Uniwersytet Warmińsko-Mazurski w Olsztynie
Wydział Biologii i Biotechnologii

Kamil Szandar

Analiza efektywności genomów organellowych w badaniach zmienności i struktury genetycznej populacji rzadkiego i zagrożonego gatunku *Pulsatilla patens* (L.) Mill.

Praca doktorska wykonana w Katedrze Botaniki i Ochrony Przyrody, Wydziału Biologii i Biotechnologii UWM w Olsztynie

Promotor: dr hab. Monika Szczecińska, prof. UWM

Promotor pomocniczy: dr Katarzyna Krawczyk

Olsztyn 2023

Badania współfinansowane przez Narodowe Centrum Nauki w ramach konkursu MINIATURA 3, DEC-2019/03/X/NZ9/01374, pt. „Genom mitochondrialny jako źródło wiedzy o strukturze genetycznej populacji *Pulsatilla patens* (L.) Mill.”

Praca doktorska realizowana w ramach projektu interdyscyplinarnych studiów doktoranckich w zakresie biologii i biotechnologii dofinansowanych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój 2014-2020, nr umowy: POWR.03.05.00-00-Z310/17-00, Program Rozwojowy Uniwersytetu Warmińsko-Mazurskiego w Olsztynie.

Spis treści

Streszczenie 2
Summary 4
1. Spis publikacji składających się na rozprawę doktorską 6
2. Wstęp 7
3. Materiał i metody 12
3.1. Materiał roślinny 12
3.2. Izolacja DNA, RNA i sekwencjonowanie 13
3.2.1. Sekwencjonowanie DNA i RNA metodą SBS (Illumina) 14
3.2.2. Sekwencjonowanie długich fragmentów (PacBio i ONT) 14
3.3. Składanie i adnotowanie genomów organellowych 14
3.4. Analizy filogenetyczne 15
3.5. Wykrywanie zmienności 16
4. Wyniki 16
4.1. Struktura i zawartość genów mitogenomu *Pulsatilla patens* 16
4.2. Sekwencje pochodzenia plastydowego oraz edycja RNA w mitogenomie *P. patens* 17
4.3. Analiza zmienności genów kodujących białka 18
4.4. Wytypowanie regionów genomu mitochondrialnego do badań filogeograficznych 18
4.5. Charakterystyka zmienności genomów chloroplastowych 19
4.6. Charakterystyka zmienności genomów mitochondrialnych 20
4.7. Analiza filogenetyczna 21
5. Podsumowanie 21
6. Literatura 23

Streszczenie

Sasanka otwarta (*Pulsatilla patens* (L.) Mill) obecnie stanowi rzadki element flory europejskiej, ale jeszcze na początku ubiegłego stulecia była gatunkiem pospolicie występującym w Europie środkowo-wschodniej, w tym w Polsce. Jednak w ciągu ostatniego 50-lecia roślina ta drastycznie ograniczyła swoje zasoby populacyjne na tym obszarze i zyskała miano gatunku zagrożonego lub krytycznie zagrożonego, chronionego Konwencją Berneńską, umieszczonego w załączniku II i IV Dyrektywy Siedliskowej. Wostatnich latach została wpisana również na listę gatunków rzadkich i zagrożonych Międzynarodowej Unii Ochrony Przyrody (IUCN). Dość szybkie tempo zanikania populacji *P. patens* spowodowało, że roślina ta stała się ciekawym obiektem badawczym, szczególnie pod względem badania procesów ewolucyjnych jakim podlegają małe izolowane geograficznie populacje jak również badania czynników które mogą ograniczać zasoby populacyjne *P. patens*.

Niniejsza rozprawa doktorska stanowi kolejny etap badań genetycznych sasanki otwartej, które mogą poszerzyć wiedzę o zasobach genomowych tej rośliny w oparciu o genomy organellowe. Do tej pory genomy organellowe wykorzystywano najczęściej do określenia różnic pomiędzy gatunkami, uwzględniając różnice w sekwencjach nukleotydowych poszczególnych genów oraz ich obecności i kolejności ułożenia w genomie. Natomiast informacje o zmienności wewnątrzgatunkowej i strukturze genetycznej populacji bazujące jednocześnie na danych z obu genomów organellowych nie były przedmiotem szerszych analiz. Dlatego też głównym celem podjętych badań była ***analiza efektywności genomów organellowych w badaniach zmienności i struktury genetycznej populacji rzadkiego i zagrożonego gatunku Pulsatilla patens* (L.) Mill.** Postawiony cel zrealizowano poprzez analizy porównawcze genomów chloroplastowych i mitochondrialnych, które obejmowały zarówno nowo zsekwencjonowane genomy jak i dostępne w bazach danych. Badania przeprowadzono na 16 osobnikach z 5 populacji, które wcześniej uwzględniono w analizach genetycznych z wykorzystaniem markerów SSR i ISJ.

Jednym z zadań badawczych zrealizowanych w ramach niniejszej rozprawy doktorskiej było zsekwencjonowanie i opisanie genomu mitochondrialnego *P. patens*. Uzyskane wyniki zaprzeczyły hipotezie występowania jednego głównego chromosomu kolistego („master circle chromosome”). Mitogenom *P. patens* wykazał budowę

wielochromosomową, na którą złożyły się trzy liniowe chromosomy flankowane przez odwrócone sekwencje powtórzone, o łącznej długości 986 613 pz. Ze względu na duże rozmiary mitogenomu do badań porównawczych mających na celu opisanie zmienności na poziomie wewnątrzgatunkowym wytypowano sekwencje genomu mitochondrialnego. Jako sekwencje genomu mitochondrialnego o dużym potencjale w badaniach filogenetycznych i filogeograficznych, które mogą stanowić alternatywne źródło informacji genetycznej względem genomu plastydowego, uznano 14 sekwencji mitochondrialnych o długości powyżej 10 kpz, które nie zwierają regionów pochodzenia plastydowego oraz długich powtórzeń. Na chromosomie pierwszym (chMt1) zidentyfikowano osiem takich sekwencji, na drugim (chMt2) cztery, a na trzecim (chMt3) dwie.

Kolejnym zadaniem badawczym były analizy porównawcze, których celem było opisanie zmienności zarówno na poziomie wewnątrz jak i międzypopulacyjnym. Uzyskane wyniki wykazały, że 14 wytypowanych sekwencji genomu mitochondrialnego wykazało znacznie niższy poziom zmienności w porównaniu do genomu plastydowego. Genomy chloroplastowe pomimo tego, że były bardziej zmienne zidentyfikowały tę samą liczbę haplotypów (11) co regiony genomu mitochondrialnego. Należy jednak podkreślić, że zaobserwowano różnice, gdyż zarówno genomy mitochondrialne jak i chloroplastowe umożliwiły molekularną identyfikację osobników, które były nierozpoznawalne na podstawie jednych lub drugich genomów.

Przeprowadzone badania wykazały, że zastosowanie w badaniach z zakresu genomiki porównawczej danych o genomach organellowych ma ogromny potencjał i rzuca nowe światło na obraz zmienności genetycznej gatunku, i może stanowić cenne uzupełnienie wiedzy o zasobach genomowych gatunków.

Summary

Pasque flower (*Pulsatilla patens* (L.) Mill) is currently a rare and endangered species of European flora, however in the beginning of the last century it was rather common across middle east Europe, including Poland. In the last fifty years *P. patens* dramatically decreased its population size in this area and became an endangered or critically endangered species, protected under Annex I of the Bern Convention, and listed in Annex II and Annex IV of the Habitats Directive of the European Union. Recently, *P. patens* has also been included in the list of rare and endangered species of the International Union for Conservation of Nature (IUCN). Rapidly declining populations of pasque flower became a major concern of scientists, which arouse new research topics involving, in particular, evolutionary processes in small, geographically isolated populations as well as factors limiting the size of *P. patens* populations.

Presented doctoral thesis is another stage of genetic research on *P. patens*, aimed to increase the knowledge on its genome resources based on the analyses of organellar genomes. To date, organellar genomes were mostly used to determine differences between species, regarding nucleotide sequences of particular genes as well as their presence and organization in the genome. On the other hand, intra-species diversity and genetic structure of populations based on the combined data from both organellar genomes were not extensively studied. Therefore, the main aim of the presented thesis was **the analysis of organellar genomes effectiveness in studies on variability and genetic structure of populations of the rare and endangered *Pulsatilla patens* (L.) Mill species.** This goal was achieved through comparative analyses of chloroplast and mitochondrial genomes, which included both newly sequenced genomes and those available in databases. The study was conducted on 16 individuals from 5 populations that were previously included in genetic analyses, using SSR and ISJ markers.

One of the research tasks carried out as part of this doctoral dissertation was the sequencing and characterization of the *P. patens* mitochondrial genome. The obtained results contradicted the hypothesis of the existence of one master circle chromosome. *P. patents* mitogenome revealed a multichromosomal structure, consisting of three linear chromosomes flanked by inverted repeated sequences, with a total length of 986,613 bp. Due to the size of mitogenome, for comparative analyses describing intra-species variability, specific mitogenome sequences of great potential in phylogenetic and

phylogeographic studies were selected. These 14 mitochondrial sequences with a length of more than 10 kb, which do not contain plastid-derived regions and long repeats, may be an alternative source of genetic information to the plastid genome. Eight such sequences were identified on the first chromosome (chMt1), four on the second (chMt2) and two on the third chromosome (chMt3).

The aim of comparative analyses, an another research task, was to describe the variability both at the intra- and inter-population level. The obtained results showed that the 14 selected mitogenome sequences demonstrated distinctively lower level of variability compared to the plastid genome. Moreover, chloroplast genomes, although more variable, identified the same number of haplotypes (11) as regions of the mitochondrial genome. It should be emphasized that combining the data from both the mitochondrial and chloroplast genomes enabled the molecular identification of individuals, which would be unrecognizable on the basis of either genome independently.

The conducted research has shown that the utility of organellar genomes data in comparative genomics studies has a great potential and sheds new light on the field of genetic variability of species. It certainly consist a valuable supplement to the knowledge about the genomic resources of species.

1. Spis publikacji składających się na rozprawę doktorską

1. Szandar K., Krawczyk K., Myszczyński K. Ślipiko M., Sawicki J., Szczecińska M. Breaking the limits - multichromosomal structure of an early eudicot *Pulsatilla patens* mitogenome reveals extensive RNA-editing, longest repeats and chloroplast derived regions among sequenced land plant mitogenomes. BMC Plant Biol 2022; 22, 109. https://doi.org/10.1186/s12870-022-03492-1

   (IF[1] = 5,260, IF[2] = 5,761; pkt.[3] = 140)

2. Szandar K., Sawicki J., Paukszto Ł., Krawczyk K., Szczecińska M. Are the Organellar Genomes Useful for Fine Scale Population Structure Analysis of Endangered Plants?—A Case Study of *Pulsatilla patens* (L.) Mill. Genes. 2023; 14(1):67. https://doi.org/10.3390/genes14010067
   (IF[1] = 4,141, IF[2] = 4,474; pkt.[3] = 100)

Razem:

Liczba punktów MNiSW: **240** pkt.

Sumaryczny IF[1]: **9,401**

Sumaryczny IF[2]: **10,235**

[1]IF – wskaźnik Impact Factor czasopisma z 2021 roku.

[2]IF – pięcioletni wskaźnik Impact Factor czasopisma z 2021 roku

MNiSW – Liczba punktów wg Komunikatu Ministra Edukacji i Nauki z dnia 21 grudnia 2021 r. o zmianie i sprostowaniu komunikatu w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych.

2. Wstęp

Postępujące w ostatnich dziesięcioleciach zmiany klimatyczne i degradacja środowiska naturalnego doprowadziły do ograniczenia zasięgów geograficznych wielu gatunków roślin. Przypuszcza się, że obecnie rzadkie i zagrożone gatunki roślin stanowią około 40% flory światowej (NIC LUGHADHA i in. 2020) Problem ich wymierania absorbuje przyrodników i naukowców na całym świecie, a zachowanie ich we florze wielu krajów stało się celem, nad którego osiągnięciem pracują całe zespoły badawcze. W badaniach tych poza danymi ekologiczno-siedliskowymi obecnie coraz częściej uwzględnia się dane dotyczące ich zasobów genowych i genomowych. Pozwala to na oszacowanie poziomu zmienności i struktury genetycznej populacji, i ułatwia wskazanie do ochrony populacji najcenniejszych.

Zmienność genetyczna jest ważnym atrybutem każdej populacji, a w przypadku gatunków zagrożonych odgrywa istotną rolę w podejmowaniu działań ochronnych, ponieważ wpływa na potencjał ewolucyjny gatunku i jego możliwości adaptacyjne do zmieniających się często stochastycznych warunków środowiska (REED i FRANKHAM 2003, JOHANSSON i in. 2007, JUMP i in. 2009, KIRK i FREELAND 2011). Dane o zasobach genowych zrewolucjonizowały badania zakresu genetyki ochrony gatunków i pozwoliły zweryfikować strategie dotyczące aktywnej ochrony wielu rzadkich i zagrożonych gatunków (HEDRICK i MILLER 1992). W większości badania te potwierdziły, że zmienność genetyczna gatunków rzadkich i zagrożonych utrzymuje się na niskim poziomie (FRANKHAM 1996, LEIMU 2006). Dodatkowo w populacjach tych zmiany wielkości zasięgu zachodzą dynamiczniej, a zróżnicowanie genetyczne tracone jest relatywnie szybciej w wyniku działającego dryfu genetycznego i potęgowane przez ograniczony przepływ genów (ECKERT i in. 2008). W krótkim czasie może to prowadzić do depresji wsobnej i zwiększenia poziomu homozygotyczności, co może ograniczać ich możliwości adaptacyjne (ELLSTARND i ELLAM 1993, Reed i FRANKHAM 2003, JOHANSON i in. 2007). Powyższe procesy działając jednocześnie potęgują tzw. wir wymierania (ang. extinction vortex).

Do tej pory w badaniach mających na celu oszacowanie zmienności genetycznej gatunków uznawanych za zagrożone wykorzystywano głównie sekwencje genomów jądrowych, w tym bardzo popularne mikrosatelity (WILSON i PROVAN 2003, WILLIAMS

i in. 2003, ORTEGO i in. 2015, SZCZECIŃSKA i in. 2016, FENG i in. 2016, HATMAKER i in. 2018, KLICHOWSKA i in. 2018, Clark i in. 2019). Jednak dzięki rozwojowi i dostępności technik molekularnych, w tym techniki sekwencjonowania DNA nowej generacji, możliwe stało się uwzględnienie na szerszą skalę danych dotyczących zmienności i struktury cały genomów organellowych.

Genomy organellowe: plastomy i mitogenomy są kluczowymi nośnikami informacji genetycznej, które występują w komórce w wielu kopiach. Pierwotnie genomy te zawierały dużo więcej genów, w toku ewolucji zaś ich znaczna część została wyeliminowana, niemniej jednak w ich strukturach pozostały najbardziej istotne geny. Jako pierwsze dla roślin zostały zsekwencjonowane i opisane genomy chloroplastowe, które w porównaniu z mitogenomami odznaczają się stosunkowo niewielkimi rozmiarami i charakteryzują się występowaniem 4 regionów: LSC, SSC i dwóch regionów IR. Ich średnia wielkość u roślin lądowych szacowana jest na około 151 kpz i u większości gatunków zawiera się między 130-170 kpz (GUO i in. 2021). Badania genomów chloroplastowych udowodniły ich zmienność na poziomie gatunku, a ich polimorficzne regiony międzygenowe są obecnie powszechnie wykorzystywane w badaniach z zakresu filogenetyki, filogeografii i ewolucji molekularnej (TABERLET i in. 1991, DEMESURE i in. 1995, 1996, NEWTON i in. 1999, PALME 2002, JOHNSON i in. 2008, HARSONO i in. 2017). Natomiast mitogenomy roślinne są dobrze znane ze swoich bardzo dużych rozmiarów i posiadają wiele cech unikalnych, które odróżniają je od mitogenomów zwierząt. Różnice w wielkości roślinnego mitogenomu i jego zmienności strukturalnej związane są głównie z akumulacją sekwencji niekodujących, zwłaszcza długich sekwencji powtarzalnych, oraz obecnością sekwencji plastydowych i jądrowych (MARIENFELD i in. 1999, DUMINIL i BESNARD 2021). W związku z trudnościami w uzyskaniu ich kompletnych sekwencji, mitogenomy roślin nie były do tej pory przedmiotem szerszych analiz, z wyjątkiem mszaków i gatunków modelowych (MYSZCZYŃSKI 2017, 2018, 2019, ASAF i in. 2016, VARRÉ i in. 2019). Jednak niedawny rozwój nowych technologii w tym sekwencjonowania długiego odczytu (ONT, PacBio) umożliwił przyspieszenie składania mitogenomów roślinnych i ujawnił, że koliste, jednochromosomowe mitogenomy nie dominują w komórkach roślin naczyniowych [KOZIK i in. 2019]. Ponadto wykazują unikalne cechy zarówno na poziome wewnątrz, jak i międzygatunkowym, co czyni je jednymi z najciekawszych, ale zarazem najtrudniejszych genomów do badań.

Do tej pory genomy organellowe wykorzystywano najczęściej do określenia różnic pomiędzy gatunkami, uwzględniając różnice w sekwencjach nukleotydowych poszczególnych genów oraz ich obecności i kolejności ułożenia w genomie. Natomiast informacje o zmienności wewnątrzgatunkowej i strukturze genetycznej populacji bazujące jednocześnie na danych z obu genomów organellowych mają charakter szczątkowy.

W związku z powyższym w niniejszej pracy doktorskiej wykorzystując najnowocześniejsze techniki biologii molekularnej podjęto badania z zakresu genomiki porównawczej mitogenomów i plastomów, które przeprowadzono dla sasanki otwartej (*Pulsatilla patens* (L.) Mill.) Wybór gatunku do badań nie jest przypadkowy. Roślina ta stanowi rzadki element flory Europejskiej i jest jednym z najbardziej zagrożonych gatunków roślin w Europie. Obecny zasięg tego gatunku jest rozdrobniony, a wielkość populacji w ostatnim 50-leciu uległa drastycznemu zmniejszeniu.

Sasanka otwarta jest gatunkiem politypowym, zróżnicowanym geograficznie i obejmującym kilka taksonów sklasyfikowanych pod tą nazwą w randze podgatunków, lub gatunków w obrębie podsekcji *Patentes* (AICHELE i SCHWECHLER 1957, SRAMKÓ i in. 2019). *Pulsatilla patens* sensu lato jest gatunkiem nizinnym o cyrkumpolarnym zasięgu geograficznym, występującym w Europie Środkowej (Niemcy, Polska), w strefie borealnej i stepowej rozciągającej się przez całą Eurazję i na obszarach prerii Ameryki Północnej. Natomiast takson występujący w Europie to *P. patens sensu stricto*, który zajmuje zachodnią część obszaru, i rozciąga się od Rumunii na południu do Finlandii na północy, a od Niemiec na zachodzie do Syberii na wschodzie, gdzie dalej na wschód zastępuje ją *P. flavescens* (Zucc.) i *P. multifida* (Pritzel). We wszystkich krajach Europy gatunek ten ma status krytycznie zagrożonego lub zagrożonego.

Uzyskanie informacji o wewnątrzgatunkowej zmienności genomu plastydowego i mitochondrialnego jest kolejnym etapem analiz genetycznych tego gatunku. Do tej pory przeprowadzono badania, które umożliwiły opisanie poziomu zmienności genetycznej jak i struktury genetycznej gatunku w Europejskiej części zasięgu, a także zsekwencjonowano, i opisano genom chloroplastowy *P. patens* (SZCZECIŃSKA i SAWICKI 2015).

Analiza zmienności loci mikrosatelitarnych dla 30 populacji ujawniła bardzo niski poziom zmienności genetycznej i bardzo wysoki poziom inbredu (SZCZECIŃSKA i in. 2016). Oszacowana średnia wartość współczynnika heterozygotyczności była relatywnie niska (Ho=0,05) i dla poszczególnych populacji zawierała się w przedziale od 0.000 do 0.125. Analiza struktury genetycznej potwierdziła obecność trzech grup genetycznych populacji: Baltic 1, Baltic 2 i Southern, dla których zaobserwowano różnice w częstościach alleli jak również stwierdzono obecność alleli specyficznych.

W ramach wcześniejszych badań zsekwencjonowano i opisano genom chloroplastowy sasanki otwartej, który podobnie jak u większości poznanych genomów roślinnych charakteryzował się strukturą kolistą, złożoną z dwóch podjednostek pojedynczej kopii (LSC i SSC) oraz dwóch regionów powtórzonych (IR) (SZCZECIŃSKA i SAWICKI 2015). Wykazano, że genom chloroplastowy z wyłączeniem drugiego regionu odwróconego (IR), koduje 113 genów z czego 79 to geny kodujące białka, 4 kodujące rRNA i 30 tRNA. Przeprowadzone analizy porównawcze budowy i struktury plastomu dwóch osobników, ujawniły, że największą zmiennością wewnątrzgatunkową charakteryzują się regiony międzygenowe *ndh*J-*ndh*K, *ndh*D-*ccs*A oraz intron genu *rps*16, które z powodzeniem mogą być wykorzystane w badaniach mających na celu odtworzenie historii ewolucyjnej tego gatunku.

W związku z powyższym w ramach przedstawionej pracy doktorskiej podjęto badania, które umożliwiły uzupełnienie wiedzy o zasobach genomowych gatunku w oparciu o genomy organellowe.

Wykorzystując w badaniach materiał genetyczny zgromadzony na potrzeby wcześniejszych analiz genetycznych gatunku przeprowadzono badania, których nadrzędnym celem była: ***analiza efektywności genomów organellowych w badaniach zmienności i struktury genetycznej populacji rzadkiego i zagrożonego gatunku Pulsatilla patens* (L.) Mill.**

Postawiony cel zrealizowano poprzez analizy porównawcze genomów chloroplastowych i mitochondrialnych, które obejmowały zarówno nowo zsekwencjonowane genomy jak i dostępne w bazach danych. Badania przeprowadzono na 16 osobnikach z 5 populacji, które wcześniej uwzględniono w analizach genetycznych z wykorzystaniem markerów SSR i ISJ. Umożliwiło to porównanie danych dotyczących

poziomu zmienności genetycznej dla loci mikrosateliatarnych i zmienności w obrębie analizowanych genomów organellowych. W ramach realizacji założonego celu weryfikacji poddano następujące **hipotezy badawcze:**

H1. Genom mitochondrialny *P. patens* charakteryzuje się kolistą strukturą wykazaną dla innych przedstawicieli rzędu *Ranunculales*.

H2. W genomie mitochondrialnym występują regiony pojedynczej kopii, które mogą być potencjalnymi markerami w badaniach populacyjnych.

H3: Genomy mitochondrialne na poziomie wewnątrz i międzypopulacyjnym wykazują niższy poziom zmienności niż genomy chloroplastowe.

H4: Genomy mitochondrialne zapewniają dodatkową rozdzielczość w wykrywaniu haplotypów *P. patens* i umożliwiają molekularną identyfikację osobników, które były nierozpoznawalne na podstawie genomów chloroplastowych.

H5. Plastomy i mitogenomy ujawniają podobną historię ewolucyjną gatunku.

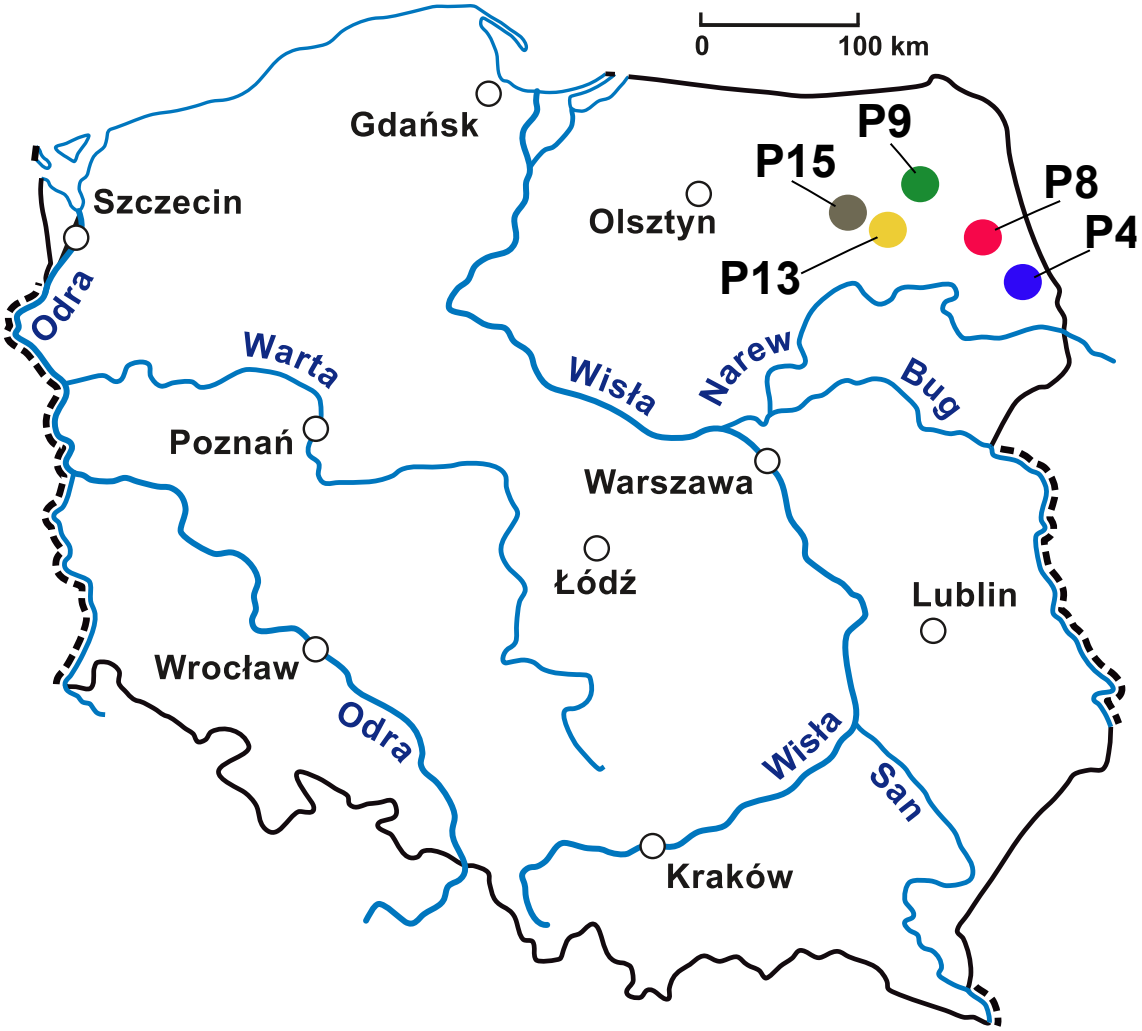
Aby zrealizować **nadrzędny cel badań** i zweryfikować postawione hipotezy badawcze wyznaczono następujące **cele szczegółowe:**

1. Zsekwencjonowanie, złożenie i opisanie genomu mitochondrialnego *Pulsatilla patens*.
2. Wytypowanie regionów genomu mitochondrialnego użytecznych w badaniach filogeograficznych sasanki otwartej.
3. Scharakteryzowanie zmienności genomów organellowych na poziomie wewnątrz i międzypopulacyjnym.
4. Sprawdzenie czy genomy mitochondrialne podobnie jak chloroplastowe mogą okazać się użytecznymi narzędziami w analizie genomiki populacji i filogeografii roślin.

3. Materiał i metody

3.1. Materiał roślinny

W badaniach wykorzystano materiał genetyczny zdeponowany w Katedrze Botaniki i Ochrony Przyrody, zebrany w latach 2009-2010 na potrzeby realizowanego projektu: NN304364438. Materiał roślinny użyty do badań obejmował łącznie 16 osobników *Pulsatilla patens*, należących do 5 populacji zlokalizowanych na terenie Polski północno-wschodniej (Ryc.1 i Tab.1). Wszystkie użyte w badaniach osobniki uwzględniono we wcześniejszych analizach genetycznych, których wyniki opublikowano w pracach dotyczących zmienności i struktury genetycznej populacji *P. patens* w zasięgu środkowoeuropejskim. (Szczecińska i in. 2016, BILSKA i SZCZECIŃSKA 2016). Wytypowane do badań osobniki należały do dwóch grup genetycznych: Baltic 1 i Baltic 2.



Ryc. 1. Rozmieszczenie geograficzne analizowanych populacji *P. patens*

Tabela 1. Charakterystyka analizowanych populacji *P. patens*

Stanowisko	Kod populacji	Kod osobnika	Współrzędne geograficzne	Grupa genetyczna
Kopna Góra (Park Krajobrazowy Puszczy Knyszyńskiej)	P4	P4-103 P4-105 P4-106	53°15′ N 23°30′ E	Baltic 1
Kopytkowo (Biebrzański Park Narodowy)	P8	P8-173 P8-197 P8-206	53°58′ N 23°56′ E	Baltic 1
Bemowo Piskie (poligon Orzysz)	P9	P9-120 P9-121 P9-122 P9-133	53°45′ N 21°59′ E	Baltic 2
Rudne (Nadleśnictwo Myszyniec)	P13	P13-1 P13-2 P13-3	53°23′ N 21°35′ E	Baltic 2
Zabiele (Nadleśnictwo Szczytno)	P15	P15-121 P15-123 P15-124	53°27′ N 21°10′ E	Baltic 1

3.2. Izolacja DNA, RNA i sekwencjonowanie

Całkowite genomowe DNA wyizolowano metodą CTAB (DOYLE i DOYLE 1990, DOYLE 1991). Ilość DNA określono fluorymetrycznie przy użyciu komercyjnego zestawu Qubit™ dsDNA BR Assay (Invitrogen, Carlsbad, NM, USA), natomiast jakość sprawdzono wykonując rozdział w 0,5% żelu agarozowym wybarwionym odczynnikiem Simple Safe (Eurx, Gdańsk, Polska). Dodatkowo, przed wykonaniem sekwencjonowania długich odczytów, wyizolowane DNA oczyszczono zestawem Genomic DNA Clean and Concentrator (Zymo, Irvine, USA).

Całkowite RNA wyizolowano, z tych samych osobników. Do izolacji RNA zastosowano zmodyfikowaną metodę fenol/SDS (AUSUBEL i in. 1987). Jakość i ilość

RNA sprawdzono bioanalizatorem (Bioanalyzer 2100, Agilent Technologies, Santa Clara, California, USA).

3.2.1. Sekwencjonowanie DNA i RNA metodą SBS (Illumina)

Do skonstruowania bibliotek genomowych wykorzystano zestaw TruSeq DNA (Illumina, San Diego, CA, USA), a następnie zsekwencjonowano na sekwenatorze HiSeqX (Illumina) przez firmę Macrogen Inc. (Seoul, Korea) uzyskując sparowane odczyty (paired-end) o długości 150 pz z insertem 350 pz. Analiza RNA-seq przeprowadzona została za pomocą zestawu Truseq RNA z chemią do redukcji rybosomalnego RNA (Ribo-Zero, Illumina), a przygotowane biblioteki zsekwencjonowano przy użyciu sekwenatora Illumina NovaSeq 6000 przez firmę Macrogen Inc. (Seoul, Korea).

3.2.2. Sekwencjonowanie długich fragmentów (PacBio i ONT)

Analizowane próby zostały przygotowane zgodnie z wytycznymi producenta do sekwencjonowania SMRTbell w systemie PacBio Sequel. Przygotowanie bibliotek i sekwencjonowanie zostało wykonane przez Macrogen Inc. (Seoul, Korea).

Biblioteki DNA do sekwencjonowania nanoporowego przygotowano za pomocą zestawu Ligation Sequencing Kit SQK-LSK109 (Oxford Nanopore Technologies), a następnie zsekwencjonowano wykorzystując komorę typu R9.4.1 umieszczoną w urządzeniu MinION MK1B (ONT). Surowe odczyty zostały wywołane przy pomocy algorytmu o wysokiej dokładności (HAC) za pomocą programu Guppy. W celu poprawy jakości odczytów sekwencjonowania długich fragmentów, odczyty nanoporowe zostały poprawione metodą hybrydowej korekcji błędów odczytu za pomocą oprogramowania FMLRC (WANG i in. 2018).

3.3. Składanie i adnotowanie genomów organellowych

Genomy chloroplastowe złożone zostały w oparciu o wcześniej opublikowany genom referencyjny (SZCZECIŃSKA i SAWICKI 2015) za pomocą programu NOVOplasty v2.8. (DIERCKXSENS i in. 2017) wyspecjalizowanego w kierunku kolistych genomów organellowych. Adnotację plastomów przeprowadzono w oparciu o istniejące genomy referencyjne w programie Geneious Prime. Rekonstrukcję genomu mitochondrialnego

przeprowadzano za pomocą szeregu programów zaimplementowanych w oprogramowaniu Geneious Prime w tym SPADes i MiniMap, wykorzystując iterowanie krótkimi i długimi odczytami. Wśród wstępnie złożonych sekwencji w oparciu o dostępne zasoby mitogenomowe roślin w bazie GenBank, zidentyfikowano kontigi zawierające geny mitochondrialne. Na wyselekcjonowanych sekwencjach przeprowadzono następnie 5-krotne mapowanie w celu ich rozszerzenia i identyfikacji sekwencji homologicznych na flankach kontigów. Obecne w genomie sekwencje powtórzone wariantowano podzbiorami składającymi się z odczytów dłuższych niż 2000 pz. Dokładna analiza odczytów nanoporowych (ONT) umożliwiła właściwą orientację trzech regionów flankujących, zmniejszając liczbę zgrupowanych sekwencji do trzech, których nie można złożyć razem ze względu na przeciwną orientację powtórzeń flankujących. Następnie, tak złożony genom zweryfikowano mapując na niego krótkie odczyty (Illumina) Mitochondrialne geny kodujące białka adnotowano przy użyciu programu MITOFY (ALVERSON i in. 2010). Zgodność z otwartymi ramkami odczytu (ORF) przewidywanymi przez program Geneious (o minimalnej długości 300 pz) sprawdzano ręcznie poprzez mapowania bibliotek RNA. Zidentyfikowane otwarte ramki odczytu (ORF) dłuższe niż 300 pz zostały sprawdzone w bazie danych GenBank przy użyciu BLASTX. Geny rRNA i tRNA zidentyfikowano odpowiednio za pomocą programu RNAmmer 1.2 (LAGESEN i in. 2007) oraz tRNAscan-SE 1.21 (LOWE i EDDY 1997). Powtórzenia zostały zidentyfikowane za pomocą wtyczki RepeatFinder do pakietu Geneious o minimalnej długości ustawionej na 500 bp i do 15% niedopasowań między powtórzeniami.

3.4. Analizy filogenetyczne

Filogenetyczna pozycja genomu mitochondrialnego *P. patens* została przeanalizowana w oparciu o trzy zestawy danych. Główny zestaw składał się z 39 genów kodujących białka (CDS), obecnych w najbardziej znanych mitogenomach roślin okrytonasiennych: *atp*1, *atp*4, *atp*6, *atp*8-9, *ccm*B, *ccm*C, *ccm*FC, *ccm*FN, *cob*, *cox*1-3, *mat*R, *mtt*B, *nad*1-7, *nad*9, *rpl*2, *rpl*5, *rpl*10, *rpl*16, *rps*1-4, *rps*8, *rps*11-14, *rps*19 i *sdh*4. Kolejne dwa zestawy danych obejmowały sekwencje genów polimeraz RNA i DNA, zidentyfikowanych w mitogenomach *P. patens* i innych gatunków roślin. Drzewa zrekonstruowano za pomocą metody najwyższej wiarygodności (ML) zaimplementowanej w programie IQ-TREE (NGUYEN i in. 2015) w oparciu o optymalny

model ewolucyjny określony za pomocą programu Modeltest (TAVARE 1986, RODRÍGUEZ i in. 1990, GU i in. 1995, POSADA i CRANDALL 1998).

Analizę filogenetyczną haplotypów cpDNA i mtDNA na poziomie wewnątrzgatunkowym przeprowadzono za pomocą metody najbliższego sąsiada. Sieć haplotypów, drzewa filogenetyczne oraz ich korelacje obliczano przy pomocy pakietów ape, phangorn oraz phytools w środowisku R (SCHLIEP 2011, PARADIS i SCHLIEP 2019, REVELL 2012).

3.5. Wykrywanie zmienności

Analizy międzygatunkowe (*Pulsatilla alpina*, *P. patens* i *P. pratensis*) zmienności genów mitochondrialnych przeprowadzono w programie Geneious Prime (moduł SNP caller) w oparciu o wyniki mapowania, zakładając minimalne pokrycie wariantu na poziomie 20 przy częstości powyżej 0,8. Wartości zróżnicowania nukleotydowego na poziomie rodzaju obliczono w programie Tassel 5.2.6 (BRADBURY i in. 2007). Analizy zmienność wewnątrzgatunkowej przeprowadzono w oparciu o uliniowanie programem MAFFT (KATOH i STANDLEY 2013). Analizę zmian nukleotydowych wykonano w programie Geneious Prime (Biomatters, Auckland, Nowa Zelandia) przy standardowych parametrach modułu SNP. Do wizualizacji zmienności częstotliwości między populacjami, rodzaju zmian (SNP lub InDel) wykorzystano pakiet oprogramowania Circos 0.69 (KRZYWIŃSKI i in. 2009).

4. Wyniki

4.1. Struktura i zawartość genów mitogenomu *Pulsatilla patens*

Przeprowadzone badania ujawniły, że mitogenom sasanki otwartej wykazuje budowę wielochromosomową, na którą składają się trzy liniowe chromosomy flankowane przez odwrócone sekwencje powtórzone o łącznej długości 986 613 pz. Powtórzenia występujące w mitogenomie *P. patens* i kształtujące jego strukturę należą do najdłuższych, jakie dotąd zostały wykryte, i liczą od 8 kpz (powtórzenie A) do ponad 87 kpz (powtórzenie D). Adnotacja mitogenomu *Pulsatilla patens* umożliwiła identyfikację zestawu genów (rRNA, tRNA i CDS) typowych dla roślin

okrytonasiennych. Według klasyfikacji SKIPPINGTON i in. (2015) zsekwencjonowany mitogenom zawiera wszystkie geny podstawowe (*atp*1, *atp*4, *atp*6, *atp*8, *atp*9, *ccm*B, *ccm*C, *ccm*Fc, *ccm*Fn, *cob*, *cox*1-3, *mat*R, *mtt*B, *nad*1-7, *nad*9) i większość genów określonych jako fakultatywne (*rpl*1, *rpl*5, *rpl*10, *rpl*16, *rps*1-4, *rps*7, *rps*11-14, *rps*19, *sdh*4). Spośród znanych z mitogenomów roślin genów nie znaleziono jedynie dwóch: *rps*10 i *sdh*3. Identycznych zestaw genów znaleziono również u trzech pozostałych gatunków *Pulsatilla*, w tym *P. alpina*, *P. pratensis* i *P. vernalis*. Dodatkowo na chromosomie pierwszym (chMt1) mitogenomu *P. patens* zlokalizowano dwa geny kodujące polimerazy: gen Pp_DNA_pol_B o długości 2850 pz (kodujący polimerazę DNA) i Pp_RNA_pol o długości 2154 pz (kodujący polimerazę RNA zależną od DNA). Oba geny są nieobecne u trzech innych gatunków *Pulsatilla*, w tym *P. alpina*, *P. pratensis* i *P. vernalis*, jak również u bardziej odległych gatunków z rzędu *Ranunculales*, takich jak *Anemone maxima* czy *Aconitum kusnezoffii*. Ich występowanie u innych gatunków roślin może świadczyć o niedawnym transferze tych genów z genomu jądrowego do mitogenomu u *P. patens*.

4.2. Sekwencje pochodzenia plastydowego oraz edycja RNA w mitogenomie *P. patens*

Obecność transferów DNA pochodzenia plastydowego w mitogenomach roślin naczyniowych jest zjawiskiem stosunkowo powszechnym (KUBO i in. 2000, TURMEL i in. 2003). Mitogenom *P. patens* zawiera ponad 35 kpz sekwencji pochodzenia plastydowego obecnych na wszystkich trzech chromosomach, co stanowi 3,6% jego całkowitej długości. Łącznie mitogenom sasanki otwartej zawiera prawie 27% jej genomu plastydowego. Plasuje to *P. patens* pomiędzy *Cucumis melo* (22,7%), a *Bambusa oldhamii* (40,9%). Występowanie unikalnego dla sasanek układu genów plastydowych w mitogenomie, przeczy hipotezie o jednokrotnym transferze chloroplastowego DNA do mitochondrium na wczesnym etapie różnicowania się roślin okrytozalążkowych.

Analizując miejsca edycji RNA, wykazano, że transkrypty mitochondrialne *P. patens* były edytowane z dużą częstotliwością, przekraczającą znany poziom u innych analizowanych roślin okrytonasiennych. Na uwagę, zasługuje fakt, że miało to miejsce pomimo przyjęcia rygorystycznych kryteriów kwalifikujących zjawisko edycji RNA (polegających na liczeniu i uwzględnianiu zdarzeń edycyjnych o częstości wariantów większych niż 0,5) oraz analizę generalnie rzadziej edytowanego transkryptomu

liściowego, a nie kwiatowego. Ujawniono wysoką łączną liczbę miejsc edycji, która wyniosła ponad 900, a genem najczęściej edytowanym był gen *nad*4 z 65 zmianami C na U. Nie wykryto, niekanonicznej, odwrotnej edycja U na C, charakterystycznej dla roślin zarodnikowych.

4.3. Analiza zmienności genów kodujących białka

Analiza porównawcza zmienności genów kodujących białka (CDS) pomiędzy gatunkami z rodzaju *Pulsatilla* ujawniła występowanie wielu polimorfizmów pojedynczego nukleotydu (SNP). Zaledwie osiem z 38 porównywanych genów nie wykazało żadnej zmienności międzygatunkowej (*nad*1, *nad*3, *nad*6, *nad*9, *cox*3, *ccm*C, *rpl*2, *mtt*B), a w kolejnych siedmiu genach wykryto tylko pojedyncze SNP (*nad*4, *nad*4L, *nad*5, *nad*7, *cox*1, *ccm*B, *rps*12, *mat*R). Pomiędzy *P. patens*, a *P. alpina* ujawniono łącznie 50 SNPs (43 dN oraz 7 dS), a pomiędzy *P. patens*, a *P. pratensis* 23 (18 dN i 5 dS).

Skonstruowano drzewo filogenetyczne na podstawie 39 genów mitochondrialnych kodujących białka, co pozwoliło na grupowanie razem gatunków *Aconitum*, *Anemone* i *Pulsatilla* (*Ranunculales*) z *Nelumbo nucifera* (*Proteales*) jako ich najbliższym krewnym, co jest zgodne z wcześniej opublikowanymi pracami (DONG i in. 2020).

4.4. Wytypowanie regionów genomu mitochondrialnego do badań filogeograficznych

Wykorzystanie genomów mitochondrialnych roślin w analizach filogenetycznych i filogeograficznych bywa kłopotliwe ze względu na rozmiary mitogenomu, obecność długich powtórzeń oraz sekwencji pochodzenia plastydowego. Wyjściem z tej sytuacji jest analiza konserwatywnych regionów mitochondrialnych, które mogą mieć ogromny potencjał w badaniach filogenetycznych i stanowić alternatywne źródło informacji genetycznej względem genomu plastydowego. Dlatego też wytypowano 14 sekwencji mitochondrialnych o długości powyżej 10 kpz, które nie zwierają regionów pochodzenia plastydowego oraz długich powtórzeń. Na chromosomie pierwszym (chMt1) zidentyfikowano osiem takich sekwencji, na drugim (chMt2) cztery, a na trzecim (chMt3) dwie. Regiony te wykorzystano w analizach zmienności na poziomie wewnątrzgatunkowym opublikowanych w pracy nr II.

4.5. Charakterystyka zmienności genomów chloroplastowych

Przeprowadzone analizy porównawcze budowy i struktury genomów chloroplastowych 16 osobników *P. patens* należących do pięciu populacji wykazały różnice dotyczące zarówno ich długości jak liczby stwierdzonych mutacji. Długość kompletnych genomów chloroplastowych dla przeanalizowanych osobników zawierała się w przedziale od 161 836 kb dla osobnika P9-122 do 161 930 dla osobników P4-103 i P4-105. Łącznie zidentyfikowano 160 mutacji punktowych, w tym 64 substytucje (SNP) i 96 indeli. Najwięcej SNP i indeli (93,12%) stwierdzono w regionach międzygenowych i intronach, zaś tylko 6,88% zaobserwowano w genach. Najbardziej zmiennym regionem okazał się region ndhD-ccsA o długości 500 pz (20.63%) z 14 SNP i 18 InDel. Zgodnie z oczekiwaniami regiony kodujące genomu chloroplastowego ujawniły niższy polimorfizm niż regiony niekodujące. Mutacje stwierdzono łącznie w 10 z 76 genów kodujących białka, a najbardziej zmiennym genem okazał się gen ndhF z dwoma substytucjami (SNP).

Oprócz ujawnionej zmienności międzypopulacyjnej plastomu *P. patens* uzyskane wyniki wskazują również na wewnątrzpopulacyjne zróżnicowanie haplotypów. Pomimo wykorzystania w analizie niewielkiej liczby osobników (zaledwie 16), genom chloroplastowy umożliwił molekularną identyfikację 11 z nich. Uzyskany w tych badaniach obraz zmienności genetycznej nie odzwierciedlił jednak wyników wcześniejszych analiz opartych na markerach SSR. Analiza zmienności plastomów na poziomie wewnątrzpopulacyjnym największą zmienność ujawniła w populacji Bemowo Piskie (P9). W populacji tej każdy osobnik reprezentował inny haplotyp, a ponadto u osobników z tej populacji stwierdzono zarówno największe różnice w długości genomu chloroplastowego, jak i największą liczbę SNP (58) i indeli (74). Populacja ta we wcześniejszych analizach opartych na markerach SSR nie charakteryzowała się największymi wartościami parametrów zróżnicowania genetycznego i ujawniła relatywnie niski poziom heterozygotyczności (Ho=0,050). Natomiast populacja Zabiele (P15), dla której stwierdzono największy poziom zróżnicowania genetycznego (Ho=102) w tych badaniach, ujawniła 2 haplotypy różniące się obecnością 4 SNP i 4 InDels. Populacją dla której nie zaobserwowano zmienności była populacja Rudne (P13).

Przeprowadzone analizy grupowania (haplotype network) wyodrębniły dwie wyraźne haplogrupy, z haplotypem pośrednim (P13) występującym u wszystkich

badanych osobników z populacji Rudne (P13). Pierwszą grupę utworzyło pięć haplotypów cpDNA obecnych w populacjach (P4, P8 i P9), u których stwierdzono aż 78 unikalnych mutacji (53,75% wszystkich wykrytych mutacji). Druga grupa składała się z czterech haplotypów cpDNA odznaczających się znacznie większym podobieństwem (14,41% mutacji).

Przeprowadzona charakterystyka molekularna analizowanych populacji i identyfikacja 11 odrębnych haplotypów potwierdza ograniczony przepływ genów między populacjami spowodowany m.in. ograniczoną produkcją nasion i faktem, że nasiona najczęściej opadają w pobliżu osobnika matczynego, nie są natomiast transportowane przez wiatr na większe odległości. Ponadto analiza wyróżniła dwie grupy haplotypów, które nie odpowiadają ich lokalizacji geograficznej.

4.6. Charakterystyka zmienności genomów mitochondrialnych

Analiza zmienności 14 wytypowanych sekwencji genomu mitochondrialnego o długości powyżej 10 kpz wykazała znacznie niższy poziom zmienności w porównaniu do genomu plastydowego, co można wytłumaczyć m.in. niezwykle wydajnym mechanizmem naprawczym. Łącznie zidentyfikowano 24 mutacje, 8 typu SNP i 16 indeli. Najwięcej mutacji, 5 SNP i 6 indeli (45,83%), stwierdzono w regionach zlokalizowanych na chromosomie chMt1, natomiast najmniejszą liczbę, 1 SNP i 5 indeli (25%) stwierdzono w regionach chromosomu chMt2.

W porównaniu do wyników otrzymanych dla genomu chloroplastowego, w przypadku analizowanych regionów mitochondrialnego DNA najbardziej heterogeniczną genetycznie populacją okazała się populacja Kopytkowo (P8), dla której wyróżniono trzy haplotypy różniące się obecnością 11 mutacji (8 indeli i 3 SNP). Podobny poziom różnic stwierdzono dla osobników z populacji Bemowo Piskie (8 indeli i 2 SNP), dla której wyróżniono trzy haplotypy, o jeden mniej w porównaniu z danymi chloroplastowymi. Natomiast najmniej zróżnicowaną populacją okazała się populacja Rudne (P13) z jedną specyficzną mutacją.

Tak samo jak w przypadku chloroplastowego DNA, analiza sieci haplotypów wyróżniła 11 haplotypów, które nie odpowiadały ich lokalizacji geograficznej. Uzyskane wyniki dotyczące molekularnej identyfikacji osobników nie pokrywają się z danymi jakie

uzyskano dla kompletnych genomów chloroplastowych. Analiza nie wyodrębniła wyraźnych haplogrup, tak jak w przypadku chloroplastowego DNA, ale potwierdziła, że najbardziej heterogeniczną populacją jest populacja Bemowo Piskie, dla której również wyodrębniono cztery haplotypy. Taką samą liczbę haplotypów, jak w przypadku danych chloroplastowych, stwierdzono również dla populacji P9 i P13. Natomiast dla populacji P8 i P4 dla mitogenomu zaobserwowano o jeden haplotyp więcej, niż dla plastomu.

4.7. Analiza filogenetyczna

W celu ustalenia relacji filogenetycznych badanych osobników przeprowadzono analizę filogenetyczną metodą najbliższego sąsiada (neighborhood-joining method - NJ). Jako grupę zewnętrzną w obu przypadkach użyto osobnika *P. patens* pochodzącego z populacji zlokalizowanej w zachodniej Ukrainie, występującej w zbiorowiskach stepu. Na podstawie uzyskanych wyników zaobserwowano niezgodność sygnału filogenetycznego dla danych mitochondrialnych i chloroplastowych. Ponadto podobnie jak w przypadku otrzymanych sieci haplotypów nie zaobserwowano zgodności ułożenia na drzewie filogenetycznym z rozmieszczeniem geograficznym.

5. Podsumowanie

Badania przeprowadzone w ramach mojej pracy doktorskiej pozwoliły na poszerzenie wiedzy z zakresu genomiki organellowej rzadkiego i zagrożonego gatunku jakim jest *Pulsatilla patens* i pozwoliły na zrealizowanie nadrzędnego celu badań jakim była "ocena efektywności genomów, chloroplastowego i mitochondrialnego, w badaniach zmienności i struktury genetycznej *P. patens*" W momencie rozpoczęcia badań, nie było znanych żadnych kompletnych sekwencji mitogenomów roślin z rzędu *Ranunculales*. Otrzymane wyniki, na podstawie których opisano budowę mitogenomu sasanki otwartej (opublikowane w 1 pracy), pozwoliły na weryfikację 1 i 2 hipotezy. Badania nie potwierdziły hipotezy, że **genom mitochondrialny *P. patens* charakteryzuje się kolistą strukturą.** Wykazano natomiast, budowę wielochromosomową, na którą składają się trzy liniowe chromosomy flankowane przez odwrócone sekwencje powtórzone. Analiza struktury mitogenomu wykazała występowanie unikalnego dla sasanek układu genów plastydowych w mitogenomie, co

przeczy hipotezie o jednokrotnym transferze chloroplastowego DNA do mitochondrium na wczesnym etapie różnicowania się roślin okrytozalążkowych. Jako sekwencje genomu mitochondrialnego o dużym potencjale w badaniach filogenetycznych, które mogą stanowić alternatywne źródło informacji genetycznej względem genomu plastydowego, uznano 14 sekwencji mitochondrialnych o długości powyżej 10 kpz, które nie zwierają regionów pochodzenia plastydowego oraz długich powtórzeń. Sekwencje te wytypowano i wykorzystano w drugim etapie analiz, który dotyczył genomiki porównawczej.

Badania z zakresu genomiki porównawczej z wykorzystaniem kompletnych genomów chloroplastowych i 14 regionów genomu mitochondrialnego ujawniły zmienność na poziomie wewnątrz jak i międzypopulacyjnym, i potwierdziły hipotezę 3, która zakładała, **że genom chloroplastowy w porównaniu z genomem mitochondrialnym charakteryzuje się większym poziomem zmienności**. Genomy chloroplastowe pomimo tego, że były bardziej zmienne zidentyfikowały tę samą liczbę haplotypów (11) co regiony genomu mitochondrialnego. Należy jednak podkreślić, że zaobserwowano różnice, gdyż zarówno genomy mitochondrialne jak i chloroplastowe umożliwiły molekularną identyfikację osobników, które były nierozpoznawalne na podstawie jednych lub drugich. Dla przykładu regiony genomu mitochondrialnego pozwoliły na molekularną identyfikację trzech haplotypów w populacji P8 i P4 podczas gdy na podstawie plastomu w populacjach tych zidentyfikowano po dwa haplotypy. Ponadto grupowanie zidentyfikowanych haplotypów zarówno dla danych chloroplastowych jak i mitochondrialnych nie znalazło odzwierciedlenia w ich zasięgu geograficznym.

Przeprowadzone badania wykazały, że zastosowanie w badaniach z zakresu genomiki porównawczej danych o genomach organellowych ma ogromny potencjał i rzuca nowe światło na obraz zmienności genetycznej gatunku, i może stanowić cenne uzupełnienie wiedzy o zasobach genomowych gatunków.

6. Literatura

1. AICHELE D., SCHWEGLEB H.-W. 1957. Die Taxonomie Der Gattung Pulsatilla. Repertorium novarum specierum regni vegetabilis. 60, 1–230, doi:10.1002/fedr.19570600103.
2. ALVERSON A.J., WEI X., RICE D.W., STERN D.B., BARRY K., PALMER J.D. 2010. Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of Citrullus Lanatus and Cucurbita Pepo (Cucurbitaceae). Mol Biol Evol, 27, 1436–1448, doi:10.1093/molbev/msq029.
3. ASAF S., KHAN A.L., KHAN A.R., WAQAS M., KANG S.-M., KHAN M.A., SHAHZAD R., SEO C.-W., SHIN J.-H., LEE I.-J. 2016. Mitochondrial Genome Analysis of Wild Rice (Oryza Minuta) and Its Comparison with Other Related Species. PLOS ONE, 11, e0152937, doi:10.1371/journal.pone.0152937.
4. AUSUBEL F.M. 1987. Current Protocols in Molecular Biology, Greene Pub. Associates, J. Wiley, order fulfillment: Brooklyn, N.Y.: Media, Pa, ISBN 978-0-471-62594-0.
5. BILSKA K., SZCZECIŃSKA M. 2016. Comparison of the Effectiveness of ISJ and SSR Markers and Detection of Outlier Loci in Conservation Genetics of Pulsatilla Patens Populations. PeerJ, 4, e2504, doi:10.7717/peerj.2504.
6. BRADBURY P.J., ZHANG Z., KROON D.E., CASSTEVENS T.M., RAMDOSS Y., BUCKLER E.S. 2007. TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. Bioinformatics, 23, 2633–2635, doi:10.1093/bioinformatics/btm308.
7. CLARK L.C., GAGLIANESE-WOODY M.R., ESTEP M.C. 2019. Development of 17 Microsatellite Markers in the Federally Endangered Species Liatris Helleri (Asteraceae). Appl Plant Sci, 7, e01250, doi:10.1002/aps3.1250.
8. DEMESURE B., SODZI N., PETIT R.J. 1995. A Set of Universal Primers for Amplification of Polymorphic Non-Coding Regions of Mitochondrial and Chloroplast DNA in Plants. Mol Ecol, 4, 129–131, doi:10.1111/j.1365-294x.1995.tb00201.x.
9. DEMESURE B., COMPS B., PETIT R.J. 1996. Chloroplast DNA Phylogeography of the Common Beech (Fagus Sylvatica L.) in Europe. Evolution, 50, 2515–2520, doi:10.2307/2410719.
10. DIERCKXSENS N., MARDULYN P., SMITS G. 2017. NOVOPlasty: De Novo Assembly of Organelle Genomes from Whole Genome Data. Nucleic Acids Research, 45, e18–e18, doi:10.1093/nar/gkw955.

11. DONG S., CHEN L., LIU Y., WANG Y., ZHANG S., YANG L., LANG X., ZHANG S. 2020. The Draft Mitochondrial Genome of Magnolia Biondii and Mitochondrial Phylogenomics of Angiosperms. PLOS ONE, 15, e0231020, doi:10.1371/journal.pone.0231020.
12. DOYLE J.J., DOYLE J.L. 1990. Isolation of Plant DNA from Fresh Tissue. Focus, 12, 13–15.
13. DOYLE J. 1991.DNA Protocols for Plants. In Molecular Techniques in Taxonomy. Hewitt G.M., Johnston A.W.B., Young J.P.W., Eds. NATO ASI Series. Springer: Berlin, Heidelberg, pp. 283–293 ISBN 978-3-642-83962-7.
14. DUMINIL J., BESNARD G. 2021. Utility of the Mitochondrial Genome in Plant Taxonomic Studies. In Molecular Plant Taxonomy: Methods and Protocols. Besse P., Ed. Methods in Molecular Biology. Springer US: New York, NY, pp. 107–118 ISBN 978-1-07-160997-2.
15. ECKERT C.G., SAMIS K.E., LOUGHEED S.C. 2008. Genetic Variation across Species' Geographical Ranges: The Central–Marginal Hypothesis and Beyond. Molecular Ecology, 17, 1170–1188, doi:10.1111/j.1365-294X.2007.03659.x.
16. ELLSTRAND N.C., ELAM D.R. 1993. Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. Annual Review of Ecology and Systematics, 24, 217–242, doi:10.1146/annurev.es.24.110193.001245.
17. FENG S., HE R., LU J., JIANG M., SHEN X., JIANG Y., WANG Z., WANG H. 2016. Development of SSR Markers and Assessment of Genetic Diversity in Medicinal Chrysanthemum Morifolium Cultivars. Frontiers in Genetics, 7.
18. FRANKHAM R. 1996. Relationship of Genetic Variation to Population Size in Wildlife. Conservation Biology, 10, 1500–1508, doi:10.1046/j.1523-1739.1996.10061500.x.
19. GU X., FU Y.X., LI W.H. 1995. Maximum Likelihood Estimation of the Heterogeneity of Substitution Rate among Nucleotide Sites. Molecular Biology and Evolution, 12, 546–557, doi:10.1093/oxfordjournals.molbev.a040235.
20. GUO Y.-Y., YANG J.-X., LI H.-K., ZHAO H.-S. 2021. Chloroplast Genomes of Two Species of Cypripedium: Expanded Genome Size and Proliferation of AT-Biased Repeat Sequences. Frontiers in Plant Science, 12.
21. HARSONO T., PASARIBU N., SOBIR S., FITMAWATI F., PRASETYA E. 2017. Phylogenetic Analysis of Indonesian Gandaria (Bouea) Using Molecular Markers of CpDNA TrnL-

F Intergenic Spacer. Biodiversitas Journal of Biological Diversity, 18, doi:10.13057/biodiv/d180108.

22. HATMAKER E.A., STATON M.E., DATTILO A.J., HADZIABDIC Đ., RINEHART T.A., SCHILLING E.E., TRIGIANO R.N., WADL P.A. 2018. Population Structure and Genetic Diversity Within the Endangered Species Pityopsis Ruthii (Asteraceae). Frontiers in Plant Science, 9.
23. HEDRICK P.W., MILLER P.S. 1992. Conservation Genetics: Techniques and Fundamentals. Ecological Applications, 2, 30–46, doi:10.2307/1941887.
24. JOHANSSON M., PRIMMER C.R., MERILÄ J. 2007. Does Habitat Fragmentation Reduce Fitness and Adaptability? A Case Study of the Common Frog (Rana Temporaria). Mol Ecol, 16, 2693–2700, doi:10.1111/j.1365-294X.2007.03357.x.
25. JOHNSON L.A., CHAN L.M., WEESE T.L., BUSBY L.D., MCMURRY S. 2008. Nuclear and CpDNA Sequences Combined Provide Strong Inference of Higher Phylogenetic Relationships in the Phlox Family (Polemoniaceae). Molecular Phylogenetics and Evolution, 48, 997–1012, doi:10.1016/j.ympev.2008.05.036.
26. JUMP A.S., MÁTYÁS C., PEÑUELAS J. 2009. The Altitude-for-Latitude Disparity in the Range Retractions of Woody Species. Trends in Ecology & Evolution, 24, 694–701, doi:10.1016/j.tree.2009.06.007.
27. KATOH K., STANDLEY D.M. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol, 30, 772–780, doi:10.1093/molbev/mst010.
28. KIRK H., FREELAND J.R. 2011. Applications and Implications of Neutral versus Non-Neutral Markers in Molecular Ecology. International Journal of Molecular Sciences, 12, 3966–3988, doi:10.3390/ijms12063966.
29. KLICHOWSKA E., ŚLIPIKO M., NOBIS M., SZCZECIŃSKA M. 2018. Development and Characterization of Microsatellite Markers for Endangered Species Stipa Pennata (Poaceae) and Their Usefulness in Intraspecific Delimitation. Mol Biol Rep, 45, 639–643, doi:10.1007/s11033-018-4192-x.
30. KOZIK A., ROWAN B.A., LAVELLE D., BERKE L., SCHRANZ M.E., MICHELMORE R.W., CHRISTENSEN A.C. 2019. The Alternative Reality of Plant Mitochondrial DNA: One Ring Does Not Rule Them All. PLOS Genetics, 15, e1008373, doi:10.1371/journal.pgen.1008373.

31. KRZYWINSKI M., SCHEIN J., BIROL İ., CONNORS J., GASCOYNE R., HORSMAN D., JONES S.J., MARRA M.A. 2009. Circos: An Information Aesthetic for Comparative Genomics. Genome Res, 19, 1639–1645, doi:10.1101/gr.092759.109.
32. KUBO T., NISHIZAWA S., SUGAWARA A., ITCHODA N., ESTIATI A., MIKAMI T. 2000. The Complete Nucleotide Sequence of the Mitochondrial Genome of Sugar Beet (Beta Vulgaris L.) Reveals a Novel Gene for TRNA(Cys)(GCA). Nucleic Acids Res, 28, 2571–2576, doi:10.1093/nar/28.13.2571.
33. LAGESEN K., HALLIN P., RØDLAND E.A., STAERFELDT H.-H., ROGNES T., USSERY D.W. 2007. RNAmmer: Consistent and Rapid Annotation of Ribosomal RNA Genes. Nucleic Acids Res, 35, 3100–3108, doi:10.1093/nar/gkm160.
34. LEIMU R., MUTIKAINEN P., KORICHEVA J., FISCHER M. 2006. How General Are Positive Relationships between Plant Population Size, Fitness and Genetic Variation? Journal of Ecology, 94, 942–952, doi:10.1111/j.1365-2745.2006.01150.x.
35. LOWE T.M., EDDY S.R. 1997. TRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. Nucleic Acids Res, 25, 955–964.
36. MARIENFELD J., UNSELD M., BRENNICKE A. 1999. The Mitochondrial Genome of Arabidopsis Is Composed of Both Native and Immigrant Information. Trends in Plant Science, 4, 495–502, doi:10.1016/S1360-1385(99)01502-2.
37. MYSZCZYŃSKI K., GÓRSKI P., ŚLIPIKO M., SAWICKI J. 2018. Sequencing of Organellar Genomes of Gymnomitrion Concinnatum (Jungermanniales) Revealed the First Exception in the Structure and Gene Order of Evolutionary Stable Liverworts Mitogenomes. BMC Plant Biol, 18, 321, doi:10.1186/s12870-018-1558-0.
38. MYSZCZYŃSKI K., BĄCZKIEWICZ A., BUCZKOWSKA K., ŚLIPIKO M., SZCZECIŃSKA M., SAWICKI J. 2017. The Extraordinary Variation of the Organellar Genomes of the Aneura Pinguis Revealed Advanced Cryptic Speciation of the Early Land Plants. Scientific Reports, 7, 9804, doi:10.1038/s41598-017-10434-7.
39. MYSZCZYŃSKI K., ŚLIPIKO M., SAWICKI J. 2019. Potential of Transcript Editing Across Mitogenomes of Early Land Plants Shows Novel and Familiar Trends. Int J Mol Sci, 20, doi:10.3390/ijms20122963.
40. NEWTON A.C., ALLNUTT T.R., GILLIES A.C.M., LOWE A.J., ENNOS R.A. 1999. Molecular Phylogeography, Intraspecific Variation and the Conservation of Tree Species. Trends in Ecology & Evolution, 14, 140–145, doi:10.1016/S0169-5347(98)01555-9.

41. NGUYEN L.-T., SCHMIDT H.A., VON HAESELER A., MINH B.Q. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution, 32, 268–274, doi:10.1093/molbev/msu300.
42. NIC LUGHADHA E., BACHMAN S.P., LEÃO T.C.C., FOREST F., HALLEY J.M., MOAT J., ACEDO C., BACON K.L., BREWER R.F.A., GÂTEBLÉ G., et al. 2020. Extinction Risk and Threats to Plants and Fungi. PLANTS, PEOPLE, PLANET, 2, 389–408, doi:10.1002/ppp3.10146.
43. ORTEGO J., AGUIRRE M.P., NOGUERALES V., CORDERO P.J. 2015. Consequences of Extensive Habitat Fragmentation in Landscape-Level Patterns of Genetic Diversity and Structure in the Mediterranean Esparto Grasshopper. Evol Appl, 8, 621–632, doi:10.1111/eva.12273.
44. PALMÉ A.E. 2002. Chloroplast DNA Variation, Postglacial Recolonization and Hybridization in Hazel, Corylus Avellana. Molecular Ecology, 11, 1769–1779, doi:10.1046/j.1365-294X.2002.01581.x.
45. PARADIS E., SCHLIEP K. 2019. Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in R. Bioinformatics, 35, 526–528, doi:10.1093/bioinformatics/bty633.
46. POSADA D., CRANDALL K.A. 1998. MODELTEST: Testing the Model of DNA Substitution. Bioinformatics, 14, 817–818, doi:10.1093/bioinformatics/14.9.817.
47. REED D.H., FRANKHAM R. 2003. Correlation between Fitness and Genetic Diversity. Conservation Biology, 17, 230–237, doi:10.1046/j.1523-1739.2003.01236.x.
48. REVELL L.J. 2012. Phytools: An R Package for Phylogenetic Comparative Biology (and Other Things). Methods in Ecology and Evolution, 3, 217–223, doi:10.1111/j.2041-210X.2011.00169.x.
49. RODRÍGUEZ F., OLIVER J.L., MARÍN A., MEDINA J.R. 1990. The General Stochastic Model of Nucleotide Substitution. Journal of Theoretical Biology, 142, 485–501, doi:10.1016/S0022-5193(05)80104-3.
50. SCHLIEP K.P. 2011. Phangorn: Phylogenetic Analysis in R. Bioinformatics, 27, 592–593, doi:10.1093/bioinformatics/btq706.
51. SKIPPINGTON E., BARKMAN T.J., RICE D.W., PALMER J.D. 2015. Miniaturized Mitogenome of the Parasitic Plant Viscum Scurruloideum Is Extremely Divergent and Dynamic and Has Lost All Nad Genes. Proceedings of the National Academy of Sciences of the United States of America, 112, E3515–E3524.

52. SRAMKÓ G., LACZKÓ L., VOLKOVA P.A., BATEMAN R.M., MLINAREC J. 2019. Evolutionary History of the Pasque-Flowers (Pulsatilla, Ranunculaceae): Molecular Phylogenetics, Systematics and RDNA Evolution. Molecular Phylogenetics and Evolution, 135, 45–61, doi:10.1016/j.ympev.2019.02.015.
53. SZCZECIŃSKA M., SAWICKI J. 2015. Genomic Resources of Three Pulsatilla Species Reveal Evolutionary Hotspots, Species-Specific Sites and Variable Plastid Structure in the Family Ranunculaceae. International Journal of Molecular Sciences, 16, 22258–22279, doi:10.3390/ijms160922258.
54. SZCZECIŃSKA M., SRAMKO G., WOŁOSZ K., SAWICKI J. 2016. Genetic Diversity and Population Structure of the Rare and Endangered Plant Species Pulsatilla Patens (L.) Mill in East Central Europe. PLOS ONE, 11, e0151730, doi:10.1371/journal.pone.0151730.
55. TABERLET P., GIELLY L., PAUTOU G., BOUVET J. 1991. Universal Primers for Amplification of Three Non-Coding Regions of Chloroplast DNA. Plant Mol Biol, 17, 1105–1109, doi:10.1007/BF00037152.
56. TAVARE S. 1986. Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. Lectures on Mathematics in the Life Sciences, 17, 57–86.
57. TURMEL M., OTIS C., LEMIEUX C. 2003. The Mitochondrial Genome of Chara Vulgaris: Insights into the Mitochondrial DNA Architecture of the Last Common Ancestor of Green Algae and Land Plants. The Plant Cell, 15, 1888–1903, doi:10.1105/tpc.013169.
58. VARRÉ J.-S., D'AGOSTINO N., TOUZET P., GALLINA S., TAMBURINO R., CANTARELLA C., UBRIG E., CARDI T., DROUARD L., GUALBERTO J.M., et al. 2019.Complete Sequence, Multichromosomal Architecture and Transcriptome Analysis of the Solanum Tuberosum Mitochondrial Genome. International Journal of Molecular Sciences, 20, 4788, doi:10.3390/ijms20194788.
59. WANG J.R., HOLT J., MCMILLAN L., JONES C.D. 2018. FMLRC: Hybrid Long Read Error Correction Using an FM-Index. BMC Bioinformatics, 19, 50, doi:10.1186/s12859-018-2051-3.
60. WILLIAMS B.L., BRAWN J.D., PAIGE K.N. 2003. Landscape Scale Genetic Effects of Habitat Fragmentation on a High Gene Flow Species: Speyeria Idalia (Nymphalidae). Mol Ecol, 12, 11–20, doi:10.1046/j.1365-294x.2003.01700.x.

61. WILSON P.J., PROVAN J. 2003. Effect of Habitat Fragmentation on Levels and Patterns of Genetic Diversity in Natural Populations of the Peat Moss *Polytrichum Commune*. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270, 881–886, doi:10.1098/rspb.2002.2324.