodatkowo przestudiowano istotny dla funkcjonowania komórek lutealnych szlak „metabolizmu kwasu arachidonowego", dla którego analiza funkcjonalna nie uzyskała zakładanego progu istotności statystycznej (publikacja 2, Rycina Suplementowa 19).

Analiza z udziałem rekordów bazy Reactome wykazała istotny udział *cis*-interakcji DEL z DEG w 3 ścieżkach molekularnych, *trans*-interakcji DEL z DEG w 20 ścieżkach oraz ASE w obrębie DEG w 1 ścieżce (publikacja 2, Tabela Suplementowa 4). Udział produktów genów zawierających DAS nie wykazał istotnego wpływu na żadną ścieżkę w bazie Reactome.

5.7. Walidacja wybranych wyników *in silico* metodą qPCR

W celu weryfikacji danych uzyskanych w wyniku sekwencjonowania RNA-Seq, wybrano cztery DEG oraz cztery DEL. Zaobserwowane, pod wpływem chemeryny, zmiany ekspresji zostały w dużej mierze potwierdzone metodą qPCR. Wzorce ekspresji genów *HSD3B7*, *IL1B* i *PTGS2* (publikacja 1, Rycina 7) oraz lncRNA *CL.9638.3*, *CL.12742.3*, *ENSSSCT00000075362* i *ENSSSCT00000078829* (publikacja 2, Rycina 8) były zgodne z wynikami RNA-Seq. Zmiany w poziomie ekspresji *CASP3* pomiędzy grupami badaną i kontrolną nie osiągnęły progu istotności statystycznej (wartość $p = 0{,}06$), jednakże

potwierdziły trend, wcześniej ujawniony podczas analizy danych RNA-Seq. Wyniki qPCR potwierdziły dokładność oraz wiarygodność zastosowanej metody wysokoprzepustowego sekwencjonowania transkryptomów.

6. PODSUMOWANIE I WNIOSKI

Cykl publikacji składających się na niniejszą rozprawę doktorską oparty jest na eksperymentach wykorzystujących wysokoprzepustowe sekwencjonowanie do przeanalizowania wpływu chemeryny na transkryptom komórek lutealnych świni domowej. Zastosowana technologia umożliwiła szczegółowe przeanalizowanie subtelnych modyfikacji genetycznych, wywołanych przez chemerynę. Wyniki otrzymane w trakcie szczegółowych analiz profili transkryptomicznych pozwoliły zobrazować efekt działania chemeryny na komórki lutealne świni domowej w środkowolutealnej fazie cyklu rujowego w warunkach *in vitro*. Dane dotyczące wpływu chemeryny na profile ekspresji genów i długich niekodujących RNA oraz ich interakcji z genami, jak również na zmienność procesów alternatywnego składania transkryptów i allelicznej specyficzności ekspresji, pozwoliły na sformułowanie następujących wniosków:

- zmiany w transkryptomie wywołane działaniem chemeryny mogą prowadzić do aktywacji szlaków transdukcji sygnałów MAPK (w tym czynników transkrypcyjnych NF-κB) i JAK/STAT oraz pośrednio do wzrostu produkcji wielu cytokin i chemokin przez komórki lutealne,
- zmiany ekspresji genów oraz procesów regulujących transkrypcję, pod wpływem chemeryny, mogą aktywować zarówno mechanizmy pro-apoptotyczne (lub nawet pro-nekroptyczne) oraz pro-przeżyciowe w komórkach lutealnych,
- chemeryna, wpływając poprzez transkryptom na regulację steroidogenezy, lokalną produkcję czynników prozapalnych, angiogenezę oraz produkcję prostaglandyn, kontroluje funkcjonowanie ciałka żółtego,
- chemeryna wydaje się być hormonem łączącym regulację procesów metabolicznych i rozrodczych.

7. BIBLIOGRAFIA

1. AKINS EL, MORRISSETTE MC. 1968. Gross ovarian changes during estrous cycle of swine. Am J Vet Res. 29:1953–1957.
2. ALESSIO E, BONADIO RS, BUSON L, CHEMELLO F, CAGNIN S. 2020. A Single Cell but Many Different Transcripts: A Journey into the World of Long Non-Coding RNAs. Int J Mol Sci. 21:302. https://doi.org/10.3390/IJMS21010302.
3. ANDREWS S. 2010. FastQC: A quality control tool for high throughput sequence data. Babraham Bioinforma. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (21.03.2021).
4. ASHBURNER M, BALL CA, BLAKE JA, BOTSTEIN D, BUTLER H, CHERRY JM, DAVIS AP, DOLINSKI K, DWIGHT SS, EPPIG JT, HARRIS MA, HILL DP, ISSEL-TARVER L, KASARSKIS A, LEWIS S, MATESE JC, RICHARDSON JE, RINGWALD M, RUBIN GM, SHERLOCK G. 2000. Gene ontology: Tool for the unification of biology. Nat Genet. 25:25–29. https://doi.org/10.1038/75556.
5. BARKER G, LIM R, RICE GE, LAPPAS M. 2012. Increased chemerin concentrations in fetuses of obese mothers and correlation with maternal insulin sensitivity. J Matern Fetal Neonatal Med. 25:2274–2280. https://doi.org/10.3109/14767058.2012.686540.
6. BATTLE A, KHAN Z, WANG SH, MITRANO A, FORD MJ, PRITCHARD JK, GILAD Y. 2015. Impact of regulatory variation from RNA to protein. Science. 347:664–667. https://doi.org/10.1126/SCIENCE.1260793.
7. BOLGER AM, LOHSE M, USADEL B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
8. BONDUE B, WITTAMER V, PARMENTIER M. 2011. Chemerin and its receptors in leukocyte trafficking, inflammation and metabolism. Cytokine Growth Factor Rev. 22:331–338. https://doi.org/10.1016/J.CYTOGFR.2011.11.004.
9. BOZAOGLU K, CURRAN JE, STOCKER CJ, ZAIBI MS, SEGAL D, KONSTANTOPOULOS N, MORRISON S, CARLESS M, DYER TD, COLE SA, GORING HHH, MOSES EK, WALDER K, CAWTHORNE MA, BLANGERO J, JOWETT JBM. 2010. Chemerin, a novel adipokine in the regulation of angiogenesis. J Clin Endocrinol Metab. 95:2476–2485. https://doi.org/10.1210/jc.2010-0042.
10. BROAD INSTITUTE. 2021. Picard Tools. http://broadinstitute.github.io/picard/ (12.05.2021).
11. CARBON S, DIETZE H, LEWIS SE, MUNGALL CJ, MUNOZ-TORRES MC, BASU S, CHISHOLM RL, DODSON RJ, FEY P, THOMAS PD, MI H, MURUGANUJAN A, HUANG X, POUDEL S, HU JC, ALEKSANDER SA, MCINTOSH BK, RENFRO DP, SIEGELE DA, ANTONAZZO G i in. 2017. Expansion of the gene ontology knowledgebase and resources: The gene ontology consortium. Nucleic Acids Res. 45:D331–D338. https://doi.org/10.1093/nar/gkw1108.
12. DE HENAU O, DEGROOT G-NN, IMBAULT V, ROBERT V, DE POORTER C, MCHEIK S, GALÉS C, PARMENTIER M, SPRINGAEL J-YY. 2016. Signaling Properties of Chemerin Receptors CMKLR1, GPR1 and CCRL2. PLoS One. 11:e0164179. https://doi.org/10.1371/journal.pone.0164179.
13. DEPRISTO MA, BANKS E, POPLIN R, GARIMELLA K V, MAGUIRE JR, HARTL C, PHILIPPAKIS AA, DEL ANGEL G, RIVAS MA, HANNA M, MCKENNA A, FENNELL TJ, KERNYTSKY AM, SIVACHENKO AY, CIBULSKIS K, GABRIEL SB, ALTSHULER D, DALY MJ. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 43:491–498. https://doi.org/10.1038/ng.806.
14. DE SOUSA ABREU R, PENALVA LO, MARCOTTE EM, VOGEL C. 2009. Global signatures of protein and mRNA expression levels. Mol Biosyst. 5:1512–1526. https://doi.org/10.1039/b908315d.
15. DOBIN A, DAVIS CA, SCHLESINGER F, DRENKOW J, ZALESKI C, JHA S, BATUT P, CHAISSON M, GINGERAS TR. 2013. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 29:15–21. https://doi.org/10.1093/bioinformatics/bts635.
16. DU X-Y, LEUNG LLK. 2009. Proteolytic regulatory mechanism of chemerin bioactivity. Acta Biochim Biophys Sin (Shanghai). 41:973–979. https://doi.org/10.1093/abbs/gmp091.
17. DU XY, ZABEL BA, MYLES T, ALLEN SJ, HANDEL TM, LEE PP, BUTCHER EC, LEUNG LL. 2009. Regulation of Chemerin Bioactivity by Plasma Carboxypeptidase N, Carboxypeptidase B (Activated Thrombin-activable Fibrinolysis Inhibitor), and Platelets. J Biol Chem. 284:751–758.

https://doi.org/10.1074/JBC.M805000200.

18. EARLY P, ROGERS J, DAVIS M, CALAME K, BOND M, WALL R, HOOD L. 1980. Two mRNAs can be produced from a single immunoglobulin μ gene by alternative RNA processing pathways. Cell. 20:313–319. https://doi.org/10.1016/0092-8674(80)90617-0.
19. EDDY SR. 2011. Accelerated profile HMM searches. PLoS Comput Biol. 7:e1002195 https://doi.org/10.1371/journal.pcbi.1002195.
20. EL-GEBALI S, MISTRY J, BATEMAN A, EDDY SR, LUCIANI A, POTTER SC, QURESHI M, RICHARDSON LJ, SALAZAR GA, SMART A, SONNHAMMER ELL, HIRSH L, PALADIN L, PIOVESAN D, TOSATTO SCE, FINN RD. 2019. The Pfam protein families database in 2019. Nucleic Acids Res. 47:D427–D432. https://doi.org/10.1093/nar/gky995.
21. ERNST MC, SINAL CJ. 2010. Chemerin: at the crossroads of inflammation and obesity. Trends Endocrinol Metab. 21:660–667. https://doi.org/10.1016/j.tem.2010.08.001.
22. FABREGAT A, JUPE S, MATTHEWS L, SIDIROPOULOS K, GILLESPIE M, GARAPATI P, HAW R, JASSAL B, KORNINGER F, MAY B, MILACIC M, ROCA CD, ROTHFELS K, SEVILLA C, SHAMOVSKY V, SHORSER S, VARUSAI T, VITERI G, WEISER J, WU G, STEIN L, HERMJAKOB H, D'EUSTACHIO P. 2018. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 46:D649–D655. https://doi.org/10.1093/nar/gkx1132.
23. FRASER L, PAUKSZTO Ł, MAŃKOWSKA A, BRYM P, GILUN P, JASTRZĘBSKI JP, PAREEK CS, KUMAR D, PIERZCHAŁA M. 2020. Regulatory potential of long non-coding rnas (LncRNAs) in boar spermatozoa with good and poor freezability. Life. 10:1–26. https://doi.org/10.3390/life10110300.
24. FRAZEE AC, PERTEA G, JAFFE AE, LANGMEAD B, SALZBERG SL, LEEK JT. 2015. Ballgown bridges the gap between transcriptome assembly and expression analysis. Nat Biotechnol. 33:243–246. https://doi.org/ 10.1038/nbt.3172.
25. GARCES MF, SANCHEZ E, ACOSTA BJ, ANGEL E, RUÍZ AI, RUBIO-ROMERO JA, DIÉGUEZ C, NOGUEIRAS R, CAMINOS JE. 2012. Expression and regulation of chemerin during rat pregnancy. Placenta. 33:373–378. https://doi.org/10.1016/j.placenta.2012.02.007.
26. GARRIDO-MARTÍN D, PALUMBO E, GUIGÓ R, BRESCHI A. 2018. ggsashimi: Sashimi plot revised for browser- and annotation-independent splicing visualization. PLOS Comput Biol. 14:e1006360. https://doi.org/10.1371/journal.pcbi.1006360.
27. GAUR U, LI K, MEI S, LIU G. 2013. Research progress in allele-specific expression and its regulatory mechanisms. J Appl Genet. 54:271–283. https://doi.org/10.1007/S13353-013-0148-Y.
28. GORALSKI KB, MCCARTHY TC, HANNIMAN EA, ZABEL BA, BUTCHER EC, PARLEE SD, MURUGANANDAN S, SINAL CJ. 2007. Chemerin, a Novel Adipokine That Regulates Adipogenesis and Adipocyte Metabolism. J Biol Chem. 282:28175–28188. https://doi.org/10.1074/jbc.M700793200.
29. GUDELSKA M, DOBRZYŃ K, KIEŻUN M, RYTELEWSKA E, KISIELEWSKA K, KAMIŃSKA B, KAMIŃSKI T, SMOLIŃSKA N. 2020. The expression of chemerin and its receptors (CMKLR1, GPR1, CCRL2) in the porcine uterus during the oestrous cycle and early pregnancy and in trophoblasts and conceptuses. Animal. 14:2116–2128. https://doi.org/10.1017/S175173112000097X.
30. JOVANOVIC M, ROONEY MS, MERTINS P, PRZYBYLSKI D, CHEVRIER N, SATIJA R, RODRIGUEZ EH, FIELDS AP, SCHWARTZ S, RAYCHOWDHURY R, MUMBACH MR, EISENHAURE T, RABANI M, GENNERT D, LU D, DELOREY T, WEISSMAN JS, CARR SA, HACOHEN N, REGEV A. 2015. Dynamic profiling of the protein life cycle in response to pathogens. Science. 347:1259038-1–7. https://doi.org/10.1126/SCIENCE.1259038.
31. KALVARI I, NAWROCKI EP, ONTIVEROS-PALACIOS N, ARGASINSKA J, LAMKIEWICZ K, MARZ M, GRIFFITHS-JONES S, TOFFANO-NIOCHE C, GAUTHERET D, WEINBERG Z, RIVAS E, EDDY SR, FINN RD, BATEMAN A, PETROV AI. 2021. Rfam 14: expanded coverage of metagenomic, viral and microRNA families. Nucleic Acids Res. 49:D192-D200. https://doi.org/10.1093/nar/gkaa1047.
32. KAMIŃSKI T, SIAWRYS G, OKRASA S, PRZAŁA J. 1999. Action of the opioid agonist FK 33-824 on porcine small and large luteal cells from the mid-luteal phase: effect on progesterone, cAMP, cGMP and inositol phosphate release. Anim Reprod Sci. 56:245–257. https://doi.org/10.1016/S0378-4320(99)00047-0.
33. KANEHISA M, FURUMICHI M, TANABE M, SATO Y, MORISHIMA K. 2017. KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 45:D353–D361. https://doi.org/10.1093/nar/gkw1092.

34. KANEHISA M, SATO Y. 2020. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci. 29:28–35. https://doi.org/10.1002/pro.3711.
35. KANG YJ, YANG DC, KONG L, HOU M, MENG YQ, WEI L, GAO G. 2017. CPC2: A fast and accurate coding potential calculator based on sequence intrinsic features. Nucleic Acids Res. 45:W12–W16. https://doi.org/10.1093/nar/gkx428.
36. KENNEDY AJ, DAVENPORT AP. 2018. International Union of Basic and Clinical Pharmacology CIII: Chemerin Receptors CMKLR1 (Chemerin$_1$) and GPR1 (Chemerin$_2$) Nomenclature, Pharmacology, and Function. Pharmacol Rev. 70:174–196. https://doi.org/10.1124/pr.116.013177.
37. KHANSEFID M, PRYCE JE, BOLORMAA S, CHEN Y, MILLEN CA, CHAMBERLAIN AJ, VANDER JAGT CJ, GODDARD ME. 2018. Comparing allele specific expression and local expression quantitative trait loci and the influence of gene expression on complex trait variation in cattle. BMC Genomics. 19:793. https://doi.org/10.1186/S12864-018-5181-0/.
38. KISIELEWSKA K, RYTELEWSKA E, GUDELSKA M, KIEŻUN M, DOBRZYŃ K, BOGUS-NOWAKOWSKA K, KAMIŃSKA B, SMOLIŃSKA N, KAMIŃSKI T. 2020a. Expression of chemerin receptors CMKLR1, GPR1 and CCRL2 in the porcine pituitary during the oestrous cycle and early pregnancy and the effect of chemerin on MAPK/Erk1/2, Akt and AMPK signalling pathways. Theriogenology. 157:181–198. https://doi.org/10.1016/j.theriogenology.2020.07.032.
39. KISIELEWSKA K, RYTELEWSKA E, GUDELSKA M, KIEŻUN M, DOBRZYŃ K, BOGUS-NOWAKOWSKA K, KAMIŃSKA B, SMOLIŃSKA N, KAMIŃSKI T. 2020b. Relative abundance of chemerin mRNA transcript and protein in pituitaries of pigs during the estrous cycle and early pregnancy and associations with LH and FSH secretion during the estrous cycle. Anim Reprod Sci. 219:106532. https://doi.org/10.1016/J.ANIREPROSCI.2020.106532.
40. KORNIENKO AE, GUENZL PM, BARLOW DP, PAULER FM. 2013. Gene regulation by the act of long non-coding RNA transcription. BMC Biol. 11:59. https://doi.org/10.1186/1741-7007-11-59.
41. KUROWSKA P, MLYCZYŃSKA E, BARBE A, MELLOUK N, DUPONT J, RAK A. 2018. The Adipokines in Domestic Animal Reproduction: Expression and Role in the Regulation of Ovarian Function. W: PAYAN-CARREIRA R. (red.). New Insights into Theriogenology. IntechOpen: 43–61. https://doi.org/10.5772/intechopen.81035.
42. KUROWSKA P, MLYCZYŃSKA E, DAWID M, SIERPOWSKI M, ESTIENNE A, DUPONT J, RAK A. 2021. Adipokines change the balance of proliferation/apoptosis in the ovarian cells of human and domestic animals: A comparative review. Anim Reprod Sci. 228:106737. https://doi.org/10.1016/J.ANIREPROSCI.2021.106737.
43. LEI J, SUN LC, HUANG S, ZHU C, LI P, HE J, MACKEY V, COY DH, HE QY. 2019. The antimicrobial peptides and their potential clinical applications. Am J Transl Res. 11:3919–3931.
44. LI A, ZHANG J, ZHOU Z. 2014a. PLEK: A tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. BMC Bioinformatics. 15:311. https://doi.org/10.1186/1471-2105-15-311.
45. LI J, MA W, ZENG P, WANG J, GENG B, YANG J, CUI Q. 2015. LncTar: a tool for predicting the RNA targets of long noncoding RNAs. Brief Bioinform. 16:806–812. https://doi.org/10.1093/bib/bbu048.
46. LI JJ, BICKEL PJ, BIGGIN MD. 2014b. System wide analyses have underestimated protein abundances and the importance of transcription in mammals. PeerJ. 2:e270. https://doi.org/10.7717/peerj.270.
47. LIU KS, LI TP, TON H, MAO XD, CHEN YJ. 2018. Advances of Long Noncoding RNAs-mediated Regulation in Reproduction. Chin Med J (Engl). 131:226–234. https://doi.org/10.4103/0366-6999.222337.
48. LIVAK KJ, SCHMITTGEN TD. 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. Methods. 25:402–408. https://doi.org/10.1006/meth.2001.1262.
49. LU Q, REN S, LU M, ZHANG Y, ZHU D, ZHANG X, LI T. 2013. Computational prediction of associations between long non-coding RNAs and proteins. BMC Genomics. 14:651. https://doi.org/10.1186/1471-2164-14-651.
50. LUANGSAY S, WITTAMER V, BONDUE B, HENAU O DE, ROUGER L, BRAIT M, FRANSSEN J-D, NADAI P DE, HUAUX F, PARMENTIER M. 2009. Mouse ChemR23 Is Expressed in Dendritic Cell Subsets and Macrophages, and Mediates an Anti-Inflammatory Activity of Chemerin in a Lung Disease Model. J

Immunol. 183:6489–6499. https://doi.org/10.4049/JIMMUNOL.0901037.

51. LUO W, BROUWER C. 2013. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics. 29:1830–1831. https://doi.org/10.1093/bioinformatics/btt285.
52. MADHRA M, GAY E, FRASER HM, DUNCAN WC. 2004. Alternative splicing of the human luteal LH receptor during luteolysis and maternal recognition of pregnancy. Mol Hum Reprod. 10:599–603. https://doi.org/10.1093/MOLEHR/GAH076.
53. MATTERN A, ZELLMANN T, BECK-SICKINGER AG. 2014. Processing, signaling, and physiological function of chemerin. IUBMB Life. 66:19–26. https://doi.org/10.1002/iub.1242.
54. MCLAREN W, GIL L, HUNT SE, RIAT HS, RITCHIE GRSS, THORMANN A, FLICEK P, CUNNINGHAM F. 2016. The Ensembl Variant Effect Predictor. Genome Biol. 17:122. https://doi.org/10.1186/s13059-016-0974-4.
55. MEDER W, WENDLAND M, BUSMANN A, KUTZLEB C, SPODSBERG N, JOHN H, RICHTER R, SCHLEUDER D, MEYER M, FORSSMANN WG. 2003. Characterization of human circulating TIG2 as a ligand for the orphan receptor ChemR23. FEBS Lett. 555:495–499. https://doi.org/10.1016/S0014-5793(03)01312-7.
56. MONNIER J, LEWÉN S, O'HARA E, HUANG K, TU H, BUTCHER EC, ZABEL BA. 2012. Expression, Regulation, and Function of Atypical Chemerin Receptor CCRL2 on Endothelial Cells. J Immunol. 189:956–967. https://doi.org/10.4049/jimmunol.1102871.
57. NAGPAL S, PATEL S, JACOBE H, DISEPIO D, GHOSN C, MALHOTRA M, TENG M, DUVIC M, CHANDRARATNA RAS. 1997. Tazarotene-induced Gene 2 (TIG2), a Novel Retinoid-Responsive Gene in Skin. J Invest Dermatol. 109:91–95. https://doi.org/10.1111/1523-1747.ep12276660.
58. NAKAGAWA S, SHIMADA M, YANAKA K, MITO M, ARAI T, TAKAHASHI E, FUJITA Y, FUJIMORI T, STANDAERT L, MARINE JC, HIROSE T. 2014. The lncRNA Neat1 is required for corpus luteum formation and the establishment of pregnancy in a subpopulation of mice. Dev. 141:4618–4627. https://doi.org/10.1242/dev.110544.
59. NAWROCKI EP, EDDY SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics. 29:2933–2935. https://doi.org/10.1093/bioinformatics/btt509.
60. NG PC, HENIKOFF S. 2003. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 31:3812–3814. https://doi.org/10.1093/nar/gkg509.
61. NOBLE JD, BALMANT KM, DERVINIS C, DE LOS CAMPOS G, RESENDE MFR, KIRST M, BARBAZUK WB. 2020. The Genetic Regulation of Alternative Splicing in Populus deltoides. Front Plant Sci. 11:590. https://doi.org/10.3389/fpls.2020.00590.
62. PAROLINI S, SANTORO A, MARCENARO E, LUINI W, MASSARDI L, FACCHETTI F, COMMUNI D, PARMENTIER M, MAJORANA A, SIRONI M, TABELLINI G, MORETTA A, SOZZANI S. 2007. The role of chemerin in the colocalization of NK and dendritic cell subsets into inflamed tissues. Blood. 109:3625–3632. https://doi.org/10.1182/BLOOD-2006-08-038844.
63. PATRO R, DUGGAL G, LOVE MI, IRIZARRY RA, KINGSFORD C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 14:417–419. https://doi.org/10.1038/nmeth.4197.
64. PERTEA M, PERTEA GM, ANTONESCU CM, CHANG T-CC, MENDELL JT, SALZBERG SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 33:290–295. https://doi.org/10.1038/nbt.3122.
65. R CORE TEAM. 2021. R: A language and environment for statistical computing. https://www.R-project.org/ (02.05.2021).
66. RAMOS M, SCHIFFER L, RE A, AZHAR R, BASUNIA A, RODRIGUEZ C, CHAN T, CHAPMAN P, DAVIS SR, GOMEZ-CABRERO D, CULHANE AC, HAIBE-KAINS B, HANSEN KD, KODALI H, LOUIS MS, MER AS, RIESTER M, MORGAN M, CAREY V, WALDRON L. 2017. Software for the integration of multiomics experiments in Bioconductor. Cancer Res. 77:e39–e42. https://doi.org/10.1158/0008-5472.CAN-17-0344.
67. RAO MRS. (red.). 2017. Long Non Coding RNA Biology. Springer Singapore, Singapur, 323 ss. https://doi.org/10.1007/978-981-10-5203-3.
68. REIMAND J, ARAK T, ADLER P, KOLBERG L, REISBERG S, PETERSON H, VILO J. 2016. g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. 44:W83–W89. https://doi.org/10.1093/nar/gkw199.

69. REVERCHON M, BERTOLDO MJ, RAMÉ C, FROMENT P, DUPONT J. 2014a. CHEMERIN (RARRES2) Decreases In Vitro Granulosa Cell Steroidogenesis and Blocks Oocyte Meiotic Progression in Bovine Species. Biol Reprod. 90:102. https://doi.org/10.1095/biolreprod.113.117044.
70. REVERCHON M, CORNUAU M, RAME C, GUERIF F, ROYERE D, DUPONT J, RAM C, GUERIF F, ROYRE D, DUPONT J. 2012. Chemerin inhibits IGF-1-induced progesterone and estradiol secretion in human granulosa cells. Hum Reprod. 27:1790–1800. https://doi.org/10.1093/humrep/des089.
71. REVERCHON M, RAMÉ C, BERTOLDO M, DUPONT J. 2014b. Adipokines and the Female Reproductive Tract. Int J Endocrinol. 2014:232454. https://doi.org/10.1155/2014/232454.
72. RIBEIRO LA, BACCI ML, SEREN E, TAMANINI C, FORNI M. 2007. Characterization and differential expression of vascular endothelial growth factor isoforms and receptors in swine corpus luteum throughout estrous cycle. Mol Reprod Dev. 74:163–171. https://doi.org/10.1002/mrd.20589.
73. ROURKE JL, DRANSE HJ, SINAL CJ. 2013. Towards an integrative approach to understanding the role of chemerin in human health and disease. 14:245–262. https://doi.org/10.1111/OBR.12009.
74. RYTELEWSKA E, KIEŻUN M, ZAOBIDNA E, GUDELSKA M, KISIELEWSKA K, DOBRZYŃ K, KAMIŃSKI T, SMOLIŃSKA N. 2021. CHEMERIN as a modulator of angiogenesis and apoptosis processes in the corpus luteum of pigs: An in vitro study. Biol Reprod. 105:1002–1015. https://doi.org/10.1093/BIOLRE/IOAB126.
75. RYTELEWSKA E, KISIELEWSKA K, KIEŻUN M, DOBRZYŃ K, GUDELSKA M, RAK A, DUPONT J, KAMIŃSKA B, KAMIŃSKI T, SMOLIŃSKA N. 2020. Expression of chemerin and its receptors in the ovaries of prepubertal and mature gilts. Mol Reprod Dev. 87:739–762. https://doi.org/10.1002/mrd.23391.
76. SALDI T, CORTAZAR MA, SHERIDAN RM, BENTLEY DL. 2016. Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. J Mol Biol. 428:2623–2635. https://doi.org/10.1016/J.JMB.2016.04.017.
77. SCHWANHÄUSSER B, BUSSE D, LI N, DITTMAR G, SCHUCHHARDT J, WOLF J, CHEN W, SELBACH M. 2011. Global quantification of mammalian gene expression control. Nature. 473:337–342. https://doi.org/10.1038/nature10098.
78. SCHWEINGRUBER C, RUFENER SC, ZÜND D, YAMASHITA A, MÜHLEMANN O. 2013. Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells. Biochim Biophys Acta. 1829:612–623. https://doi.org/10.1016/j.bbagrm.2013.02.005.
79. SHANNON P, MARKIEL A, OZIER O, BALIGA NS, WANG JT, RAMAGE D, AMIN N, SCHWIKOWSKI B, IDEKER T. 2003. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res. 13:2498–2504. https://doi.org/10.1101/gr.1239303.
80. SMOLIŃSKA N, KIEŻUN M, DOBRZYŃ K, RYTELEWSKA E, KISIELEWSKA K, GUDELSKA M, ZAOBIDNA E, BOGUS-NOWAKOWSKA K, WYRĘBEK J, BORS K, KOPIJ G, KAMIŃSKA B, KAMIŃSKI T. 2019. Expression of Chemerin and Its Receptors in the Porcine Hypothalamus and Plasma Chemerin Levels during the Oestrous Cycle and Early Pregnancy. Int J Mol Sci. 20:3887. https://doi.org/10.3390/ijms20163887.
81. STAUDT AC, WENKEL S. 2011. Regulation of protein function by 'microProteins'. EMBO Rep. 12:35–42. https://doi.org/10.1038/embor.2010.196.
82. SUN J, LIN Y, WU J. 2013. Long Non-Coding RNA Expression Profiling of Mouse Testis during Postnatal Development. PLoS One. 8:e75750. https://doi.org/10.1371/journal.pone.0075750.
83. TAYLOR DH, CHU ETJ, SPEKTOR R, SOLOWAY PD. 2015. Long non-coding RNA regulation of reproduction and development. Mol Reprod Dev. 82:932–956. https://doi.org/10.1002/mrd.22581.
84. THATCHER SR, ZHOU W, LEONARD A, WANG BB, BEATTY M, ZASTROW-HAYES G, ZHAO X, BAUMGARTEN A, LI B. 2014. Genome-Wide Analysis of Alternative Splicing in Zea mays: Landscape and Genetic Regulation. Plant Cell. 26:3472–3487. https://doi.org/10.1105/tpc.114.130773.
85. TRINCADO JL, ENTIZNE JC, HYSENAJ G, SINGH B, SKALIC M, ELLIOTT DJ, EYRAS E. 2018. SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. Genome Biol. 19:40. https://doi.org/10.1186/s13059-018-1417-1.
86. WARNES GR, BOLKER B, BONEBAKKER L, GENTLEMAN R, LIAW WHA, LUMLEY T, MAECHLER M, MAGNUSSON A, MOELLER S, SCHWARTZ M, VENABLES B. 2021. Gplots: Various R Programming Tools for Plotting Data. https://www.scienceopen.com/document?vid=0e5d8e31-1fe4-492f-a3d8-8cd71b2b8ad9 (14.06.2021).

87. WANG ET, SANDBERG R, LUO S, KHREBTUKOVA I, ZHANG L, MAYR C, KINGSMORE SF, SCHROTH GP, BURGE CB. 2008. Alternative isoform regulation in human tissue transcriptomes. Nature. 456:470–476. https://doi.org/10.1038/nature07509.
88. WANG J, PAN Y, SHEN S, LIN L, XING Y. 2017. rMATS-DVR: rMATS discovery of differential variants in RNA. Bioinformatics. 33:2216–2217. https://doi.org/10.1093/bioinformatics/btx128.
89. WANG L, PARK HJ, DASARI S, WANG S, KOCHER JP, LI W. 2013a. CPAT: Coding-potential assessment tool using an alignment-free logistic regression model. Nucleic Acids Res. 41:e74. https://doi.org/10.1093/nar/gkt006.
90. WANG Q, KIM JY, XUE K, LIU JY, LEADER A, TSANG BK. 2012. Chemerin, a novel regulator of follicular steroidogenesis and its potential involvement in polycystic ovarian syndrome. Endocrinology. 153:5600–5611. https://doi.org/10.1210/en.2012-1424.
91. WANG Q, LEADER A, TSANG BK. 2013b. Inhibitory roles of prohibitin and chemerin in FSH-induced rat granulosa cell steroidogenesis. Endocrinology. 154:956–967. https://doi.org/10.1210/en.2012-1836.
92. WICKHAM H. 2021. ggplot2: Elegant Graphics for Data Analysis. https://ggplot2.tidyverse.org (15.06.2021).
93. WITTAMER V, FRANSSEN J-DD, VULCANO M, MIRJOLET J-FF, LE POUL E, MIGEOTTE I, BRÉZILLON S, TYLDESLEY R, BLANPAIN C, DETHEUX M, MANTOVANI A, SOZZANI S, VASSART G, PARMENTIER M, COMMUNI D. 2003. Specific Recruitment of Antigen-presenting Cells by Chemerin, a Novel Processed Ligand from Human Inflammatory Fluids. J Exp Med. 198:977–985. https://doi.org/10.1084/jem.20030382.
94. WU J, MAO X, CAI T, LUO J, WEI L. 2006. KOBAS server: A web-based platform for automated annotation and pathway identification. Nucleic Acids Res. 34:W720–W724. https://doi.org/10.1093/nar/gkl167.
95. WUCHER V, LEGEAI F, HÉDAN B, RIZK G, LAGOUTTE L, LEEB T, JAGANNATHAN V, CADIEU E, DAVID A, LOHI H, CIRERA S, FREDHOLM M, BOTHEREL N, LEEGWATER PAJ, LE BÉGUEC C, FIETEN H, JOHNSON J, ALFÖLDI J, ANDRÉ C, LINDBLAD-TOH K, HITTE C, DERRIEN T. 2017. FEELnc: A tool for long non-coding RNA annotation and its application to the dog transcriptome. Nucleic Acids Res. 45:e57. https://doi.org/10.1093/nar/gkw1306.
96. YAO RW, WANG Y, CHEN LL. 2019. Cellular functions of long noncoding RNAs. Nat Cell Biol. 21:542–551. https://doi.org/10.1038/s41556-019-0311-8.
97. YATES AD, ACHUTHAN P, AKANNI W, ALLEN J, ALLEN J, ALVAREZ-JARRETA J, AMODE MR, ARMEAN IM, AZOV AG, BENNETT R, BHAI J, BILLIS K, BODDU S, MARUGÁN JC, CUMMINS C, DAVIDSON C, DODIYA K, FATIMA R, GALL A, GIRON CG i in. 2020. Ensembl 2020. Nucleic Acids Res. 48:D682–D688. https://doi.org/10.1093/nar/gkz966.
98. ZABEL BA, ALLEN SJ, KULIG P, ALLEN JA, CICHY J, HANDEL TM, BUTCHER EC. 2005. Chemerin Activation by Serine Proteases of the Coagulation, Fibrinolytic, and Inflammatory Cascades. J Biol Chem. 280:34661–34666. https://doi.org/10.1074/JBC.M504868200.
99. ZABEL BA, NAKAE S, ZÚÑIGA L, KIM J-YY, OHYAMA T, ALT C, PAN J, SUTO H, SOLER D, ALLEN SJ, HANDEL TM, SONG CH, GALLI SJ, BUTCHER EC, ZÚÑIGA L, KIM J-YY, OHYAMA T, ALT C, PAN J, SUTO H, SOLER D, ALLEN SJ, HANDEL TM, SONG CH, GALLI SJ, BUTCHER EC. 2008. Mast cell–expressed orphan receptor CCRL2 binds chemerin and is required for optimal induction of IgE-mediated passive cutaneous anaphylaxis. J Exp Med. 205:2207–2220. https://doi.org/10.1084/jem.20080300.
100. ZABEL BA, KWITNIEWSKI M, BANAS M, ZABIEGLO K, MURZYN K, CICHY J. 2014. Chemerin regulation and role in host defense. Am J Clin Exp Immunol. 3:1–19.