PRACE PRZEGLĄDOWE

Genomika protistów – bardzo zróżnicowanych, ale słabo poznanych eukariotów

Paweł Mackiewicz[1], Przemysław Gagat[1], Andrzej Bodył[2]

[1]Zakład Genomiki, Wydział Biotechnologii, Uniwersytet Wrocławski, Wrocław
[2]Zakład Bioróżnorodności i Taksonomii Ewolucyjnej, Instytut Zoologiczny, Uniwersytet Wrocławski, Wrocław

Genomics of Protists – very diversified but poorly studied eukaryotes

Summary

Initially, most eukaryotic sequence projects were devoted to typical animals, fungi and plants. Now, more and more effort is put into sequencing protist genomes. Protists are an artificial assemblage that usually contains unicellular eukaryotes coming from different phylogenetic lineages, much more diversified and widespread than higher Eukaryota. The sequenced protist genomes are essential for reconstruction of the Tree of Life and understanding significant events in eukaryotic evolution and diversification. Many protists are parasites and pathogens of medical and economic significance, and play an important ecological role as primary producers and crucial links in food webs. A lot of protists also serve as model organisms in many biological fields and are becoming important in biotechnology. 37 protist genome projects were published until the beginning of 2010 and 217 are ongoing. Knowledge coming from these projects will be helpful in more efficient protection from pathogenic protists and their elimination. Sequenced genomes of ecologically important protists could help to understand many environmental phenomena and even to control them. Thanks to the newly sequenced genomes we can discover unknown enzymes and metabolic pathways that will be useful in many branches of biotechnology.

Key words:
genome, genomics, unicellular eukaryote, protist, parasite, model organism.

Adres do korespondencji

Paweł Mackiewicz,
Zakład Genomiki,
Wydział Biotechnologii,
Uniwersytet Wrocławski,
ul. Przybyszewskiego 63/77,
51-148 Wrocław;
e-mail:
pamac@smorfland.uni.wroc.pl

biotechnologia
4 (91) 91–130 2010

1. Sekwencjonowanie genomów i protisty

Można przyjąć, że opublikowanie w 1995 r. pierwszej kompletnej sekwencji wolno żyjącej bakterii chrobotwórczej – *Haemophilus influenzae* (1) zapoczątkowało nowy okres w dziejach biologii nazwany erą poligenomową (2). Od tego czasu, dzięki rozwojowi nowych technik sekwencjonowania (3-5), liczba kompletnie poznanych genomów zaczęła intensywnie rosnąć, a wiele zjawisk zachodzących na różnych poziomach organizacji biologicznej zaczęto rozpatrywać w kontekście całego genomu (2). Genomy organizmów prokariotycznych są znacznie częściej sekwencjonowane niż genomy eukariotów ze względu na ich relatywnie małe rozmiary i prostą organizację. Według bazy GOLD (6, http://www.genomesonline.org), do początku 2010 r., opublikowano 1033 sekwencje genomów prokariotów i tylko 127 sekwencji przedstawicieli Eukaryota, przy czym te ostatnie genomy nie zawsze zostały w pełni zsekwencjonowane. Zgodnie z tą samą bazą danych, rozpoczętych i nie opublikowanych projektów sekwencjonowania genomów prokariotów było w tym samym okresie 3824, natomiast eukariotów – 1248.

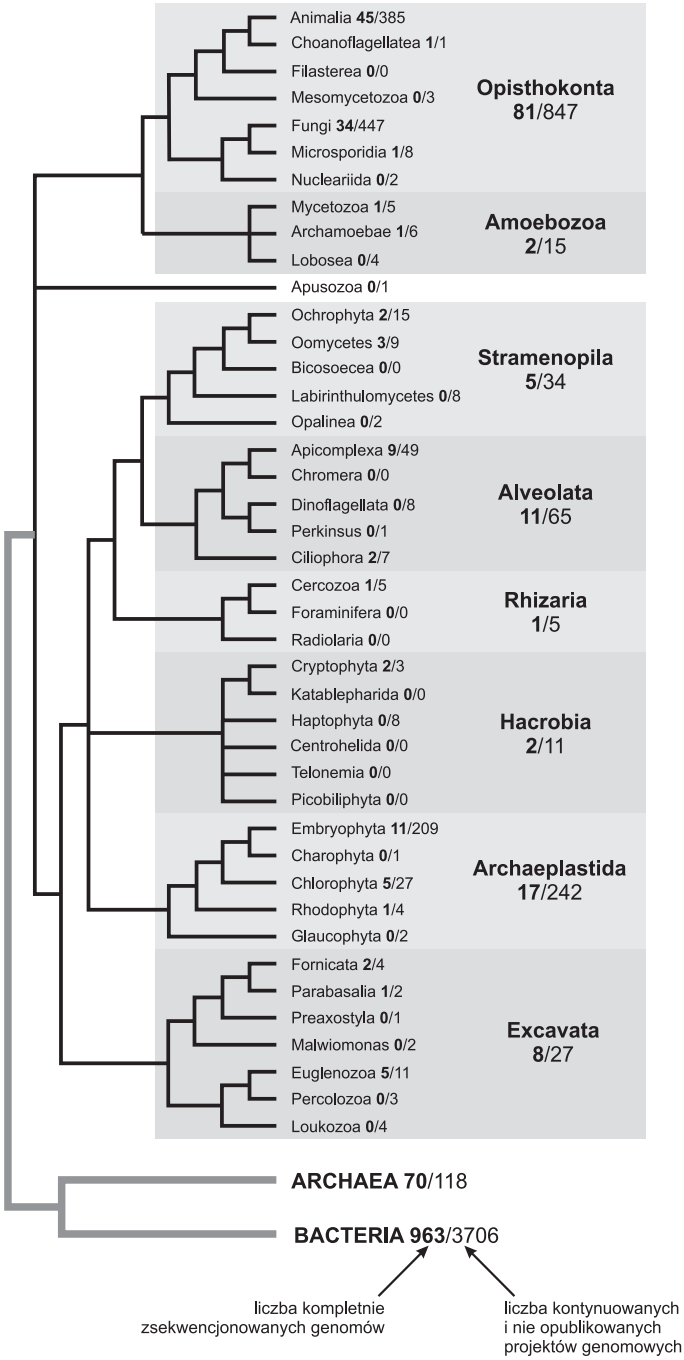
W przypadku eukariotów, większość projektów sekwencjonowania genomów była początkowo poświęcona organizmom modelowym, takim jak: drożdże *Saccharomyces cerevisiae* (7), nicień *Caenorhabditis elegans* (8), muszka owocowa *Drosophila melanogaster* (9) oraz roślina wyższa *Arabidopsis thaliana* (10). Zsekwencjonowano również genom człowieka (11,12). Wszystkie te gatunki – podobnie jak wiele innych, których genomy zsekwencjonowano później – reprezentują trzy główne grupy wyróżniane w obrębie domeny Eukaryota, tj. grzyby, rośliny i zwierzęta (13,14). Pozostałe eukarioty – niebędące typowymi grzybami, roślinami lub zwierzętami, i zaliczane przez różnych autorów do historycznej grupy Protista lub Protoctista (13-15) – nie były początkowo eksplorowane na szeroką skalę pod względem genomowym.

Podział eukariotów na cztery królestwa: protisty (Protista lub Protoctista), grzyby (Fungi), zwierzęta (Animalia) oraz rośliny (Plantae) był najbardziej popularny do początku lat osiemdziesiątych XX w. (16). Jednak w bardziej współczesnych koncepcjach postuluje się istnienie – zwłaszcza w obrębie protistów – szeregu nowych grup (17-21). Uwzględniając najnowsze odkrycia filogenetyki molekularnej (22-29), można wyodrębnić osiem głównych linii ewolucyjnych (tzw. supergrup) w obrębie Eukaryota: Opisthokonta, Amoebozoa, Archaeplastida, Hacrobia, Rhizaria, Alveolata, Stramenopila (Heterokonta) i Excavata (30-35). Do wielu z tych linii należą gatunki bardzo zróżnicowane morfologicznie, z których większość jest ciągle słabo poznana pod względem ultrastrukturalnym i molekularnym. W tabeli 1 przedstawiono krótką charakterystykę poszczególnych supergrup, a na rysunku obecne poglądy na powiązania filogenetyczne pomiędzy nimi z podaniem liczby opublikowanych kompletnych sekwencji ich genomów oraz liczby kontynuowanych i nie opublikowanych jeszcze projektów.

Tabela 1

Charakterystyka głównych linii rozwojowych eukariotów (supergrup)

Supergrupa	Formy życiowe i sposób odżywiania się	Charakterystyka	Przykłady przedstawicieli lub podgrup
Opsithokonta	wielokomórkowe (niektóre o złożonej budowie tkankowej), jednokomórkowe; heterotrofy	płaskie grzebienie mitochondrialne; jedna wić wyrastającą w dystalnej części komórki; mitochondria przekształcone w mitosomy (mikrosporidia)	zwierzęta (Animalia = Metazoa); grzyby (Fungi); choanoflagellaty; pasożyty zwierząt: *Ichthyosporea*, Microsporidia
Amoebozoa	ameby, wielojądrowa śluźnia (śluzorośla); heterotrofy	brak wici; płatowate i wachlarzowate pseudopodia; rozgałęzione rurkowate grzebienie mitochondrialne; mitochondria zmodyfikowane do mitosomów (anaeroby)	*Amoeba proteus*; śluzorośla (Mycetozoa); pasożyty: *Acanthamoeba*, *Entamoeba histolytica*
Archaeplastida (Plantae)	jednokomórkowe, kolonijne, wielokomórkowe (niektóre o złożonej budowie tkankowej); autotrofy	plastydy dwubłonowe nabyte od sinic; celulozowa ściana komórkowa; płaskie grzebienie mitochondrialne	glaukofity (Glaucophyta), krasnorosty (Rhodophyta), rośliny zielone (Viridiplantae) obejmujące zielenice i rośliny wyższe
Hacrobia	wiciowce, ameby; autotrofy, heterotrofy	pseudopodia wzmocnione włóknami osiowymi (słonecznice); plastyd czterobłonowy nabyty od krasnorostów (kryptofity, haptofity); istotny udział w planktonie	słonecznice (Centrohelida); kryptofity (Cryptophyta); haptofity (Haptophyta) – w przeszłości utworzyły kredę; Picobiliphyta – składnik pikoplanktonu
Rhizaria	ameby, wiciowce, amebowiciowce; heterotrofy, autotrofy (chlorarachniofity)	długie, cienkie pseudopodia tworzące sieć; cylindryczne grzebienie mitochondrialne; pancerzyki (promienice, otwornice); plastyd czterobłonowy nabyty od zielenic (Chlorarachniophyta)	promienice (Radiolaria) i otwornice (Foraminifera) – istotna rola skałotwórcza; pasożyty roślin (Phytomyxea) i zwierząt, szczególnie morskich bezkręgowców (Ascetosporea)
Alveolata	jednokomórkowe; heterotrofy, autotrofy (bruzdnice)	złożona powłoka komórki (amfisema) wzmocniona skomplikowanym systemem spłaszczonych pęcherzyków (alweole); rurkowate grzebienie mitochondrialne; charakterystyczna budowa wici i rzęsek; plastydy wielobłonowe różnego pochodzenia (bruzdnice); plastyd niefotosyntetyczny – apikoplast (apikompleksy)	orzęski (Ciliphora); pasożyty zwierząt (Apicomplexa), np. *Eimeria*, *Plasmodium*, *Toxoplasma*; bruzdnice (Dinoflagellata) – istotny składnik planktonu, endosymbionty koralowców rafotwórczych, powodują zakwity morskie związane z produkcją toksyn
Stramenopila (Heterokonta)	jednokomórkowe, kolonijne, wielokomórkowe pseudotkankowe (brunatnice), pseudogrzybnia komórczakowa (lęgniowce); autotrofy, heterotrofy	dwie wici nierównej długości; plastydy czterobłonowe pochodzenia krasnorostowego; pancerzyki krzemionkowe (okrzemki)	brunatnice (Phaeophyta); okrzemki (Bacillariophyta) – istotny składnik planktonu; pasożyty roślin – lęgniowce (Oomycetes): *Phytophthora*, *Pythium*; pasożyt zwierząt: *Blastocystis*
Excavata	wiciowce; heterotrofy; autotrofy (eugleniny)	charakterystyczna rynienka pokarmowa; plastyd trójbłonowy nabyty od zielenic (eugleniny); mitochondria przekształcone w mitosomy lub hydrogenosomy (anaeroby i niektóre mikroaerofile)	*Reclinomonas*; *Euglena*; pasożyty zwierząt: *Trypanosoma*, *Leishamania*, *Trichomonas*, *Gardia*, *Naegleria*



Rys. Dotychczasowe poglądy na drzewo życia ze szczególnym uwzględnieniem głównych grup eukariotów. Podano również liczby kompletnie zsekwencjonowanych genomów (czcionka pogrubiona) oraz liczby kontynuowanych i nie opublikowanych projektów (czcionka normalna) do początku 2010 r.

Ostatnio coraz więcej projektów genomowych skoncentrowanych jest na protistach (rys., tab. 2), dając coraz jaśniejszy obraz olbrzymiego zróżnicowania strukturalnego i funkcjonalnego genomów eukariotów. Protisty stanowią przeważającą część wszystkich przedstawicieli Eukaryota, dlatego poznanie ich genomów jest niezbędne do wyjaśnienia związków filogenetycznych pomiędzy różnymi liniami eukariotycznymi, a tym samym do odtworzenia drzewa życia (ang. *Tree of Life*, http:// tolweb.org/tree). Znajomość genomów protistów pozwala także zrozumieć wiele aspektów ewolucji eukariotów. Wspomnieć tu należy o: 1) zmianach wielkości oraz różnych specjalizacjach ich genomów, 2) ewolucji intronów spliceosomalnych, 3) zmianach w kodzie genetycznym, 4) poziomym transferze informacji genetycznej, 5) zmianach w szlakach metabolicznych, 6) kształtowaniu się związków pasożytniczych i symbiotycznych oraz 7) nabywaniu organelli komórkowych i ich wtórnych modyfikacjach (36--43). Wiele protistów ma olbrzymie znaczenie medyczne oraz ekologiczne. Należy tutaj wiele uciążliwych i groźnych pasożytów (np. zarodziec malaryczny *Plasmodium falciparum* wywołujący malarię) oraz organizmów fotosyntetyzujących, które odgrywają istotną rolę w łańcuchach troficznych i obiegu pierwiastków w ekosystemach naturalnych (np. okrzemki Bacillariophyta). Poznanie genomów protistów pasożytniczych może okazać się pomocne w zrozumieniu mechanizmów dokonywanych przez nie infekcji, a następnie w ich zapobieganiu oraz zwalczaniu wywołujących je pasożytów. Z kolei, zsekwencjonowanie genomów różnych glonów pozwoli zapewne na lepsze zrozumienie sieci zależności ekologicznych, a w przyszłości na kontrolę i ochronę zagrożonych ekosystemów. Ponadto genomy licznych protistów stanowią bogate źródło nieznanych jeszcze genów oraz szlaków i cykli metabolicznych, które mogą znaleźć zastosowanie w różnych gałęziach biotechnologii.

2. Opublikowane projekty sekwencjonowania genomów protistów

Według wspomnianej już bazy GOLD (6), na początku 2010 r., wśród 127 opublikowanych kompletnych sekwencji genomów eukariotów, 37 genomów reprezentowało szeroko rozumianą grupę Protista (rys., tab. 2). Stanowi to tylko 29% wszystkich projektów, chociaż protisty dominują na drzewie życia, a rośliny, zwierzęta i grzyby stanowią jedynie jego niewielką część (rys.). Największym zainteresowaniem wśród wszystkich organizmów eukariotycznych cieszyły się genomy zwierząt (45 projektów), grzybów (34 projekty) i roślin wyższych (11 projektów), natomiast wśród protistów poznano najwięcej genomów alweolatów – 11, przy czym aż 9 apikompleksów pasożytniczych.

W tabeli 2 zamieszczono wykaz i krótką charakterystykę opublikowanych sekwencji genomów protistów. Do sekwencjonowania wybierano przede wszystkim gatunki: 1) pasożytnicze o dużym znaczeniu medycznym i gospodarczym, 2) ważne ekologicznie, 3) modelowe, 4) istotne w biotechnologii oraz 5) zasadnicze w zrozumieniu ewolucji życia na naszej planecie.

Tabela 2

Wykaz opublikowanych genomów protistów

Gatunek	Supergrupa	Podgrupa	Wielkość genomu [Mb]	Liczba genów kodujących białka	Główne znaczenie	Rok opublikowania	Publikacja
1	2	3	4	5	6	7	8
Babesia bovis T2Bo	Alveolata	Apicomplexa	8,2	3671	pasożyt zwierząt	2007	(52,62)
Bigelowiella natans CCMP621 (nukleomorf)	Rhizaria	Cercozoa	0,373	331	ewolucja	2006	(108)
Chlamydomonas reinhardtii CC-503 cw92 mt+	Archaeplastida	Chlorophyta	120	15 143	modelowy/biotechnologia	2007	(102)
Cryptosporidium hominis TU502	Alveolata	Apicomplexa	9,2	3994	pasożyt człowieka	2004	(58)
Cryptosporidium parvum Iowa II	Alveolata	Apicomplexa	9,1	3807	pasożyt człowieka i zwierząt	2004	(53,59)
Cyanidioschyzon merolae 10D	Archaeplastida	Rhodophyta	16,5	4775	modelowy/biotechnologia	2004	(103,104)
Dictyostelium discoideum AX4	Amoebozoa	Mycetozoa	34	12 500	modelowy	2005	(113)
Encephalitozoon cuniculi GB-M1	Opisthokonta	Microsporidia	2,9	1997	pasożyt człowieka i zwierząt	2001	(44)
Entamoeba histolytica HM-1:IMSS	Amoebozoa	Archamoebae	20	8160	pasożyt człowieka	2005	(48,49)
Giardia intestinalis (*lamblia*) A WB ATCC 50803	Excavata	Fornicata	11,7	6470	pasożyt człowieka	2007	(46)
Giardia intestinalis (*lamblia*) B GS ATCC 50581	Excavata	Fornicata	11	6768	pasożyt człowieka	2009	(47)
Guillardia theta (nukleomorf)	Hacrobia	Cryptophyta	0,551	464	ewolucja	2001	(106)
Hemiselmis andersenii CCMP644 (nukleomorf)	Hacrobia	Cryptophyta	0,572	472	ewolucja	2007	(107)
Leishmania infantum JPCM5	Excavata	Euglenozoa	32,1	8154	pasożyt człowieka	2007	(74)
Leishmania major Friedlin	Excavata	Euglenozoa	32,8	8298	pasożyt człowieka	2005	(71,74,75)
Leishmania braziliensis MHOM/BR/75M2904	Excavata	Euglenozoa	32	8153	pasożyt człowieka	2007	(74)

1	2	3	4	5	6	7	8
Micromonas pusilla CCMP1545	Archaeplastida	Chlorophyta	21,9	10 575	ekologiczne	2009	(92)
Micromonas pusilla RCC299	Archaeplastida	Chlorophyta	20,9	10 056	ekologiczne	2009	(92)
Monosiga brevicollis MX1	Opisthokonta	Choanoflagellatea	41,6	9200	ewolucja	2008	(114)
Ostreococcus lucimarinus CCE9901	Archaeplastida	Chlorophyta	13,2	7651	ekologiczne/modelowy	2007	(90)
Ostreococcus tauri OTH95	Archaeplastida	Chlorophyta	12,6	7892	ekologiczne/modelowy	2006	(90,91)
Paramecium tetraurelia d4-2 (makronukleus)	Alveolata	Ciliophora	72	39 642	modelowy	2006	(109)
Phaeodactylum tricornutum CCAP1055/1	Stramenopila	Ochrophyta	27,4	10 402	ekologiczne/biotechnologia	2008	(93)
Phytophthora infestans T30-4	Stramenopila	Oomycetes	240	17 797	pasożyt roślin	2009	(82)
Phytophthora ramorum Pr102, UCD Pr4	Stramenopila	Oomycetes	65	14 451	pasożyt roślin	2006	(82,83)
Phytophthora sojae P6497	Stramenopila	Oomycetes	95	16 988	pasożyt roślin	2006	(82,83)
Plasmodium falciparum 3D7	Alveolata	Apicomplexa	23,3	5403	pasożyt człowieka	2002	(52-55)
Plasmodium knowlesi H, Pk1(A+)	Alveolata	Apicomplexa	23,5	5188	pasożyt człowieka i zwierząt	2008	(56)
Plasmodium vivax Sal-I	Alveolata	Apicomplexa	26,8	5433	pasożyt człowieka	2008	(55)
Plasmodium yoelii yoelii 17XNL	Alveolata	Apicomplexa	23,1	5878	pasożyt zwierząt	2002	(52,57)
Tetrahymena thermophila SB210 (makronukleus)	Alveolata	Ciliophora	103,9	27 424	modelowy/biotechnologia	2006	(110)
Thalassiosira pseudonana CCMP1335	Stramenopila	Ochrophyta	32,4	11 776	ekologiczne/biotechnologia	2004	(93,94)
Theileria annulata Ankara C9	Alveolata	Apicomplexa	8,3	3792	pasożyt zwierząt	2005	(52,60)
Theileria parva Muguga	Alveolata	Apicomplexa	8,3	4035	pasożyt zwierząt	2005	(52,60,61)
Trichomonas vaginalis G3	Excavata	Parabasalia	160	59 681	pasożyt człowieka	2007	(45)
Trypanosoma brucei TREU927/4 GUTat10.1	Excavata	Euglenozoa	26	9068	pasożyt człowieka	2005	(71,72)
Trypanosoma cruzi CL Brener TC3	Excavata	Euglenozoa	60,4	12 000	pasożyt człowieka	2005	(71,73)

2.1. Genomy protistów pasożytniczych

Protisty pasożytnicze wyewoluowały niezależnie w różnych liniach filogenetycznych, takich jak: Alveolata, Excavata, Amoebozoa, Stramenopila i Opisthokonta (33). Pomimo adaptacji do odmiennych nisz ekologicznych (np. specjalizacja do różnych żywicieli), ich genomy wykształciły na drodze konwergencji wiele cech wspólnych. Z jednej strony, uległy one redukcji i utraciły wiele genów, zwłaszcza tych związanych z metabolizmem, które okazały się zbędne w bogatym w różne związki chemiczne środowisku żywiciela. Z drugiej jednak strony, doszło w nich do ekspansji rodzin genów zaangażowanych w przystosowanie do pasożytniczego trybu życia. Dotyczy to przede wszystkim genów umożliwiających skuteczną infekcję oraz ochronę przed układem odpornościowym gospodarza. Istotne znaczenie w ich ewolucji odegrał również poziomy transfer genów od bakterii, który umożliwił przemodelowanie i wzbogacenie metabolizmu tych pasożytów. Z kolei, charakterystyczne dla nich geny są dobrymi kandydatami do wykorzystania w celach diagnostycznych, produkcji szczepionek i opracowania nowych metod leczenia.

2.1.1. Genomy pasożytów pozbawionych typowych mitochondriów: *Encephalitozoon cuniculi*, *Entamoeba histolytica*, *Trichomonas vaginalis* i *Giardia intestinalis*

Ważnymi z punktu widzenia medycznego i ekonomicznego są patogenne protisty anaerobowe i mikroaerofilne, które reprezentowane są przez: 1) mikrosporidium *Encephalitozoon cuniculi* (44), 2) ekskawaty *Trichomonas vaginalis* (45) i 3) *Giardia (lamblia) intestinalis* (46,47) oraz 4) amebę *Entamoeba histolytica* (48,49). *E. cuniculi* charakteryzuje się dużym potencjałem zoonotycznym, infekując różne gatunki ssaków, w tym człowieka. Chorują przede wszystkim osobniki o obniżonej odporności, np. dotknięte AIDS czy zażywające cyklosporynę. Główne objawy to problemy gastryczne i neurologiczne, które związane są z miejscem ekspansji tego pasożyta. *T. vaginalis* (rzęsistek pochwowy) bytuje w układzie moczowo-płciowym człowieka, powodując rocznie 170 milionów zakażeń. Ostre infekcje rzęsistkiem są często powodem wzrostu ryzyka zarażenia wirusem HIV-1. *G. intestinalis* (giardia lub lamblia) wywołuje lambliozę, infekując – najczęściej u dzieci – jelito i pęcherzyk żółciowy, co powoduje 250 milionów objawowych zakażeń rocznie. *E. histolytica* (pełzak czerwonki) jest również pasożytem układu pokarmowego, który wywołuje pełzakowicę u 50 milionów ludzi rocznie.

Na podstawie analiz genomów *E. cuniculi* (44), *T. vaginalis* (45), *G. intestinalis* (46,47) oraz *E. histolytica* (48, 49) (zob. także tab. 2) wykazano brak lub ubóstwo genów zaangażowanych w syntezę nukleotydów, kwasów tłuszczowych i aminokwasów, co wynika z pasożytniczego trybu życia i dostarczania wielu związków przez organizm gospodarza. Geny zaangażowane w biosyntezę aminokwasów i kwasów tłuszczowych występują tylko u *T. vaginalis*. U *T. vaginalis* i *G. intestinalis* nie znale-

ziono również genów kodujących miozynę, co wskazuje na istnienie osobliwego mechanizmu cytokinezy, który nie wymaga pierścienia aktyno-miozynowego.

Największej redukcji uległy genomy: *G. intestinalis* oraz *E. cuniculi*. Ten drugi genom jest jednocześnie najmniejszym (2,9 Mb) z dotychczas zsekwencjonowanych genomów organizmów eukariotycznych, wyłączając genomy nukleomorfów (zob. podrozdz. 2.3.3). Genomy *G. intestinalis* i *E. cuniculi* są zwarte pod względem struktury i składu genów. Ich geny są bardzo ubogie w introny: u *E. cuniculi* znaleziono ich 13, natomiast u *G. intestinalis* – zaledwie 4. Ponadto genom *E. cuniculi* charakteryzuje się zredukowaną długością zarówno sekwencji międzygenowych, jak i sekwencji samych białek w porównaniu do ich homologów u innych eukariotów. Olbrzymia zależność *G. intestinalis* i *E. cuniculi* od swoich gospodarzy objawia się znacznymi uproszczeniami na poziomie ich metabolizmu komórkowego. *E. cuniculi* nie jest zdolny do przeprowadzania fermentacji alkoholowej i mlekowej. Z kolei, obecność enzymów zaangażowanych w produkcję i rozkład trehalozy sugeruje, że jest to główny rezerwuar węglowodanów u tego pasożyta. Presja selekcyjna na redukcję genomów *E. cuniculi* i *G. intestinalis* doprowadziła również do uproszczenia ich maszynerii replikacyjnej i transkrypcyjnej, a także transportu wewnątrzkomórkowego. U *E. cuniculi* kolejnym uproszczeniom uległy aparat translacyjny oraz potranslacyjna modyfikacja białek, natomiast u *G. intestinalis* – zestaw genów związanych z poliadenylacją, degradacją mRNA oraz cytoszkieletem aktynowym.

Każdy z genomów omawianych pasożytów posiada swoje charakterystyczne cechy. Na przykład genom *G. intestinalis* jest tetraploidalny i występuje w postaci dwóch jąder. Dla genomu *E. histolytica* charakterystyczne są natomiast długie, tandemowo powtórzone sekwencje tRNA, które stanowią 10% wielkości genomu i pełnią najprawdopodobniej funkcje strukturalne. Z kolei, interesującą cechą genomu *T. vaginalis* jest aż 65% zawartość sekwencji powtórzonych i niedawno rozprzestrzenionych jednostek transpozonowych, co doprowadziło do wzrostu wielkości tego genomu. Niezwykle duża jest również liczba przewidzianych genów, która wynosi prawie 60 tysięcy i jest jedną z najwyższych wśród eukariotów. Jednak introny znaleziono zaledwie w 65 genach.

Duży wpływ na ukształtowanie się metabolomu *E. histolytica*, *G. intestinalis* oraz *T. vaginalis* miały wielokrotne poziome transfery genów od bakterii. W ich genomach zidentyfikowano odpowiednio 96, 100 i 152 geny o takim pochodzeniu. Wśród nich, najwięcej jest genów kodujących enzymy związane z metabolizmem węglowodanów i aminokwasów. Umożliwiają one zwiększenie zakresu przetwarzanych związków, które są łatwo dostępne w bogatym środowisku gospodarzy.

W prawie wszystkich genomach charakteryzowanych gatunków pasożytów, ale z wyjątkiem bardzo zredukowanego genomu *E. cuniculi*, doszło do niezależnych ekspansji wielu rodzin białek związanych z adaptacją do pasożytniczego trybu życia i specyficznego środowiska gospodarza. U *T. vaginalis* szczególnie liczne – jak na eukarioty jednokomórkowe – są kinazy kodowane przez około 900 genów. Białka te są również bogato reprezentowane w zwartych genomach *G. intestinalis* i *E. histolytica*. Duża liczba kinaz może świadczyć o złożonym systemie przekształcania sygnałów

u tych pasożytów. U *T. vaginalis* doszło ponadto do ekspansji różnych typów białek powierzchniowych, takich jak: BspA, GP63 i VSP, które kodowane są przez około 800 genów. U *G. intestinalis* zmienne białka powierzchniowe (VSP) również tworzą dużą rodzinę białek. U *E. histolytica* liczne są lektyny wielopodjednostkowe GalGalNAc, a ponadto znaleziono geny kodujące białka powierzchniowe typu BspA. Wszystkie te białka są istotne w adhezji i inwazji gospodarzy oraz ochronie *E. histolytica* przed ich enzymami trawiennymi i układem odpornościowym. U *E. histolytica* i *T. vaginalis* zidentyfikowano dodatkowo peptydy homologiczne, które tworzą pory w komórkach gospodarza i prowadzą ostatecznie do ich lizy. Genomy *E. histolytica* i *T. vaginalis* kodują szczególnie liczne białka związane z transportem pęcherzykowym, które są ważne w procesach, takich jak: 1) sekrecja białek patogennych, 2) fagocytoza bakterii i komórek gospodarza oraz 3) endocytoza białek i innych cząsteczek znajdujących się w otaczającym środowisku gospodarza. Do innych bogato reprezentowanych rodzin białek u *T. vaginalis*, które związane są z pasożytniczym trybem życia, należą: 1) peptydazy (potencjalne czynniki wirulencji, niektóre z nich są kandydatami na inhibitory peptydazy anty-HIV), 2) transportery węglowodanów i aminokwasów (umożliwiające pobieranie tych związków od gospodarza) oraz 3) kotransportery kationowo-chlorkowe (regulujące zmiany osmotyczne środowiska gospodarza). U *T. vaginalis* zidentyfikowano także enzymy zaangażowane w syntezę lipofosfoglikanu tworzącego glikokaliks na powierzchni komórki, który odpowiada za przyleganie tego pasożyta do komórek gospodarza. Z kolei, genom *E. histolytica* koduje: 1) liczne proteazy cysteinowe, odpowiedzialne za degradację macierzy zewnątrzkomórkowej gospodarza, 2) receptory chrakteryzujące się obecnością siedmiu obszarów transbłonowych oraz 3) trimery białek G, które regulują proces encystacji.

Wiele z wymienionych białek może służyć do celów diagnostycznych, produkcji szczepionek lub jako potencjalne cele dla leków. Dzięki sekwencjom genomów *E. cuniculi*, *E. histolytica*, *G. intestinalis* i *T. vaginalis* dokładniej poznano molekularne podłoże ich patogenezy i infekcji, co będzie pomocne w leczeniu i zapobieganiu wywoływanym przez nie chorobom.

Jedną z istotnych cech wspólnych *E. cuniculi*, *E. histolytica*, *G. intestinalis* i *T. vaginalis* jest brak typowych mitochondriów, które uległy przekształceniu w hydrogenosomy (*T. vaginalis*) lub mitosomy (pozostałe gatunki) (50,51). Hydrogenosomy uczestniczą w produkcji ATP i wodoru na drodze oddychania beztlenowego. Analizy genomu *T. vaginalis* doprowadziły do wykrycia jeszcze jednej – dotychczas nie znanej – funkcji tego organellum, która związana jest z metabolizmem aminokwasów. W mitosomach natomiast zachodzi dojrzewanie białek żelazowo-siarkowych (50,51). Przypuszcza się, że mogą one także chronić pasożyty przez stresem oksydacyjnym. Dowodem na dawną obecność typowych mitochondriów u przodków opisywanych gatunków jest występowanie w ich genomach jądrowych genów homologicznych z genami mitochondrialnymi (50,51). Zarówno hydrogenosomy, jak i mitosomy są pozbawione enzymów cyklu Krebsa oraz łańcucha transportu elektronów, a geny kodujące te białka zostały utracone przez genomy jądrowe.

Bardzo ciekawe wyniki uzyskano porównując genomy szczepów patogennych zaliczanych do jednego gatunku – *G. intestinalis* (46,47). Spośród 7 genotypów zsekwencjonowano genomy dwóch: GS i WB, które zakażają człowieka. Genomy tych szczepów wykazują odpowiednio 77 i 78% identyczności w sekwencjach kodujących białka na poziomie nukleotydowym i aminokwasowym. Genotyp GS charakteryzuje się większym poziomem polimorfizmu allelicznego (0,53%) niż genotyp WB (0,01%), przy czym oba genomy są tetraploidalne. Może to świadczyć o różnych mechanizmach tworzenia i redukcji zmienności w każdym z tych szczepów. W szczepie GS zidentyfikowano 28 unikatowych genów, natomiast w szczepie WB – tylko 3 takie geny. Mogą one zostać wykorzystane w celach diagnostycznych oraz identyfikacji poszczególnych szczepów. Niektóre z tych genów wykazują pochodzenie bakteryjne, a funkcja wielu z nich jest nieznana. Interesującą cechą, różniącą szczepy GS i WB, jest zmienny repertuar białek VSP. Dodatkowe różnice dotyczą białek: NEK i HCMP. Wszystkie te cechy mogą wyjaśniać obserwowane różnice biologiczne i kliniczne pomiędzy obu szczepami *G. intestinalis*. Ponadto promotory kilku enzymów szczepu GS, które związane są z syntezą ściany cysty, nie posiadają typowych miejsc wiązania czynników transkrypcyjnych Myb2, co może tłumaczyć słabszą encystację tego szczepu i sugerować inny mechanizm jej regulacji. Zaobserwowane różnice pomiędzy szczepami GS i WB uprawniają do traktowania ich jako odrębnych gatunków i mogą pomóc w skuteczniejszym zapobieganiu i leczeniu wywoływanych przez nie chorób.

2.1.2. Genomy apikompleksów: *Babesia, Cryptosporidium, Plasmodium, Theileria*

Wśród opublikowanych obecnie genomów protistów, najliczniejszą grupę stanowią przedstawiciele Apicomplexa (sporowce właściwe) (33). Są one endopasożytami (często wewnątrzkomórkowymi) bezkręgowców i kręgowców. Spośród znanych ponad 5000 gatunków, opublikowanych zostało zaledwie 9 genomów, należących do: 1) czterech gatunków z rodzaju *Plasmodium* (52-57), 2) dwóch gatunków z rodzaju *Cryptosporidium* (53,58,59), 3) dwóch gatunków z rodzaju *Theileria* (52,60,61) oraz 4) *Babesia bovis* (52,62) (zob. także tab. 2). Genomy te zostały zsekwencjonowane ze względu na ich olbrzymie znaczenie medyczne, gospodarcze i społeczne. Na przykład z powodu malarii, za którą odpowiadają zarodźce z rodzaju *Plasmodium*, każdego roku umiera 2,7 miliona ludzi, a 300-500 milionów choruje. Szczególnie niebezpieczny jest gatunek *P. falciparum* będący największym zabójcą ludzi wśród apikompleksów. Co 30 sekund pozbawia on życia jedno dziecko w Afryce. Pozostałe wymienione apikompleksy, odpowiedzialne są za kryptosporidiozę u ludzi (*Cryptosporidium* spp.) oraz choroby bydła: babeszjozę (*Babesia bovis*), gorączkę wschodniego wybrzeża Afryki (*Theileria parva*) i teileriozę tropikalną (*T. annulata*). Straty wyrządzane przez te gatunki szacuje się na setki milionów dolarów rocznie.

Opublikowane genomy: *Plasmodium* (52-57), *Cryptosporidium* (53,58,59), *Theileria* (52,60,61) oraz *Babesia* (52,62) są zróżnicowane pod względem ich wielkości i struk-

tury. U *Plasmodium* mają one wielkość około 25 Mb, natomiast u *Cryptosporidium*, *Theileria* i *Babesia* są prawie trzy razy mniejsze. Jednak potencjał kodujący genomu *Plasmodium* (liczącego około 5500 genów) jest zaledwie w 30% większy od zdolności kodujących: *Cryptosporidium*, *Theileria* i *Babesia*. Rozbieżności w wielkości genomu i liczbie kodowanych genów pomiędzy *Plasmodium* a *Cryptosporidium, Theileria* i *Babesia* wynikają z krótszych regionów międzygenowych i gęstszego upakowania genów u trzech ostatnich rodzajów. Najwięcej intronów posiadają genomy: *Theileria* i *Babesia*. U *Theileria* geny podzielone stanowią aż 75% całego genomu, u *Plasmodium* – ponad 50%, natomiast u *Cryptosporidium* – zaledwie 5%. Zwiększenie liczby intronów może mieć znaczenie dla tworzenia nowych wariantów genów (i kodowanych przez nie białek) na drodze tasowania eksonów, wpływając w ten sposób na dużą zmienność antygenową komórek *Theileria* i *Babesia*.

Wszystkie zsekwencjonowane genomy przedstawicieli Apicomplexa są ubogie w sekwencje kodujące tRNA (39-47 kopii) i rRNA (2-7 kopii), co pokazuje ich tendencję do redukcji. Tendencję tę, która wynika z przystosowania do pasożytniczego trybu życia, można również dostrzec w zaniku wielu enzymów zaangażowanych w podstawowe szlaki metaboliczne, takie jak: 1) glukoneogeneza, 2) utlenianie kwasów tłuszczowych, 3) cykl mocznikowy, 4) synteza aminokwasów, 5) hemu, 6) puryn i 7) kwasów tłuszczowych. Podobnie jak u innych pasożytów, genomy apikompleksów charakteryzują się zwiększoną liczbą par AT. W ekstremalnych przypadkach sięga ona aż 81% u *P. falciparum*, natomiast u *P. vivax* wynosi – 58%. Stosunkowo niska, jak na apikompleksy, zawartość par AT u *P. vivax* może być związana z izochoryczną strukturą jego chromosomów, w których tylko zmienne części telomerów są bogate w pary AT, natomiast obszary konserwatywne – o dużym potencjale kodującym – obfitują w pary GC.

Jednym z zaskakujących odkryć, będących wynikiem sekwencjonowania genomów apikompleksów, jest brak możliwości przyporządkowania aż 40% genów do odpowiednich grupy ortologicznych, co sugeruje, że kodują one nowe rodziny białek. Powstawanie nowych genów jest najprawdopodobniej przyczyną wyodrębniania się nowych gatunków w obrębie Apicomplexa. Gatunki te różnią się pomiędzy sobą niejednokrotnie zaledwie kilkoma genami, które kodują białka chroniące ich komórki przed odpowiedzią ze strony układu odpornościowego gospodarza, np. białka powierzchniowe o dużej zmienności antygenowej (PfEMP1) u *P. falciparum* czy glikoproteiny podobne do mucyny u *C. parvum*. Częsta lokalizacja subtelomerowa tych genów (szczególnie u *Plasmodium*) dodatkowo zwiększa ich zmienność antygenową na skutek częstych rekombinacji.

Sekwencjonowanie genomów przedstawicieli Apicomplexa jest szczególnie istotne, ponieważ rośnie ich oporność na dostępne leki, natomiast ich wektory (komary – w przypadku *Plasmodium*, a kleszcze – w przypadku *Theileria* i *Babesia*) również nabywają odporności na insektycydy. Informacje zawarte w genomach apikompleksów pozwoliły już na wyselekcjonowanie najlepszych kandydatów na szczepionki (np. białka adhezyjne, odpowiedzialne za łączenie się apikompleksów z komórkami gospodarza), jak również stanowią podstawę do prac nad nowymi lekami

(55,57,63,64). Źródłem nowych rozwiązań biotechnologicznych może być także poznanie nowych rodzin białek u apikompleksów oraz zrozumienie, w jaki sposób radzą one sobie z układem odpornościowym gospodarza, czego przykładem może być transformacja leukocytów przez przedstawicieli *Theileria* (65).

Zdecydowana większość apikompleksów posiada charakterystyczne organellum zwane apikoplastem (66), które jest zmodyfikowanym plastydem o pochodzeniu krasnorostowym lub zielenicowym (67,68). Apicoplast nie jest aktywny fotosyntetycznie i został zaadoptowany do pasożytniczego trybu życia, uczestnicząc w syntezie izoprenoidów, kwasów tłuszczowych i hemu (66). Jego genom koduje bardzo niewiele genów, dlatego prawie wszystkie białka apikoplastu kodowane są przez genom jądrowy. Białka te importowane są przez system wewnątrzbłonowy za pomocą charakterystycznych presekwencji dwuczęściowych, które składają się z peptydu sygnałowego i peptydu tranzytowego (69). Apikoplastu nie zidentyfikowano dotychczas u gatunków *Cryptosporidium*. Jednak w genomach jądrowych *C. hominis* oraz *C. parvum* znaleziono geny, które mogą świadczyć o jego wcześniejszej obecności (58,70).

2.1.3. Genomy świdrowców: *Trypanosoma* i *Leishmania*

Głównymi powodami sekwencjonowania 5 genomów świdrowców z rodzajów: *Trypanosoma* (71-73) i *Leishmania* (71,74,75) (zob. także tab. 2) były ważne powody medyczne i weterynaryjne. *T. brucei* wywołuje śpiączką afrykańską, *T. cruzi* – chorobę Chagasa, natomiast *L. major*, *L. infantum* i *L. braziliensis* – różne typy leiszmanioz. Łącznie *Trypanosoma* i *Leishmania* infekują miliony, a zabijają dziesiątki tysięcy ludzi. Niestety, brak szczepionek – w połączeniu z toksycznością stosowanych leków – nie daje nadziei na poprawę sytuacji w najbliższych latach.

Genomy przedstawicieli *Trypanosoma* (71-73) i *Leishmania* (71,74,75), podobnie jak genomy apikompleksów (zob. poprzedni rozdz.), charakteryzują się dużą różnorodnością strukturalną. Genom *T. brucei* składa się z 11 par dużych chromosomów o wielkości powyżej 1 Mb, natomiast genomy *T. cruzi* i *Leishmania* z odpowiednio 28 i 35-36 par mniejszych chromosomów. *T. brucei* posiada dodatkowo 3-5 średnich chromosomów (200-500 kb) i około 100 minichromosomów (50-100 kb), co w sumie daje zadziwiającą liczbę około 120 chromosomów. Bogate w chromosomy kariotypy *T. cruzi* i *Leishmania* odzwierciedlają stan ancestralny, natomiast duże chromosomy u *T. brucei* są wynikiem fuzji mniejszych chromosomów. Genom *T. cruzi* jest dwa razy większy od genomów pozostałych świdrowców. Posiada on w przybliżeniu 25-30% więcej genów niż pozostałe gatunki *Trypanosoma* oraz gatunki *Leishmania*, a co najmniej 50% jego genomu składa się z sekwencji powtórzonych, takich jak: 1) retrotranspozony, 2) powtórzenia subtelomerowe oraz 3) rodziny wielogenowe kodujące białka powierzchniowe.

Pomimo tych różnic, genomy *Trypanosoma* i *Leishmania* są podobne pod względem organizacyjnym. Występują w nich olbrzymie klastry policistronowe, składa-

jące się z dziesiątek, a nawet setek genów, pomiędzy którymi następuje przełączanie kodowania nici DNA, tzn. nić sensowa staje się antysensową, a nić antysensowa sensową (76-78). Tworzenie olbrzymich klastrów genów jest jedną z głównych przyczyn zdumiewająco wysokiego poziomu syntenii u świdrowców. Pomimo 200-500 milionów lat niezależnej ewolucji, 68% genów u *T. brucei* i 75% u *L. major* zachowało swoje konserwatywne ułożenie. Na utrzymanie syntenii wpływają też bardzo rzadkie rekombinacje mejotyczne, ponieważ rozmnażanie płciowe zdarza się u świdrowców wyjątkowo rzadko.

Najbardziej zmienne geny, które charakterystyczne są dla poszczególnych gatunków świdrowców i odpowiadają za oddziaływania z komórkami gospodarza, występują pomiędzy blokami konserwatywnych genów. U *Trypanosoma*, najliczniejsze z nich, zaangażowane są w produkcję glikoprotein powierzchniowych, umożliwiających uniknięcie reakcji ze strony układu odpornościowego gospodarza poprzez generowanie dużej zmienności antygenowej. U *T. brucei* jedna z takich glikoprotein, zwana zmienną glikoproteiną powierzchniową (VSG), kodowana jest przez 1000 sekwencji. VSG tworzy płaszcz glikoproteinowy zbudowany z 10 milionów identycznych makrocząsteczek, które osłaniają niezmienne epitopy pasożyta przed systemem immunologicznym gospodarza. Większość sekwencji kodujących VSG to pseudogeny ułożone tandemowo w częściach subtelomerowych chromosomów. Stanowią one swoisty rezerwuar potencjalnej zmienności antygenowej, który wykorzystywany jest do tworzenia aktywnych genów zlokalizowanych w specjalnych miejscach ekspresji (BESs, ang. *Bloodstream Expression Sites*). Niektóre z tych genów powstają w sposób mozaikowy z połączenia różnych genów *VSG*. Aktywacja nowego genu *VSG* polega na zastąpieniu starego genu nowym w miejscu BES. Mechanizm ten umożliwia ciągłe tworzenie nowego rodzaju VSG, co powoduje permanentny stan infekcji, który nieleczony prowadzi do śmierci. *T. cruzi* posiada znacznie więcej różnych genów kodujących białka powierzchniowe niż pozostałe świdrowce. Wśród nich należy wymienić: 1) mucyny, 2) białka powierzchniowe związane z mucynami (MASP), 3) proteazy gp63 oraz 4) bardzo zmienne transsialidazy (TS), znajdujące się pod szczególnie silną presją immunologiczną.

Na powierzchni komórek *Leishmania* występują glikoproteiny kodowane przez duże rodziny genów, takie jak: *amastin*, *gp46* i *gp63*. Jednak powierzchnia ich komórek zdominowana jest przez charakterystyczne kompleksy węglowodanowe, które uczestniczą w różnych etapach procesu infekcji. Kompleksy te zawierają: 1) lipofosfoglikan (LPG), 2) fosfolipidy glikoinozytolowe (GIPL) oraz 3) proteofosfoglikan (PPG). Obecność dwóch białek zawierających domenę czynnika hamującego migrację makrofagów (MIF), które kodowane są przez geny *LmjF33.1740* i *LmjF33.1750*, może stanowić jeden z wielu sposobów na przeżycie komórek *Leishmania* w organizmie gospodarza poprzez hamowanie odpowiedzi komórkowej ze strony jego układu odpornościowego.

Pasożytniczy tryb życia powoduje u świdrowców z jednej strony ekspansję genów kodujących makrocząsteczki oddziałujące z organizmem gospodarza, nato-

miast z drugiej prowadzi do znacznej redukcji zdolności metabolicznych. Przykładem może być utrata enzymów odpowiedzialnych za syntezę: hemu, większości aminokwasów oraz puryn. Spośród wszystkich zsekwencjonowanych genomów świdrowców, najbardziej zredukowanym metabolizmem charakteryzuje się *T. brucei*. Wynika to z jedynie zewnątrzkomórkowego trybu życia i, w konsekwencji, prostszego cyklu rozwojowego. Konieczność adaptacji do większej liczby nisz ekologicznych, zarówno zewnątrz-, jak i wewnątrzkomórkowego środowiska makrofagów, sprawia, że liczba ortologów u *T. cruzi* i *Leishmania* jest wyższa niż u *T. brucei*.

Świdrowce posiadają także uproszczoną maszynerię transkrypcyjną, a brak mechanizmów jej kontroli ujawnia się w postaci duplikacji genów jako środka na zwiększenie ich ekspresji. Duplikacje genów i ich dywergencja pozwalają również na zwiększanie zmienności. Innymi cechami charakterystycznymi świdrowców, które pozostają w bezpośrednim związku z ich genomami, są: 1) transsplicing połączony z poliadenylacją, 2) duży potencjał modyfikacji potranslacyjnych białek, 3) redukcja szlaków przekazywania sygnałów komórkowych przy jednoczesnej ekspansji kinaz białkowych oraz 4) zdolność do modyfikacji histonów, wskazująca na wczesne pojawienie się modelowania chromatyny u eukariotów. Należy również wspomnieć o dominującej roli cytoszkieletu tubulinowego, której towarzyszy brak homologów białek filamentów pośrednich oraz redukcja cytoszkieletu aktyno-miozynowego. Głównymi czynnikami kształtującymi ewolucję genomów *Trypanosoma* były insercje nowych fragmentów chromosomów oraz substytucje istniejących już odcinków, które rozrywały bloki synteniczne generując geny specyficzne gatunkowo. Z kolei, w obrębie linii *Leishmania*, zmiany polegały głównie na utracie genów, które ulegały wcześniej przekształceniu w pseudogeny.

Zsekwencjonowanie genomów świdrowców pozwoliło na wyselekcjonowanie dużej liczby białek nabytych od bakterii na drodze poziomego transferu genów (72,79), które są doskonałymi kandydatami do badań nad nowymi lekami. Wyprodukowanie skutecznych leków, a także szczepionek, jest pilne, ponieważ w ostatnich latach obserwujemy wzrost przypadków zachorowań na trypanosomatozy i leiszmaniozy. Opublikowane genomy pomogą również wyjaśnić zagadki związane z ewolucją świdrowców. Wciąż bowiem trwają debaty nad tym, czy posiadały one w swojej przeszłości plastyd (79-81).

2.1.4. Genomy gatunków *Phytophtora* – uciążliwe pasożyty roślin wyższych, które są blisko spokrewnione z glonami posiadającymi plastydy złożone

Przedstawiciele rodzaju *Phytophtora* są uciążliwymi, grzybopodobnymi pasożytami roślin należącymi do lęgniowców (Oomycetes) (33). Wspólnie z okrzemkami, brunatnicami, złotowiciowcami i pokrewnymi grupami glonów, tworzą one gromadę Stramenopila (33). Dotychczas poznano kompletne sekwencje genomów trzech gatunków *Phytophtora*: 1) *P. infestans* (zaraza ziemniaczana) – infekującej głównie

ziemniaki (82), 2) *P. sojae* – zarażającej soję (82,83) oraz 3) *P. ramorum* – infekującej 100 różnych gatunków roślin, w tym dęby i różaneczniki (82,83) (zob. także tab. 2). Pasożyty te powodują ogromne straty w rolnictwie i ekosystemach naturalnych. Epidemia zarazy ziemniaczanej w Irlandii w połowie XIX w. była przyczyną wielkiego głodu i masowych migracji ludności do USA.

Przedstawiciele rodzaju *Phytophtora* charakteryzują się szczególnie szybkimi zdolnościami adaptacyjnymi, co ma odzwierciedlenie w strukturze ich genomów (82,83). Porównanie genomów trzech gatunków *Phytophtora* ujawniło obecność konserwatywnych bloków sekwencji bogatych w kolinearnie ułożone geny, które rozdzielone są regionami zawierającymi mniej genów, ale o większej zawartości sekwencji powtórzonych i jednostek transpozonowych. To właśnie one odpowiedzialne są za rearanżacje i niezwykłą zmienność tych obszarów. Udział sekwencji powtórzonych w genomach *Phytophtora* wzrasta wraz z ich wielkością: od 28% (u *P. ramorum*), poprzez 39% (u *P. sojae*) do 74% (u *P. infestans*). Ten ostatni gatunek posiada również od 5 do 10 razy więcej różnych ruchomych elementów genetycznych niż dwa pozostałe gatunki. Najbardziej interesujące jest to, że w tych dynamicznych regionach zgromadzone są przede wszystkim szybko ewoluujące geny, które kodują efektorowe białka sekrecyjne odpowiedzialne za patogenność i infekcyjność przedstawicieli *Phytophtora* oraz zmianę fizjologii swoich gospodarzy. Białka te reprezentowane są przez tzw. efektory apoplastowe wydzielane do przestrzeni międzykomórkowej, takie jak: 1) hydrolazy degradujące tkanki, 2) inhibitory enzymów gospodarza i 3) toksyny nekrotyczne. Inną grupą efektorowych białek sekrecyjnych są białka cytoplazmatyczne, reprezentowane przez RXLR i CRN, które transportowane są do wnętrza komórek roślin przez specjalną ssawkę. Białka te są szczególnie liczne i zróżnicowane u *P. infestans*. Każdy gatunek *Phytophtora* ma własny unikatowy repertuar genów determinujących jego specyficzność względem gospodarza, sposób infekcji i charakter patogenności. Szybka ewolucja tych genów jest napędzana przez częste rekombinacje niealleliczne oraz duplikacje tandemowe, co generuje dużą zmienność genetyczną, zapewniającą plastyczność adaptacyjną gatunkom *Phytophtora*.

Chociaż przedstawiciele *Phytophtora* są blisko spokrewnieni z glonami zaliczanymi do Stramenopila (33,84,85), to jednak nie znaleziono u nich plastydu. Zgodnie z hipotezą chromalweolata, wspólny przodek stramenopili, alweolatów, kryptofitów i haptofitów, nabył na drodze wtórnej endosymbiozy plastyd od krasnorostu, który został następnie utracony w różnych liniach heterotroficznych (np. u orzęsków), włączając część form pasożytniczych, takich jak przedstawiciele rodzaju *Phytophtora* (84,85). W genomach gatunków *Phytophtora* zidentyfikowano geny pochodzące od krasnorostów i sinic (83), co mogłoby świadczyć o wcześniejszej obecności plastydu u tych pasożytów. Jednak w przeprowadzonych analizach statystycznych nie wykazano, aby ich udział był wystarczający do poparcia hipotezy chromalweolata (86).

2.2. Protisty ważne ekologicznie

Protisty o dużym znaczeniu ekologicznym, to przede wszystkim różne glony fotosyntetyzujące będące istotnym składnikiem planktonu, które są odpowiedzialne za wiązanie ditlenku węgla i produkcję pierwotną w ekosystemach mórz i oceanów. Genomy gatunków o takim znaczeniu poznano w dwóch grupach: u zielenic i okrzemek. U glonów tych wyewoluowały różne strategie adaptacyjne (np. związane z występowaniem odmiennych aparatów fotosyntetycznych czy osobliwych selenoenzymów), natomiast w ich genomach dominują geny zaangażowane w metabolizm komórkowy, które ułatwiają optymalne wykorzystanie zasobów środowiska.

2.2.1. Niewielkie genomy pikoeukariotów: *Micromonas* i *Ostreococcus*

Pikoeukarioty są formami planktonowymi, charakteryzującymi się niewielkimi rozmiarami komórek – od 0,2 do 2 μm. Należą do nich m.in. fotosyntetyzujace prasinofity (Prasinophyceae), które odgałęziają się u podstawy drzewa filogenetycznego roślin zielonych (87). Są to prosto uorganizowane glony jednokomórkowe z jednym plastydem i mitochondrium, pozbawione ściany komórkowej i wici. Pikoeukarioty są jednak szeroko rozprzestrzenione, od mórz tropikalnych po polarne, i odgrywają dużą rolę w ich produkcji pierwotnej, procesach biogeochemicznych i łańcuchach troficznych (88,89). Stanowią one istotne wskaźniki kondycji ekosystemów morskich i, w konsekwencji, zmian klimatycznych. Obecnie znamy kompletne sekwencje genomów czterech parsinofitów. Są to dwa gatunki *Ostreococcus*: *O. lucumarinus* (90) i *O. tauri* (90,91) oraz dwa szczepy *Micromonas pusilla* oznaczone jako: CCMP1545 i RCC299 (92) (zob. także tab. 2).

Cechą charakterystyczną genomów *Ostreococcus* (90,91) i *Micromonas* (92) są ich niewielkie rozmiary, przy czym najbardziej zredukowane genomy posiadają gatunki *Ostreococcus*. Należy podkreślić, że w genomach *Micromonas* znaleziono geny związane z rozmnażaniem płciowym i tworzeniem ściany komórkowej. Odkrycie to jest nieoczekiwane, ponieważ procesów tych nie zaobserwowano dotychczas u tego prasinofita. Genomy *Micromonas* kodują więcej rodzin transporterów błonowych oraz białek zapewniających ochronę przed metalami ciężkimi i wolnymi rodnikami niż genomy *Ostreococcus*. Jednak w przeciwieństwie do genomów *Ostreococcus*, są one ubogie (szczep CCMP1545) lub pozbawione (szczep RCC299) transpozonów.

Zsekwencjonowane genomy dwóch szczepów *M. pusilla* są prawie identyczne (97%) pod względem sekwencji 18S rRNA, mają jednak tylko 90% wspólnych genów, co wskazuje, że mogą one być osobnymi gatunkami. Hipoteza ta jest zgodna z istniejącymi pomiędzy nimi dużymi różnicami ekologicznymi. Szczep CCMP1545 został wyizolowany z wód umiarkowanych, natomiast RCC299 – z wód tropikalnych. Geny specyficzne dla poszczególnych szczepów pochodzą najprawdopodobniej z poziomych transferów od bakterii i opistokontów. Szczep CCMP1545 nie posiada

niektórych rodzin białek (np. transporterów dla związków azotu), które są niezbędne dla szczepu RCC299, żyjącego w wodach bardziej oligotroficznych. Ponadto szczepy *M. pusilla* różnią się: 1) organizacją przełączników RNA (tzw. ryboprzełączników), 2) składem komponentów RNAi, 3) liczbą intronów oraz 4) obecnością sekwencji powtórzonych.

Genomy *Ostreococcus* są bardzo zredukowane, co związane jest z miniaturyzacją ich komórek (około 1 μm). Genomy te cechują się szczególnie dużą gęstością kodowania. Bardzo mało jest sekwencji powtórzonych. Introny są stosunkowo krótkie, a liczba zduplikowanych genów jest niewielka. Ponad 300 genów uległo fuzji, a wiele genów z dużych rodzin białek zostało utraconych, np. geny zaangażowane w biogenezę wici oraz ściany komórkowej czy geny kodujące czynniki transkrypcyjne i białka będące składnikami chromatyny.

Najciekawszą adaptacją *Ostreococcus* jest stosunkowo duża liczba białek zawierających selenocysteinę, która kodowana jest przez kodon TGA. U *O. lucumarinus* białek tych jest 20, u *O. tauri* – 19, podczas gdy w 10 razy większym genomie zielenicy *Chlamydomonas reinhardtii* – tylko 10. Selenoenzymy są bardziej reaktywne niż enzymy bez selenu, dlatego komórka zawierająca te białka może ograniczyć liczbę innych enzymów. Analizy genomów *Ostreococcus* wskazują na istnienie kilku interesujących cech fenotypowych, takich jak: 1) optymalizacja w wykorzystaniu zasobów środowiska (np. występowanie transporterów amonowych), 2) obecność unikatowego metabolizmu metali oraz 3) adaptacja do niskiej zawartości żelaza w środowisku. Ich aparat fotosyntetyczny charakteryzuje się brakiem typowych białek układu antenowego LHCII, natomiast obecnością – specyficznych dla prasinofitów – białek wiążących chlorofil oraz białek związanych z fotosyntezą typu C4. *O. lucumarinus* przystosowany jest do większej intensywności światła i, w związku z tym, posiada dwie kopie genu ferrytyny, podczas gdy niewyspecjalizowany *O. tauri* – tylko jedną.

Na podstawie analiz porównawczych genomów *O. lucumarinus* i *O. tauri* wykazano, że 18 z 20 ich chromosomów charakteryzuje się wysokim poziomem syntenii i konserwatywności. Jednak chromosom 2 jest bogaty w introny i charakteryzuje się dużym poziomem rearanżacji genów, za co odpowiedzialne są występujące licznie na nim transpozony. Przypuszcza się, że może to być chromosom płciowy. Dodatkowo, po jednym chromosomie u każdego gatunku, nie wykazuje prawie żadnych podobieństw i koduje geny związane z biosyntezą, modyfikacją lub transportem węglowodanów. Geny te najprawdopodobniej pochodzącą z poziomych transferów. Mogą one odgrywać rolę w zmianie glikozylacji powierzchni komórki, stanowiąc w ten sposób ochronę przed fagami oraz protistami fagotroficznymi. Jeden z chromosomów u *O. lucumarinus* powstał w wyniku fuzji dwóch innych. Opisane zmiany chromosomów są najprawdopodobniej związane ze specjacją w obrębie rodzaju *Ostreococcus*.

2.2.2. Zróżnicowane genomy okrzemek: *Phaeodactylum tricornutum* i *Thalassiosira pseudonana*

Okrzemki (z grupy Stramenopila) są glonami jednokomórkowymi zasiedlającymi zarówno wody śródlądowe, jak i morskie (33,93,94). Mają one szczególne znaczenie ekologiczne. Ich rola w globalnym obiegu węgla i globalnej fotosyntezie jest porównywana do produktyności wszystkich lasów deszczowych razem wziętych. Charakterystyczną cechą ich komórek jest obecność krzemionkowej ściany komórkowej. Ponadto okrzemki posiadają plastyd czterobłonowy, który został nabyty od krasnorostu na drodze endosymbiozy wtórnej (93-96).

Dotychczas poznano genomy dwóch gatunków okrzemek: *Thalassiosira pseudonana* (93,94) oraz *Phaeodactylum tricornutum* (93). Pierwszy z nich należy do okrzemek promienistych, natomiast drugi – do okrzemek pierzastych. Chociaż linie te oddzieliły się zaledwie 90 milionów lat temu, to struktura i organizacja genomów *Thalassiosira* i *Phaeodactylum* są bardzo różne. Nie stwierdzono większych obszarów syntenii. Ponad 40% genów nie posiada homologów w drugim genomie. W genomie *T. pseudonana* występuje więcej intronów niż u *P. tricornutum* (odpowiednio 1,5 i 0,79 intronów na gen). Z kolei, genom *Phaeodactylum* charakteryzuje się większą aktywnością retrotranspozonów niż genom *Thalassiosira*. Ich udział w każdym z genomów wynosi odpowiednio 3,8 i 1,1%. Retrotranspozony odegrały z pewnością istotną rolę w rearanżacji tych genomów. Przypuszcza się, że wysokie tempo różnicowania się genomów i genów u *T. pseudonana* i *P. tricornutum* umożliwiło odniesienie przez nie sukcesu ewolucyjnego i intensywną ekspansją w morzach i oceanach.

Poznanie genomów *Phaeodactylum* i *Thalassiosira* pozwoliło na zidentyfikowanie nowych genów, zaangażowanych w: 1) transport kwasu krzemowego, 2) tworzenie ściany komórkowej, 3) pobieranie żelaza, 4) biosyntezę wielonienasyconych kwasów tłuszczowych oraz 5) wykorzystywanie związków azotu. Znajomość szczegółów biosyntezy nanostruktur w pancerzykach okrzemek, opartych na analizach ich genomów, może okazać się istotna w dalszym rozwoju nanotechnologii bazujących na krzemie (97,98). Zaskakujące było odkrycie u nich kompletnego cyklu mocznikowego, zintegrowanego z metabolizmem komórkowym. Szczególnie liczne są rodziny: 1) kinaz histydynowych, 2) cyklin (regulujących cykl komórkowy), 3) białek związanych z metabolizmem poliamidów długołańcuchowych (uczestniczących w tworzeniu ściany komórkowej) oraz 4) czynników transkrypcyjnych białek szoku termicznego (ulegających ekspresji w pewnych warunkach wzrostu i stanowiących połowę wszystkich czynników transkrypcyjnych).

Interesujące jest występowanie w genomach *T. pseudonana* i *P. tricornutum* ponad 300. genów nabytych w wyniku poziomego transferu z różnych grup prokariotów, o rząd wielkości więcej niż w genomach innych eukariotów, ale podobnie jak u bakterii. Nabyte geny istotnie kształtują funkcjonowanie komórek okrzemek i są odpowiedzialne za wykorzystanie węgla organicznego oraz azotu. Uczestniczą one ponadto w: 1) cyklu mocznikowym, 2) metabolizmie poliamin, 3) tworzeniu ściany ko-

mórkowej, 4) nietypowych mechanizmach replikacji, 5) naprawie i rekombinacji DNA oraz 6) odbieraniu sygnałów ze środowiska opartych na kinazach histydynowych.

Plastyd okrzemek został nabyty od krasnorostu (95,96). Jednak, w przeciwieństwie do plastydu kryptofitów (zob. podrozdz. 2.3.3), utracił on jądro komórkowe endosymbionta. Ponadto genomy plastydowe okrzemek kodują znacznie mniej genów niż genomy plastydowe krasnorostów (99). Wszystko to wskazuje na intensywny transfer genów z genomów endosymbionta do genomu jądrowego gospodarza. Na podstawie analiz zsekwencjonowanych genomów *Phaeodactylum* i *Thalassiosira* wykazano, że zawierają one wiele genów mających pochodzenie krasnorostowe (93,94). Geny te kodują białka z charakterystycznymi dwuczęściowymi presekwencjami, które składają się z peptydu sygnałowego oraz peptydu tranzytowego, i odpowiadają za ich import do plastydu czterobłonowego (100). Nieoczekiwane wyniki przyniosły dalsze analizy filogenomiczne, które wykazały obecność ponad dwukrotnie więcej genów pochodzenia zielenicowego niż krasnorostowego (101). Dane te można zinterpretować w dwojaki sposób: albo przodek okrzemek posiadał endosymbionta lub plastyd o pochodzeniu zielenicowym, albo też geny te zostały nabyte od zielenic w wyniku wielu poziomych transferów. Przyszłe studia powinny pomóc w sprecyzowaniu ich pochodzenia.

2.3. Modelowe gatunki glonów i ich zredukowane jądra występujace jako nukleomorfy

Aby zrozumieć funkcjonowanie komórek eukariotycznych zdolnych do przeprowadzania fotosyntezy, zsekwencjonowano genomy relatywnie prosto uorganizowanych glonów jednokomórkowych, takich jak: zielenice i krasnorosty. Dzięki zwartości ich genomów oraz obecności najbardziej istotnych genów, można łatwo przeanalizować wiele istotnych zjawisk i procesów zachodzących na poziomie molekularnym oraz zrozumieć ich funkcjonowanie w kontekście całej komórki eukariotycznej. Skrajnym przypadkiem uproszczenia genomów zielenic i krasnorostów są nukleomorfy – silnie zredukowane jądra komórkowe, które, wraz z całą komórką, zostały pochłonięte przez inne protisty heterotroficzne, co doprowdziło do powstania plastydów wielobłonowych i wyłonienia się nowych linii ewolucyjnych w obrębie Eukaryota.

2.3.1. Genom modelowej zielenicy *Chlamydomonas reinhardtii*

C. reinhardtii (zawłotnia) jest zielenicą jednokomórkową, służącą za bardzo dobry system modelowy do badania zarówno zjawiska fotosyntezy, jak i funkcjonowania wici, które są nieobecne w komórkach roślin wyższych (33,102). Zielenica ta jest bardzo dobrym obiektem do badań, ponieważ charakteryzuje się krótkim czasem

trwania generacji i może rosnąć również w ciemności, wykorzystując octan jako źródło węgla i energii.

Cechą charakterystyczną genomu *Chlamydomonas* (102) (zob. również tab. 2) jest duża zawartość intronów w porównaniu do innych eukariotów jednokomórkowych oraz roślin wyższych. U *C. reinhardtii* liczba eksonów na gen wynosi 8,3, natomiast u rośliny wyższej *A. thaliana* – 5,2. Introny w genach *Chlamydomonas* są średnio dłuższe (373 bp) niż introny w genach *Arabidopsis* (164 bp). Sugeruje się, że duży udział w ich modyfikacjach mogły mieć transpozony. W genomie *C. reinhardtii* genów podzielonych jest 92%, podczas gdy w genomie *A. thaliana* – 79%.

W wyniku analiz porównawczych genomu *Chlamydomonas* z innymi zsekwencjonowanymi genomami eukariotów zidentyfikowano dużą liczbę genów odpowiedzialnych za biogenezę plastydów (349 genów) oraz wici (195 genów), przy czym wiele z nich nie było wcześniej znanych. Podobnie jak u *A. thaliana*, bardzo licznymi rodzinami białek są kinazy i cytochromy P-450. Jednak najliczniejszą rodziną, również w porównaniu z innymi organizmami, jest klasa III cyklaz guanylowych i adenylowych. Nukleotydy cykliczne, tworzone przez te enzymy, uczestniczą w przekształcaniu sygnałów w komórce i są istotne u *C. reinhardtii* w procesach: rozmnażania, funkcjonowania wici oraz przystosowywania się do zmiennych warunków troficznych w środowisku. Adaptację do takich warunków umożliwiają również zróżnicowane transportery błonowe, specyficzne dla poszczególnych związków, wśród których najliczniejsze są transportery: fosforanowe PiT i potasowe KUP. W obrębie transporterów błonowych zidentyfikowano także transportery charakterystyczne dla komórek zwierząt (np. regulujące poziom kationów wapnia i innych jonów), które związane są z funkcjonowaniem wici.

2.3.2. Uproszczony genom termofilnego i kwasolubnego krasnorostu *Cyanidoschyzon merolae*

C. merolae jest bardzo małym (2 μm) krasnorostem jednokomórkowym, żyjącym w gorących źródłach siarkowych o niskim pH (33,103,104). Posiada on minimalny zestaw organelli komórkowych oraz mały i prosto zorganizowany genom, co czyni go idealnym organizmem modelowym do badania funkcjonowania i ewolucji komórki eukariotycznej.

Osobliwą cechą genomu *Cyanidoschyzon* (103,104) (zob. również tab. 2) jest bardzo wysoka używalność dinukleotydów CpG w porówaniu do innych eukariotów. Jednostki transpozonowe stanowią zaledwie 0,7% jego genomu, natomiast obecne są kopie nieznanych rozproszonych elementów powtórzonych, które są aktywne transkrypcyjnie i stanowią 5% genomu. Podstawowy zestaw istotnych genów czyni *C. merolae* bardzo dobrym modelem do badań proteomicznych. Jedynie w 26 genach stwierdzono występowanie intronów, a aż 25 z nich zawiera tylko jeden intron. Znaleziono jedynie 3 kopie rDNA, bez tandemowych powtórzeń typowych dla euka-

riotów, oraz tylko 30 genów tRNA. Stwierdzono zmniejszony udział genów związanych z: 1) biosyntezą metabolitów wtórnych, 2) transportem komórkowym, 3) katabolizmem oraz 4) transdukcją sygnałów, co wiąże się ze słabym zróżnicowaniem komórki *Cyanidoschyzon*. Cechą charakterystyczną genomu tego krasnorostu jest najmniejszy znany klaster genów kodujących histony, który zlokalizowany jest tylko na jednym chromosomie.

W genomie *C. merolae* występują tylko dwa geny dynamin, związane z podziałem mitochondriów i plastydów. Nie stwierdzono natomiast genów dla dyneiny i miozyny, co może świadczyć o prostym mechanizmie cytokinezy. Aparat importu białek do plastydu nie zawiera trzech komponentów: Toc159, Tic40 i Tic55, które obecne są u roślin wyższych. *Cyanidoschyzon* posiada również najprostszy zestaw chlorofili i karotenoidów wśród eukariotów fotosyntetyzujących. Podobnie jak u roślin wyższych, znaleziono u niego ten sam mozaikowy (tj. pochodzenia sinicowego i eukariotycznego) zestaw enzymów cyklu Calvina, potwierdzając w ten sposób koncepcję jednorazowego nabycia plastydu przez wspólnego przodka zielenic i krasnorostów (105).

Genom mniejszy od genomu *C. merolae* odkryto u zielenicy *Ostreococcus* (zob. podrozdz. 2.2.1.), jednak nie jest on w takim samym stopniu uproszczony strukturalnie jak genom *Cyanidoschyzon*. Ponieważ *C. merolae* żyje w ekstremalnych warunkach środowiskowych, nie wiadomo, czy uproszczenie jego genomu jest cechą pierwotną, czy raczej wtórną, wynikającą z przystosowania do specyficznych warunków życia. W uzyskaniu odpowiedzi na to pytanie może pomóc poznanie genomu kolejnego krasnorostu zamieszkującego gorące źródła – *Galdieria sulphuraria*. Bardziej szczegółowe analizy białek *Cyanidoschyzon* będą istotne w poznaniu ich adaptacji strukturalnych do wysokiej temperatury i dużej kwasowości. Dane te mogą znaleźć ważne zastosowanie w biotechnologii, pomagając w projektowaniu bardziej stabilnych białek.

2.3.3. Genomy nukleomorfów – skrajnie zredukowane jądra komórkowe endosymbiontów pochodzących od krasnorostów i zielenic

Nukleomorfy to niezwykłe organelle komórkowe występujące w przestrzeni periplastydowej niektórych typów plastydów czterobłonowych (42,106-108). Są one zredukowanymi jądrami komórkowymi glonów, które zostały pochłonięte przez jednokomórkowe eukarioty heterotroficzne, a następnie przekształcone w plastydy złożone. Zjawisko to określane jest jako endosymbioza wtórna, w przeciwieństwie do endosymbiozy pierwotnej, polegającej na pochłonięciu cyjanobakterii, czyli organizmu prokariotycznego (95,96). Nukleomorfy zostały zachowane tylko w dwóch liniach filogenetycznych: u kryptofitów, które weszły w endosymbiozę z krasnorostem oraz u chlorarachniofitów, które nabyły swój plastyd od zielenicy (42,106-108).

Chociaż nukleomorfy kryptofitów i chlorarachniofitów wyewoluowały niezależnie, to jednak ich genomy są uderzająco podobne zarówno pod względem ich wielkości, jak i struktury. Obecnie dysponujemy kompletnymi sekwencjami nukleomorfów dwóch kryptofitów: *Guillardia theta* (106) i *Hemiselmis andersenii* (107) oraz jednego chlorarachinofita: *Bigelowiella natans* (108) (zob. także tab. 2). Nukleomorfy są jednymi z najmniejszych genomów, nawet w porównaniu z genomami prokariotów. Ich wielkość (< 1 Mb), a także potencjał kodujący, uległy drastycznej redukcji w porównaniu do spokrewnionych z nimi form wolnożyjących. Na przykład wielkość genomu nukleomorfu *B. natans* i liczba jego genów są odpowiednio 300- i 50-krotnie mniejsze niż u zielenicy *C. reinhardtii* (zob. podrozdz. 2.3.1). Zarówno nukleomorfy kryptofitów, jak i chlorarachniofitów zawierają tylko 3 chromosomy, dostarczając w ten sposób interesujący przykład konwergencji ewolucyjnej. Sugeruje się, że mniejsze, ale liczniejsze chromosomy mogłyby zostać zagubione podczas podziału mitotycznego, natomiast w przypadku większych chromosomów istniałby problem z ich upakowaniem w nukleomorfie. Trzy chromosomy, o wielkości od 95 do 210 kb, jak się wydaje, stanowią optymalne rozwiązanie dla zapewnienia stabilności genomu nukleomorfu.

Materiał genetyczny nukleomorfów wykazuje wiele cech genomów znajdujących się pod silną presją selekcyjną na redukcję ich wielkości. Wśród cech tych należy wymienić podwyższoną zawartość par AT (około 75%) i bardzo wysoką gęstość kodowania. Część genów nukleomorfów transkrybowana jest w postaci policistronowych jednostkek transkrypcyjnych, a niektóre są nawet genami zachodzącymi. Ponadto ekstremalna redukcja genomów nukleomorfów pociąga za sobą skrajne skrócenie sekwencji kodujących białka. Podobnie jak u mikrosporidii *E. cuniculi* (zob. podrozdz. 2.1.1), niemal wszystkie kodowane polipeptydy są krótsze od ich homologów u innych organizmów. Wciąż nie jest znana funkcja wielu genów (prawie połowy w nukleomorfie *G. theta*), a większość genów o znanych funkcjach związana jest z podtrzymywaniem aktywności nukleomorfu i rybosomów, znajdujących się w kompartymencie periplastydowym.

Nukleomorfy kryptofitów i chlorarachniofitów charakteryzują się podobną zawartością procentową genów uczestniczących w różnych procesach, takich jak: transkrypcja, translacja czy biogeneza plastydu. Wyjątek stanowią geny zaangażowane w metabolizm RNA, których jest zdecydowanie więcej u *B. natans* niż u *G. theta* i *H. andersenii*, co ma związek z odmienną ewolucją intronów w nukleomorfach chlorararchinofitów i kryptofitów. Nukleomorf *B. natans* zawiera 852 introny (średnio 3,1 intronu na gen), które uległy miniaturyzacji do długości 18-21 nt. Jest interesujące, że introny te są bardzo konserwatywne, ponieważ porównanie ich rozmieszczenia z nukleomorfem innego chlorarachniofita – *Gymnochlora stellata* wskazuje na zachowanie około 94% z nich. Introny w nukleomorfach chlorarachnofitów są bardzo stabilne i ulegają utracie wolniej niż introny zielenic czy roślin wyższych. W przeciwieństwie do nukleomorfów *B. natans* i *G. stellata*, nukleomorf kryptofita *G. theta* zachował zaledwie 17 intronów, podczas gdy nukleomorf *H. andersenii*

utracił całkowicie introny razem z maszynerią splicingową. Dane te wskazują na istnienie dwóch strategii ewolucyjnych: pierwsza z nich polega na zachowywaniu intronów z jednoczesną ich miniaturyzacją (chlorarachniofity), natomiast druga na – ich całkowitej eliminacji (kryptofity).

Nukleomorfy kryptofitów kodują 30, natomiast chlorarachniofitów 17 białek, które są importowane do ich plastydów. Białka te syntetyzowane są w kompartymencie periplastydowym, a następnie przechodzą przez dwie błony właściwego plastydu do stromy. Jest niezwykle interesujace, że nukleomorfy są utrzymywane tylko dla ekspresji zaledwie od kilkunastu do kilkudziesięciu białek plastydowych, czyli jeszcze mniej niż w przypadku genomów plastydowych (99). W innych typach plastydów wtórnych (np. u stramenopili; zob. podrozdz. 2.2.2.), jądro endosymbionta zostało całkowicie utracone (95,96). Być może utrzymywanie nukleomorfów wynika z trudności w transferze ich genów plastydowych do jądra gospodarza (42).

Nukleomorfy są niezwykle interesującymi przykładami miniaturyzacji jąder komórkowych (genomy typu *bonsai*), a ich dokładne poznanie ma fundamentalne znaczenie dla zrozumienia zjawiska endosymbiozy (42). Ze względu na swoją prostą organizację, nukleomorfy mogą być także pomocne w zrozumieniu bardziej złożonych procesów zachodzących w jądrze komórkowym, wciąż bowiem są one funkcjonalnym, choć silnie zredukowanym jądrem, które utrzymuje kontakt z cytoplazmą i plastydem swojej dawnej komórki, jak i z komórką oraz jądrem nowego gospodarza (42).

2.4. Wysoko uorganizowane protisty mogące służyć jako modele dla organizmów wielokomórkowych

Projekty sekwencjonowania genomów protistów o złożonej budowie i zaawansowanych procesach komórkowych mają na celu opracowanie systemów modelowych, służących do zrozumienia zjawisk zachodzących u znacznie bardziej skomplikowanych organizmów wielokomórkowych. Z tego powodu zsekwencjonowano następujące genomy: 1) orzęsków *Paramecium tetraurelia* (pantofelek), 2) *Tetrahymena thermophila*, 3) ameby *Dictyostelium discoideum* oraz 4) choanoflagellata *Monosiga brevicollis*. Ich genomy kodują wiele genów obecnych również u wielokomórkowców. U orzęsków, podobnie jak u zwierząt, funkcje somatyczne i rozrodcze są rozdzielone i realizowane przez dwa różne jądra: makronukleus i mikronukleus. *D. discoideum* wykształcił, niezależnie od roślin i zwierząt, formy wielokomórkowe będące wynikiem specyficznych interakcji międzykomórkowych. Z kolei, *M. brevicollis* należy do tej samej linii filogenetycznej co zwierzęta.

2.4.1. Bogate w geny genomy orzęsków: *Paramecium tetraurelia* i *Tetrahymena thermophila*

Genom orzęsków jest szczególnie interesujący, ponieważ jest on złożony z dwóch jąder: diploidalnego mikronukleusa – zaangażowanego w rozmnażanie płciowe oraz poliploidalnego makronukleusa – pełniącego funkcje somatyczne (43,109,110). Makronukleus powstaje z mikronukleusa w wyniku skomplikowanych procesów amplifikacji i delecji różnych fragmentów DNA (m.in. sekwencji powtórzonych i transpozonów) oraz dodawania telomerów. Analizy genomowe *Tetrahymena* wskazują, że wycinanie materiału genetycznego służy do pozbycia się obcego DNA z genomu. Z powodu małej ilości DNA powtórzonego, makronukleus jest łatwiejszy w składaniu zsekwencjonowanych fragmentów niż mikronukleus, dlatego dotychczas tylko jego sekwencje są znane. Opublikowano kompletne sekwencje dwóch makronukleusów: *P. tetraurelia* (109) i *T. thermophila* (110) (zob. także tab. 2).

U *Paramecium* i *Tetrahymena* zidentyfikowano bardzo dużą liczbę genów jak na organizmy jednokomórkowe: ponad 27 tysięcy – u *T. thermophila* i prawie 40 tysięcy – u *P. tetraurelia*, co znacznie przewyższa liczbę genów znalezionych u wielu wyższych eukariotów. Tak duża liczba genów jest u *Tetrahymena* wynikiem licznych duplikacji, które zaszły w linii prowadzącej do tego gatunku, natomiast u *Paramecium* – konsekwencją przynajmniej trzech kolejnych duplikacji całego genomu (WGD, ang. *Whole-Genome Duplication*). Cechą charakterystyczną tych całościowych duplikacji genomu, która jest rzadko spotykana w zduplikowanych genomach innych eukariotów (jak drożdże, ryby, rośliny wyższe), jest duża konserwatywność w ułożeniu genów w zduplikowanych chromosomach. Inną interesującą cechą genomów orzęsków jest duża liczba zachowanych zduplikowanych genów. Tylko 32% z nich utraciło swojego paraloga. Stwierdzono, że geny kodujące białka zaangażowane w te same szlaki metaboliczne lub tworzące te same kompleksy wykazują takie same reguły zanikania. Wynika to z konieczności zachowania właściwych stosunków stechiometrycznych pomiędzy oddziałujących białkami. Wykazano również, że geny kodujące: 1) białka rybosomów, 2) komponenty cytoszkieletu, 3) histony, 4) czynnik elongacji translacji i 5) dehydrogenazę bursztynianową, które charakteryzują się wysokim stopniem ekspresji, były preferencyjnie zachowywane po kolejnych rundach duplikacji. Wiele z nich posiada wszystkie z 8 możliwych kopii. Ostatnia całościowa duplikacja genomu *P. tetraurelia* pokrywa się w czasie z eksplozją specjacji w obrębie rodzaju *Paramecium*, która doprowadziła do wyodrębnienia się piętnastu gatunków siostrzanych. W przypadku *T. thermophila* duplikacjom podlegały przede wszystkim geny związane z reagowaniem i odpowiedzią na czynniki środowiska. Kodowane przez nie białka odpowiadają za: 1) przekazywanie sygnałów (np. kinazy), 2) wykorzystanie różnych zasobów środowiska (np. proteazy oraz transportery: ABC, jonowe VIC i ATPazy typu P), 3) zmianę kształtu i ruch komórki (np. kinezyny i dyneiny) oraz 4) wewnątrzkomórkowy transport błonowy (np. białka Rab). Jest interesujące, że u *T. thermophila* jako kodon stop wykorzystywany jest jedynie kodon UGA, który dodatkowo, w niektórych

genach, może służyć do kodowania selenocysteiny, co czyni ten gatunek pierwszym znanym, u którego wszystkie 64 kodony mogą być tłumaczone na aminokwasy.

Dokładniejsze przeszukiwania genomów orzęsków zaowocowały znalezieniem szesnastu genów pochodzących od glonów (111), co zgodnie z hipotezą chromalweolata, miałoby świadczyć o występowaniu plastydu u przodka Alveolata, do których należą orzęski (33,84,85). Jednak obecność tych genów można równie dobrze wytłumaczyć poziomymi transferami genów od glonów pochłanianych przez orzęski jako pokarm (112).

Orzęski, a szczególnie *Tetrahymena*, są łatwe w hodowli i charakteryzują się obecnością wielu procesów komórkowych, które zachodzą również u eukariotów wyższych (włączajac człowieka), a są nieobecne u innych jednokomórkowych organizmów modelowych, takich jak drożdże: *S. cerevisiae* i *Schizosaccharomyces pombe*. To czyni je idealnym modelem do różnorodnych analiz biologicznych, biomedycznych i biotechnologicznych.

2.4.2. Genom ameby *Dictyostelium discoideum* o ciekawych zachowaniach społecznych

D. discoideum jest amebą należącą do śluzorośli, która wykazuje interesujące zachowania socjalne, polegające na wyjątkowych zdolnościach do organizowania się pojedynczych komórek i tworzenia form wielokomórkowych z charakterystycznymi ciałami owocowymi zbudowanymi z trzonków, które zakończone są skupieniami zarodników (33,113). Organizacyjnie zajmuje ona stanowisko pośrednie pomiędzy jedno- i wielokomórkowymi eukariotami. Z tego powodu *Dictyostelium* jest bardzo dobrym organizmem modelowym do badania zjawisk związanych z: 1) kształtowaniem się wielokomórkowości, 2) ruchem komórek, 3) przekazywaniem sygnałów, 4) interakcjami międzykomórkowymi i 5) adhezją komórkową. Komórki tej ameby są również funkcjonalnym odpowiednikiem makrofagów ssaków i mogą służyć za model do ich badania

Genom *D. discoideum* (113) (zob. także tab. 2) jest bogaty w: 1) pary AT (77%), 2) transpozony występujące w klastrach i 3) różne sekwencje powtórzone, z których jedna tworzy centromery. Koduje on aż 390 genów tRNA, które zorganizowane są w pary. Z kolei, geny rRNA występują na elemencie pozachromosomowym o budowie palindromowej i wielkości 88 kb. Częściowe kopie tych genów występują na końcach chromosomów i tworzą nietypowe telomery. Genom *Dictyostelium* charakteryzuje się dużą gęstością kodowania (62%) oraz dużą liczbą genów, która porównywalna jest z liczbą genów u organizmów wielokomórkowych. Jednak w dokładniejszych analizach wykazano, że 20% genów zostało niedawno zduplikowanych. Introny i sekwencje międzygenowe są krótkie. Średnia długość intronu wynosi 146 bp, podczas gdy u człowieka – 3365 bp. W 34% białek występują długie ciągi powtarzających się aminokwasów, mogące służyć jako model do badania np. chorób neurodegeneracyjnych u człowieka, które związane są z ekspansją takich powtórzeń. Za pomocą *D. discoideum*

mogą być badane również inne choroby człowieka, ponieważ w jego genomie zidentyfikowano 64 geny homologiczne do genów człowieka związanych z różnymi chorobami. Genów tych jest więcej niż u innego gatunku modelowego – drożdży *S. cerevisiae*.

Genom *Dictyostelium* koduje szczególnie dużo syntaz poliketydowych, odpowiedzialnych za produkcję antybiotyków i metabolitów wtórnych, które eksportowane są na zewnątrz komórki przez liczne transportery ABC, pełniące również funkcje detoksykacyjne. W wykorzystaniu zasobów środowiska dużą rolę odgrywają także enzymy związane z degradacją celulozy. Najciekawsze są jednak białka odpowiedzialne za asocjacje komórek *D. discoideum*, które tworzą formy wielokomórkowe, wykształcone niezależnie od roślin i zwierząt. Wiele z tych białek – charakterystycznych tylko dla zwierząt – związanych jest z adhezją komórek oraz przekazywaniem sygnałów opartych na: 1) białkach G, 2) domenie SH2, 3) transporterach ABC oraz 4) licznie reprezentowanych kinazach. *Dictyostelium* zawiera również duży zestaw innych białek „zwierzęcych" związanych z cytoszkieletem oraz białka „roślinne" uczestniczące w odkładaniu i metabolizowaniu celulozy. Taki hybrydowy proteom *D. discoideum* wskazuje, że jego przodek, stojący u podstawy linii filogenetycznych prowadzących do roślin, zwierząt i ameb, wyposażony był w zróżnicowany zestaw genów.

2.4.3. Genom jednokomórkowego choanoflagellata *Monosiga brevicollis* blisko spokrewnionego ze zwierzętami

Grupą protistów, która jest najbliżej spokrewniona ze zwierzętami są choanoflagellaty znane również jako wiciowce kołnierzykowe. Kompletna sekwencja genomu jednego z nich, *M. brevicollis* (114) (zob. także tab. 2), pozwala na zrozumienie, w jaki sposób kształtowała się wielokomórkowość, oraz umożliwia wyobrażenie sobie, jak mógł wyglądać jednokomórkowy przodek zwierząt.

Chociaż genom *Monosiga* jest zwarty, to kodowane geny są tylko nieco mniej bogate w introny niż geny człowieka (odpowiednio 6,6 oraz 7,7 intronów na gen). Jednak introny *M. brevicollis* są znacznie krótsze (średnia długość – 174 bp) niż introny zwierząt i człowieka (zob. podrozdz. 2.4.2). Najciekawszym wynikiem analiz genomu *Monosiga* było zidentyfikowanie 78 domen białek, które występują tylko u Choanoflagellata i Metazoa. Wiele z nich związanych jest z przyleganiem komórek (domeny kadheryny, lektyny C, integryny α) i macierzą zewnątrzkomórkową (domeny kolagenowe lamininy, ependyminy, reeliny). Wśród tych domen białkowych, jest pięć domen immunoglobulin, dużo mniej w porównaniu z 150 do 1500 domenami obecnymi u tkankowców. Dane te wskazują na późną ekspansję rodziny białek immunoglobulin, już po dywergencji choanoflagellatów i zwierząt.

U *M. brevicollis* szczególnie liczne są domeny szlaku fosfotyrozyny, a zwłaszcza kinaz tyrozynowych (około 120), które są typowe dla zwierząt. Dane te świadczą o ich wczesnym powstaniu, jeszcze przed rozdzieleniem się linii wiciowców kołnierzykowych i zwierząt. Nie znaleziono jednak składników innych – typowo zwierzę-

cych – szlaków przekształcania sygnałów wewnątrz komórki oraz wielu czynników transkrypcyjnych związanych z rozwojem. Odkryto jedynie pewne komponenty szlaków Notch i hedgehog oraz reprezentantów rodzin czynników transkrypcyjnych: p53, Myc i Sox/TCF.

Na podstawie analiz genomu *Monosiga* wskazuje się, że przodek zwierząt posiadał już kilka istotnych strukturalnych i funkcjonalnych preadaptacji, potrzebnych do różnicowania się komórek i rozwinięcia zaawansowanych relacji pomiędzy nimi. Jednak białka o architekturze typowej dla tkankowców, które powstały w wyniku tasowania eksonów, pojawiły się później, po rozdzieleniu się linii choanoflagellatów i zwierząt.

Niedawno w genomie *M. brevicollis* znaleziono ponad 100 genów pochodzących od glonów (115). Czy to oznacza, że jednokomórkowi przodkowie zwierząt posiadali plastyd? Nie możemy obecnie wykluczyć takiej możliwości, ale należy pamiętać, że geny te mogły zostać nabyte na drodze wielokrotnych poziomych transferów, którym nie musiały towarzyszyć stabilne endosymbiozy plastydowe (114).

3. Rozpoczęte i przyszłe projekty sekwencjonowania genomów protistów

Według bazy GOLD (6), na początku 2010 r., wśród 1248 kontynuowanych i nie opublikowanych projektów sekwencjonowania genomów eukariotów, 217 projektów było poświęconych protistom. Tak jak przedstawiono to na rysunku, protisty te reprezentują wszystkie główne linie filogenetyczne eukariotów, chociaż reprezentacja ta nie jest jednakowa. Dominują projekty poświęcone alweolatom (w szczególności apikompleksom pasożytniczym) oraz zielenicom i stramenopilom z grupy Ochrophyta, które mają znaczenie ekologiczne. Najsłabiej reprezentowane są Rhizaria i Hacrobia, obejmujące gatunki słabo poznane zarówno pod względem taksonomicznym, jak i molekularnym. Jednak prawie 40% wszystkich projektów poświęconych protistom, to projekty EST, a nie genomowe. Dla porównania, projekty EST nie protistów stanowią około 25%. Projekty EST pozwolą niewątpliwie na szybkie i tanie uzyskanie potencjalnych sekwencji kodujących białka, ale nie odsłonią całej struktury genomów analizowanych protistów. Na stronie internetowej bazy GOLD (http://www.genomesonline.org) można znaleźć łącza do instytucji sekwencjonujących określone genomy oraz stron poświęconych poszczególnym projektom, na których zgromadzono informacje dotyczące sekwencjonowanych organizmów, stanu poznania ich genomów oraz uzasadnienia realizowanych projektów. Poniżej charakteryzujemy niektóre z tych projektów.

3.1. Genomy pasożytów

Najwięcej nowych projektów genomowych dotyczy pasożytów człowieka i ssaków, w tym wielu gatunków udomowionych. Protisty te należą do: 1) Apicomplexa (rodzaje: *Babesia*, *Crytposporidium*, *Eimeria*, *Neospora caninum*, *Plasmodium* oraz

Toxoplasma gondii), 2) Amoebozoa (*Acanthamoeba castellanii*), 3) Excavata (rodzaje: *Leishmania*, *Trypanosoma* oraz *Trichomonas vaginalis*) i 4) Stramenopila (*Blastocystis hominis*). Pomimo poznania genomów niektórych z tych pasożytów, w wielu przypadkach, sekwencjonowane są nadal genomy spokrewnionych z nimi gatunków lub należących do nich szczepów. Porównania zsekwencjonowanych genomów pozwolą znaleźć subtelne różnice na poziomie genomowym, decydujące o różnej inwazyjności i patogenności poszczególnych gatunków i szczepów. Ze względu na problemy z hodowaniem protistów pasożytniczych, do sekwencjonowania wybrane są także gatunki nieinwazyjne, blisko spokrewnione z formami pasożytniczymi, lub gatunki pasożytujące na organizmach łatwych w hodowli (np. ameby: *Entamoeba dispar*, *E. invadens*, *E. moshkovskii*, *E. terrapinae* czy ekskawat *Naegleria gruberi*).

Duże znaczenie ma również poznanie genomów gatunków pasożytujących na innych zwierzętach niż ssaki, które odgrywają ważną rolę w gospodarce, rolnictwie, hodowli i ekosystemach naturalnych. Do takich pasożytów należą: 1) skoczkowiec *Batrachochytrium dendrobatidis* (powodujący śmiertelne uszkodzenia w układzie oddechowym i nerwowym płazów), 2) orzęsek *Ichthyophthirius multifiliis* (odpowiedzialny za bardzo zaraźliwą i częstą chorobę ryb zwaną ichtioftiriozą), 3) alweolat *Perkinsus marinus* (powodujący duże spadki liczebności populacji ostryg) oraz 4) ekskawaty *Spironucleus vortens* (zasiedlający przewód pokarmowy ryb) i 5) *Crithidia deanei* (pasożytująca na owadach, w tym na pszczołach). Istotne jest poznanie genomów mikrosporidii, infekujących owady, skorupiaki i ryby. Szczególnie ważne są mikrosporidie z rodzaju *Nosema*, które powodują ogromne straty w hodowlach jedwabników i pszczół. Kilka projektów poświęconych jest pasożytom roślin należącym do Oomycetes (*Aphanomyces*, *Peronospora*, *Phytophthora*), które wyrządzają poważne szkody w rolnictwie i ekosystemach naturalnych.

Dzięki wymienionym projektom będzie możliwe odkrycie nowych genów związanych z patogennością, a także zrozumienie specyficznych mechanizmów infekcji, co może przyczynić się do bardziej skutecznego zapobiegania zakażeniom wywoływanym przez te pasożyty i ich efektywnego zwalczania.

3.2. Genomy gatunków istotnych ekologicznie

Niezwykle istotne są rozpoczęte projekty sekwencjonowania genomów gatunków o dużym znaczeniu ekologicznym, które należą do różnych grup glonów, takich jak: 1) bruzdnice (Dinoflagellata), 2) krasnorosty (Rhodophyta), 3) zielenice (Chlorophyta), 4) haptofity (Haptophyta) oraz 5) stramenopile z grupy Ochrophyta (w tym okrzemki). Wiele z tych gatunków jest kosmopolitycznych, a część należy do pikoplanktonu (116). Należy podkreślić ich niezmiernie duży udział w produkcji pierwotnej ekosystemów oraz obiegu węgla (a zwłaszcza wiązaniu ditlenku węgla) (117,118), co ma daleko idące konsekwencje dla klimatu. Wśród glonów tych, szczególne miejsce zajmują okrzemki, które wiążą 20% globalnego węgla nieorganiczne-

go, generując jednocześnie 40% węgla organicznego produkowanego w oceanach (119), co ma podstawowe znaczenie dla funkcjonowania ekosystemów morskich. Na uwagę zasługuje projekt genomowy poświęcony okrzemce *Fragilariopsis cylindrus*, stanowiącej podstawę ekosystemów mórz polarnych, gdzie fotosynteza roślin wyższych jest bardzo ograniczona. Regiony polarne są najbardziej czułe na zmiany klimatu i odpowiadają na to topnieniem lodowców. Celem projektu jest poznanie przystosowań *F. cylindrus* do bytowania w niskich temperaturach i jej reakcji na zmiany klimatyczne. Innym istotnym glonem polarnym jest haptofit *Phaeocystis antarctica*, którego genom jest również sekwencjonowany. Oprócz roli w globalnym obiegu węgla, uczestniczy on również w cyklu siarki, emitując 3-dimetylosulfoniopropionian (DMSP), który jest prekursorem dimetylosiarczku (DMS) (120,121). DMS jest bardzo lotny, a jego aerozole wpływają na formowanie się chmur i, w konsekwencji, na klimat. Poznanie szlaków metabolicznych zaangażowanych w tworzenie DMSP i DMS może pomóc przewidywać stopień zaawansowania procesu globalnego ocieplenia.

Niektóre gatunki glonów intensywnie się dzielą tworząc tzw. zakwity, co nie jest obojętne dla ekosystemów naturalnych, powodując ich zubożenie przez: zmniejszenie natężenia światła, spadek poziomu tlenu i produkcję toksyn (122). Z gatunków tworzących zakwity, obecnie sekwencjonowane są genomy: okrzemki *Pseudo-nitzschia multiseries* oraz bruzdnic, takich jak: *Karenia brevis* i *Alexandrium tamarense*. Produkują one neurotoksyny wywołujące choroby, a nawet przypadki śmiertelne u człowieka. Ponadto są one odpowiedzialne za masowe ginięcie zwierząt morskich, w tym gatunków hodowlanych, np. małży. W przypadku glonu *Aureococcus anophagefferens* (należącego do Ochrophyta) zakwity również doprowadzają do masowego obumierania planktonu, zostery morskiej i małży, co powoduje duże straty w gospodarce morskiej.

Aby zrozumieć i kontrolować całe cykle troficzne, zapoczątkowano projekty genomowe heterotroficznych (*Paraphysomonas*, *Spumella*) i miksotroficznych (*Ochromonas*) stramenopili z grupy Ochrophyta, które odżywiją się fitoplanktonem i regulują jego ilość, a, w konsekwencji, wpływają na obieg węgla w przyrodzie. Genomy haptofitów: *Emiliania huxleyi* i *Isochrysis galbana* wybrano do sekwencjonowania m.in. z powodu ich dużego znaczenia jako źródła pokarmu w akwakulturach: małży, krewetek i innych morskich organizmów. Zakłada się, że wiedza o ich genomach umożliwi zwiększenie produktywności tych hodowli. Sekwencjonuje się również genomy różnych ekotypów zielenicy *Ostreococcus* o zróżnicowanych wymaganiach środowiskowych, np. związanych z intensywnością światła. Projekty te mają wyjaśnić, jakie cechy ich małych genomów (zob. podrozdz. 2.2.1) wiążą się z adaptacją do zróżnicowanych nisz ekologicznych.

Dużym problemem ekologicznym pozostaje od wielu lat blaknięcie koralowców (123,124), spowodowane pozbywaniem się lub obumieraniem bruzdnic symbiotycznych z rodzaju *Symbiodinium* zwanych zooksantellami. Zooksantelle odpowiedzialne są za odżywianie polipów i deponowanie węglanu wapnia. Dlatego spadek ich liczebności prowadzi do obumierania rafy jako całego ekosystemu. Przypuszcza się,

że rozpoczęte projekty genomowe wyjaśnią naturę związku symbiotycznego pomiędzy koralowcami a zooksantellami, a tym samym mechanizmy blednięcia koralowców, co może pomóc w zapobieganiu tym katastrofalnym dla ekosystemów morskich zjawiskom.

Celem sekwencjonowania genomów gatunków ważnych ekologicznie jest również znalezienie dobrych markerów do oceny zmian zachodzących w ekosystemach. Ponadto dane te mogą pomóc w lepszym poznaniu, a być może nawet w przewidywaniu i kontrolowaniu, wielu ważnych procesów ekologicznych.

3.3. Genomy organizmów modelowych

Część projektów poświęcona jest tzw. organizmom modelowym, służącym do badania różnorodnych zjawisk biologicznych na poziomie komórkowym i molekularnym. Przykładem mogą być projekty poświęcone protistom fotosyntetyzujacym, takim jak: ekskawat *Euglena gracilis* oraz zielenice: *C. reinhardtii* i *Volvox carteri*.

Interesujące będzie porównanie genomu *C. reinhardtii* (zob. podrozdz. 2.3.1) z genomem *V. carteri*. Pierwszy z tych gatunków jest jednokomórkowy i wytwarza gamety podobne morfologicznie, natomiast drugi tworzy kolonie z komórkami wyspecjalizowanymi do funkcji wegetatywnych i generatywnych, przy czym komórki płciowe zróżnicowane są na nieruchliwą komórkę jajową i mobilny plemnik. Pomimo tych różnic, oba gatunki rozdzieliły się stosunkowo niedawno (około 50 milionów lat temu) i wykazują duże podobieństwo sekwencji kodujących. Dlatego istotne będzie wykrycie zmian genomowych, które doprowadziły do tak dużego zróżnicowania fenotypowego oraz specjalizacji komórkowej i płciowej pomiędzy *C. reinhardtii* i *carteri*.

Chociaż poznano już sekwencję genomu makronukleusa u orzęska *T. thermophila* (zob. podrozdz. 2.4.1), to jednak jest on sekwencjonowany w jego różnych szczepach. Pozwoli to na poznanie zróżnicowania molekularnego makronukleusa oraz cech charakterystycznych dla poszczególnych szczepów, hodowanych w różnych warunkach m.in. stresowych. Uzyskane dzięki temu informacje będą pomocne w licznych zastosowaniach ekotoksykologicznych tego orzęska. Poważnym wyzwaniem jest zsekwencjonowanie mikronukleusa ze względu na dużą zawartość w nim sekwencji powtórzonych. Jednak poznanie jego sekwencji da pełny obraz struktury i funkcji całego genomu jądrowego u orzęsków.

Zsekwencjonowanie genomów dwóch przedstawicieli śluzorośli: *Physarum polycephalum* i *Dictyostelium purpureum* pomoże z pewnością poznać podłoże genetyczne wielu interesujących procesów komórkowych. *P. polycephalum* jest organizmem wzorcowym w badaniach ruchu komórki, natomiast *D. purpureum*, podobnie jak *D. discoideum* o poznanym już genomie (patrz podrozdz. 2.4.2), służy jako model do badania interakcji międzykomórkowych.

3.4. Genomy użyteczne w biotechnologii

Wiele genomów sekwencjonowanych jest ze względu na ich potencjalne wykorzystanie do celów biotechnologicznych. Zielenica *Botryococcus braunii* wytwarza węglowodory (aż do 85% suchej masy), które mogą być przekształcane na biopaliwo (125,126). Niestety, dotychczas bardzo niewiele wiadomo o genach (i kodowanych przez nie enzymach), które związane są z syntezą tych związków. Na uwagę zasługuje również skoczkowiec *Piromyces* sp. E2 (grupa Chytridiomycota) (33). Bytuje on w układzie pokarmowym ssaków roślinożernych rozkładając pokarm roślinny, a jego mitochondria uległy przekształconiu w hydrogenosomy produkujące wodór i ATP (zob. podrozdz. 2.1.1). Niektóre geny tego gatunku są już wykorzystywane w biotechnologii poprzez ich ekspresjonowanie w komórkach drożdży *S. cerevisiae* w celu produkcji etanolu (127,128). Poznanie genomu *Piromyces* pozwoli także na jego szersze wykorzystanie do produkcji wodoru m.in. jako źródła energii. Wspomniano już, że sekwencjonowane są również genomy różnych szczepów zielenicy *C. reinhardtii* (jeden szczep został już zsekwencjonowany, zob. podrozdz. 2.3.1). Planuje się, aby gatunek ten wykorzystywać zarówno w oczyszczaniu gleby z metali ciężkich, jak i do wytwarzania wodoru, stosowanego do produkcji biopaliw (129).

Genom krasnorostu *Chondrus crispus* (chrząstnica kędzierzawa lub mech irlandzki) jest sekwencjonowany ze względu na jego duże znaczenie w produkcji karagenu zawierającego polisacharydy oraz niewielkie ilości jodu i bromu (130,131). Związek ten stosowany jest w wielu gałęziach przemysłu (np. piwowarskiego, spożywczego i kosmetycznego) oraz w lecznictwie do wytwarzania galaret i żeli. Ponadto odznacza się on dużą aktywnością biologiczną będąc m.in. immunostymulatorem. Duże nadzieje wiąże się również z poznaniem genomu innego krasnorostu – *Porphyra purpurea*. Planuje się zwiększenie jego produkcji, ponieważ jest on powszechnie wykorzystywany w krajach azjatyckich w celach konsumpcyjnych, np. do zawijania sushi (132). Interesujące będzie także poznanie zestawu genów wspomnianego już termofilnego i kwasolubnego krasnorostu *G. sulphuraria* (zob. podrozdz. 2.3.2), które umożliwiają jemu wzrost na ponad 50 różnych źródłach węgla, w tym na rzadkich węglowodanach i alkoholach cukrowych (133). Geny te mogą znaleźć zastosowanie w różnych działach biotechnologii.

Należy jeszcze wspomnieć o znaczeniu biotechnologicznym orzęska *T. thermophila*. Poznanie genomu jego wielu szczepów będzie bardzo pomocne w konstrukcji linii komórkowych ekspresjonujących różne białka, a szczególnie celulazy (konieczne w produkcji biopaliw) oraz białka chelatujące metale (istotne w biorekultywacji ścieków przemysłowych skażonych toksycznymi metalami ciężkimi).

3.5. Genomy ważne ewolucyjnie

Istotny – z punktu widzenia ewolucyjnego – jest projekt dotyczący powstania wielokomórkowości u zwierząt (OMP, ang. *Origins of Multicellularity Project*) (134). Celem tego przedsięwzięcia jest wyjaśnienie zjawiska wielokomórkowości w swietle genomów jednokomórkowych protistów, które są blisko spokrewnione z organizmami tkankowymi. Sekwencjonowaniu poddano w związku z tym genomy gatunków należących do: Apusozoa, Mesomycetozoa, Nucleariida i Choanoflagellatea (dotychczas znany jest kompletny genom tylko jednego gatunku należącego do Choanoflagellatea, zob. podrozdz. 2.4.3). Poznanie genomu *Amastigomonas* sp., z grupy Apusoza, pomoże rozwiązać niejasne związki filogenetyczne tej grupy z innymi eukariotami (19,135-137).

Interesujące wnioski ewolucyjne dotyczące ewolucji roślin wyższych mogą zostać sformułowane podczas analizy genomu *Chara vulgaris* należącej do ramienic – grupy glonów najbliżej spokrewnionej z roślinami lądowymi (138). Planuje się poznanie genomów jądrowych kryptofitów (np. *Guillardia theta*) i chlorarachniofitów (np. *Bigelowiella natans*), które nabyły plastydy na drodze wtórnej endosymbiozy odpowiednio od krasnorostów i zielenic (42,95,96). Genomy nukleomorfów (zredukowane jądra eukariotycznych dawców plastydów) zostały już poznane (zob. podrozdz. 2.3.3), a zsekwencjonowanie genomów jądrowych tych glonów, razem z genomami przedstawicieli innych grup zawierających plastydy złożone (np. haptofity i bruzdnice), da pełniejszy obraz zjawisk zachodzących w genomach podczas wtórnych endosymbioz (np. transfer genów z genomów endosymbionta do genomu jądrowego gospodarza) (36,39, 139,140). Dostępność kompletnych sekwencji genomów kryptofitów, a także haptofitów, dostarczy dużej liczby sekwencji kodujących, niezbędnych do przeprowadzenia rzetelnych analiz filogenetycznych mających na celu wyjaśnienie wciąż niejasnych powiązań Hacrobia z pozostałymi supergrupami eukariotów (22,24-28).

Posiadanie kompletnych sekwencji genomów wielu gatunków, reprezentujących różne linie ewolucyjne w obrębie domeny Eukaryota, umożliwi tworzenie drzew filogenetycznych na bardzo dużych zbiorach sekwencji, dzięki czemu uzyskiwane wyniki będą bardziej wiarygodne. Dostępność tak obszernego źródła informacji oferuje większy wybór sekwencji do analiz filogenetycznych, wśród których znacznie łatwiej jest znaleźć sekwencje konserwatywne i ortologiczne, które są najbardziej adekwatne do tego typu badań. Kompletne sekwencje wielu genomów umożliwią również odrzucenie genów, które podlegały duplikacjom lub poziomemu transferowi i zaburzają analizy filogenetyczne. Jedynie zakrojone na szeroką skalę projekty sekwencjonowania nowych genomów, w połączeniu z nowymi metodami filogenomicznymi (141-143), pozwolą na dokładniejsze wyjaśnienie związków filogenetycznych pomiędzy różnymi liniami ewolucyjnymi w obrębie Eukaryota. Konieczne jest wobec tego poznanie sekwencji genomów głównych organizmów, tj. reprezentantów wszystkich ważniejszych linii filogenetycznych.

Dobrym przykładem na znaczenie projektów genomowych w odtwarzaniu prawdziwych powiązań filogenetycznych jest mikrosporidium *E. cuniculi*: jednokomórkowy, ob-

ligatoryjny pasożyt wewnątrzkomórkowy ssaków, w tym człowieka (zob. podrozdz. 2.1.1). Początkowo uważano, że gatunek ten, podobnie jak inne mikrosporidie, nigdy nie posiadał mitochondrium (144). Na wczesnych drzewach filogenetycznych, opartych głównie na małej i dużej podjednostce jądrowego rDNA, pasożyty te grupowały się z innymi – przypuszczalnie również niemitochondrialnymi – eukariotami zaliczanymi obecnie do Excavata i Amoebozoa (144-147). Stało się to podstawą do wyodrębnienia nadkrólestwa Archezoa (144). Zakładano, że eukarioty te oddzieliły się jeszcze przed nabyciem mitochondriów i reprezentują fazę przedmitochondrialną w ewolucji Eukaryota. Do ostatecznego odrzucenia tej hipotezy przyczyniły się filogenezy oparte na danych pochodzących z kompletnie zsekwencjonowanego genomu *E. cuniculi* (44,148). Dzięki kompletnej sekwencji jego genomu wybrano prawie 100 genów o najwolniejszym tempie dywergencji i uzyskano filogenezy jednoznacznie grupujące te geny z genami grzybów, a nie z genami ekskawatów czy ameb. Wcześniejsze analizy filogenetyczne, na których mikrosporidie odgałęziały się u podstawy drzewa Eukaryota, konstruowano w oparciu na szybciej ewoluujących sekwencjach, które są szczególnie podatne na tzw. efekt przyciągania długich gałęzi i dają fałszywy obraz związków filogenetycznych (149-151). Ponadto u wszystkich przedstawicieli Archezoa, włączając Microsporidia, odkryto zmodyfikowane mitochondria w postaci mitosomów i hydrogenosomów (zob. podrozdz. 2.1.1). Na podstawie tych danych jasno pokazano, że przodek wszystkich eukariotów posiadał mitochondrium (152,153).

3.6. Niezwykle duże genomy bruzdnic i ameb

Sekwencjonowanie genomów niektórych protistów jest dużym wyzwaniem ze względu na ich wielkość. Szczególnie duże genomy posiadają bruzdnice (Dinoflagellata) – jednokomórkowe glony odgrywające istotną rolę ekologiczną (tab. 1). Typowy genom bruzdnic jest większy niż dotychczas zsekwencjonowane genomy eukariotów (154). Najmniejszy genom *Symbiodinium* spp. ma wielkość 2,9 Gb, a największy *Prorocentrum micans* – aż 244,5 Gb. Problemem jest nie tyle sama wielkość genomu, co duża frakcja DNA powtórzonego oraz wysoka liczba (nawet do 5000) zduplikowanych genów o bardzo podobnych sekwencjach (155), co utrudnia składanie krótkich, zsekwencjonowanych fragmentów genomu. Szacowana liczba genów kodujących białka w najmniejszym (*Symbiodinium* spp.) i największym (*P. micans*) genomie bruzdnic wynosi odpowiednio 38,2 i 87,7 tys. (155), czyli więcej niż u wysoko uorganizowanych organizmów tkankowych. Jednak, biorąc pod uwagę liczne duplikacje obserwowane w genomach bruzdnic, można sądzić, że tak duża liczba genów odzwierciedla raczej nadmiar informacji niż funkcjonalne zróżnicowanie i złożoność kodowanych proteomów. Ponieważ zduplikowane geny są często przekształcane w pseudogeny i ostatecznie tracone, albo też ulegają dywergencji funkcjonalnej, utrzymywanie tak dużej liczby kopii genów w genomie bruzdnic może jednak wskazywać na pewne ich związki z adaptacją do szerokiego zakresu strategii tro-

ficznych realizowanych w różnych niszach ekologicznych. Dlatego poznanie przynajmniej części ich genomów może odsłonić nieznane jeszcze zależności pomiędzy strukturą genomu a zdolnościami adaptacyjnymi organizmów.

Znaną od dawna osobliwą cechą bruzdnic jest brak nukleosomów i histonów. U protistów tych za kondensację DNA odpowiadają białka HLP (ang. Histone-Like Protein), które pod względem struktury drugorzędowej przypominają białka bakteryjne: HLP oraz HU (156-158). Podział jądra jest nietypowy, np. nie zanika błona jądrowa, a ich chromosomy są stale skondensowane, nawet w interfazie (159,160). Poznanie genomu bruzdnic może pomóc w zrozumieniu jak zorganizowana, kopiowana i odczytywana jest informacja genetyczna u tych organizmów.

Jeszcze większym przedsięwzięciem może być sekwencjonowanie genomów niektórych ameb (Amoebozoa). Wielkości największych genomów zostały oszacowane na: 290 Gb – u *Amoeba proteus*, 670 Gb – u *Polychaos dubium* (= *A. dubia*) i 1370 Gb – u *Chaos chaos* (161). Należy jednak do tych szacunków podchodzić z ostrożnością, ponieważ były one oparte na niezbyt dokładnych analizach biochemicznych ilości całkowitego DNA komórkowego, który najprawdopodobniej zawierał znaczne ilości DNA mitochondrialnego. Inne badania przeprowadzone na *A. proteus* różnią się o rząd wielkości w porównaniu do wcześniejszych obliczeń (162). W badaniach tych wykazano również, że genom *A. proteus* składa się z 500 do 1000 małych chromosomów i najprawdopodobniej jest poliploidalny.

4. Podsumowanie

Protisty stanowią dominujące i bardzo zróżnicowane grupy na drzewie życia eukariotów. Poznanie ich genomów jest istotne dla jego kompletnego zrekonstruowania i zrozumienia podstawowych etapów w ewolucji eukariotów, takich jak: 1) powstanie jądra komórkowego, systemu wewnątrzbłonowego i cytoszkieletu, 2) nabycie i ewolucja mitochondriów oraz plastydów, 3) ewolucja płciowości, 4) powstanie wielokomórkowości oraz 5) ewolucja szlaków rozwojowych. Należy tu również wspomnieć o bardziej generalnych zagadnieniach, czyli naturze i mechanizmach powstawania nowych gatunków oraz przyczynach zdarzeń makroewolucyjnych, które prowadzą do wyodrębniania się nowych linii filogenetycznych. Genomy organizmów modelowych wzbogacą znacznie naszą wiedzę o strukturze i funkcjonowaniu genomów, ekspresji poszczególnych genów oraz wielu innych zjawiskach zachodzących na poziomie molekularnym. Szczególnie istotne jest poznanie genomów pasożytów tak, aby dokładnie zrozumieć mechanizmy ich infekcji, a następnie opracować skuteczne sposoby zapobiegania im i leczenia wywoływanych chorób. Dzięki poznaniu genomów protistów istotnych ekologicznie lepiej zrozumiemy skomplikowane procesy zachodzące w różnych ekosystemach. Nowe sekwencje genomów protistów pozwolą również odkryć nieznane dotąd białka oraz szlaki i cykle metaboliczne, które mogą w istotny sposób wpłynąć na rozwój biotechnologii. Odczyta-

nie sekwencji genomów jest jednak dopiero początkiem w zrozumieniu całej złożoności eukariotów i przełożeniu tej wiedzy na praktyczne zastosowania.

Literatura

1. Fleischmann R. D., Adams M. D., White O., Clayton R. A., Kirkness E. F., Kerlavage A. R., Bult C. J., Tomb J., Dougherty B. A., Merrick J. M., et al., (1995), Science, 269, 496-512.
2. Mackiewicz P., Zakrzewska-Czerwińska J., Cebrat S., (2005), Biotechnologia, 3, 7-21.
3. Shendure J., Mitra R. D., Varma C., Church G. M., (2004), Nat. Rev. Genet., 5, 335-344.
4. Metzker M. L., (2005), Genome Res., 15, 1767-1776.
5. Metzker M. L., (2010), Nat. Rev. Genet.,11, 31-46.
6. Liolios K., Chen I. M., Mavromatis K., Tavernarakis N., Hugenholtz P., Markowitz V. M., Kyrpides N. C., (2010), Nucleic Acids Res., 38 (Database issue), D346-D354.
7. Goffeau A., Barrell B. G., Bussey H., Davis R. W., Dujon B., Feldmann H., Galibert F., Hoheisel J. D., Jacq C., Johnston M., et al., (1996), Science, 274, 563-567.
8. *C. elegans* Sequencing Consortium, (1998), Science, 282, 2012-2018.
9. Adams M. D., Celniker S. E., Holt R. A., Evans C. A., Gocayne J. D., Amanatides P. G., Scherer S. E., Li P. W., Hoskins R. A., Galle R. F., et al., (2000), Science, 287, 2185-2195.
10. *Arabidopsis* Genome Initiative, (2000), Nature, 408, 796-815.
11. Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., et al., (2001), Nature, 409, 860-921.
12. Venter J. C., Adams M. D., Myers E. W., Li P. W., Mural R. J., Sutton G. G., Smith H. O., Yandell M., Evans C. A., Holt R. A., Gocayne J. D., et al., (2001), Science, 291, 1304-1351.
13. Margulis L., Schwartz K. V., (1988), *Five Kingdoms. An Illustrated Guide to the Phyla of Life on Earth*, W. H. Freeman and Company, New York.
14. Cavalier-Smith T., (1998), Biol. Rev., 73, 203-266.
15. Scamardella J. M., (1999), Inter. Microbiol., 2, 207-221.
16. Cavalier-Smith T., (1981), Biosystems, 14, 461-481.
17. Cavalier-Smith T., (2002), Int. J. Syst. Evol. Microbiol., 52, 297-354.
18. Burki F., Pawlowski J., (2006), Mol. Biol. Evol., 23, 1922-1930.
19. Moreira D., von der Heyden S., Bass D., Lopez-Garcia P., Chao E., Cavalier-Smith T., (2007), Mol. Phylogenet. Evol., 44, 255-266.
20. Patron N. J., Inagaki Y., Keeling P. J., (2007), Curr. Biol., 17, 887-891.
21. Cavalier-Smith T., (2010), Biol. Lett., 6, 342-345.
22. Hackett J. D., Yoon H. S., Li S., Reyes-Prieto A., Rummele S. E., Bhattacharya D., (2007), Mol. Biol. Evol., 24, 1702-1713.
23. Rodriguez-Ezpeleta N., Brinkmann H., Burger G., Roger A. J., Gray M. W., Philippe H., Lang B. F., (2007), Curr. Biol., 17, 1420-1425.
24. Shalchian-Tabrizi K., Minge M. A., Espelund M., Orr R., Ruden T., Jakobsen K. S., Cavalier-Smith T., (2008), PLoS One, 3, e2098.
25. Burki F., Shalchian-Tabrizi K., Pawlowski J., (2008), Biol. Lett., 4, 366-369.
26. Burki F., Inagaki Y., Bråte J., Archibald J. M., Keeling P. J., Cavalier-Smith T., Sakaguchi M., Hashimoto T., Horak A., Kumar S., et al., (2009), Gen. Biol. Evol., 27, 231-238.
27. Minge M. A., Silberman J. D., Orr R. J., Cavalier-Smith T., Shalchian-Tabrizi K., Burki F., Skjaeveland A., Jakobsen K. S., (2009), Proc. Biol. Sci., 276, 597-604.
28. Okamoto N., Chantangsi C., Horák A., Leander B. S., Keeling P. J., (2009), PLoS One, 4, e7080.
29. Hampl V., Hug L., Leigh J. W., Dacks J. B., Lang B. F., Simpson A. G., Roger A. J., et al., (2009), Proc. Natl. Acad. Sci. USA., 106, 3859-3864.
30. Baldauf S. L., (2003), Science, 300, 1703-1706.

31. Cavalier-Smith T., (2004), Proc. Biol. Sci., 271, 1251-1262.
32. Simpson A. G., Roger A. J., (2004), Curr. Biol., 14, R693-R696.
33. Adl S. M., Simpson A. G., Farmer M. A., Andersen R. A., Anderson O. R., Barta J. R., Bowser S. S., Brugerolle G., Fensome R. A., Fredericq S., et al., (2005), J. Eukaryot. Microbiol., 52, 399-451.
34. Keeling P. J., Burger G., Durnford D. G., Lang B. F., Lee R. W., Pearlman R. E., Roger A. J., Gray M. W., (2005), Trends Ecol. Evol., 20, 670-676.
35. Baldauf S. L., (2008), J. Syst. Evol., 46, 263-273.
36. Timmis J. N., Ayliffe M. A., Huang C. Y., Martin W., (2004), Nat. Rev. Genet., 5, 123-135.
37. Cavalier-Smith T., (2005), Ann. Bot., 95, 147-175.
38. Cavalier-Smith T., (2006), Proc. Biol. Sci., 273, 1943-1952.
39. Keeling P. J., Palmer J. D., (2008), Nat. Rev. Genet., 9, 605-618.
40. Martin W., Koonin E. V., (2006), Nature, 440, 41-45.
41. Bodył A., Mackiewicz P., Stiller J. W., (2009), Bioessays, 31, 1219-1232.
42. Moore C. E., Archibald J. M., (2009), Annu. Rev. Genet., 43, 251-264.
43. McGrath C. L., Zufall R. A.., Katz L. A., (2006), in: *Genomics and Evolution of Microbial Life*, Eds. Katz L. A., Bhattacharya D., Oxford University Press, New York.
44. Katinka M. D., Duprat S., Cornillot E., Méténier G., Thomarat F., Prensier G., Barbe V., Peyretaillade E., Brottier P., Wincker P., et al., (2001), Nature, 414, 450-453.
45. Carlton J. M., Hirt R. P., Silva J. C., Delcher A. L., Schatz M., Zhao Q., Wortman J. R., Bidwell S. L., Alsmark U. C., Besteiro S., et al., (2007), Science, 315, 207-212.
46. Morrison H. G., McArthur A. G., Gillin F. D., Aley S. B., Adam R. D., Olsen G. J., Best A. A., Cande W. Z., Chen F., Cipriano M. J., et al., (2007), Science, 317, 1921-1926.
47. Franzén O., Jerlström-Hultqvist J., Castro E., Sherwood E., Ankarklev J., Reiner D. S., Palm D., Andersson J. O., Andersson B., Svärd S. G., (2009), PLoS Pathog., 5, e1000560.
48. Loftus B., Anderson I., Davies R., Alsmark U. C., Samuelson J., Amedeo P., Roncaglia P., Berriman M., Hirt R. P., Mann B. J., et al., (2005), Nature, 433, 865-868.
49. Caler E., Lorenzi H., (2010), in*: Anaerobic Parasitic Protozoa: Genomics and Molecular Biology*, Eds. Clark C. G., Johnson P. J., Adam R. D., Caister Academic Press Ltd., Norwich, GB.
50. van der Giezen M., (2009), J. Eukaryot. Microbiol., 56, 221-231.
51. Shiflett A. M., Johnson P. J., (2010), Annu. Rev. Microbiol., (w druku).
52. Lau A. O., (2009), Mol. Biochem. Parasitol., 164, 1-8.
53. Templeton T. J., Iyer L. M., Anantharaman V., Enomoto S., Abrahante J. E., Subramanian G. M., Hoffman S. L., Abrahamsen M. S., Aravind L., (2004), Genome Res., 14, 1686-1695.
54. Gardner M. J., Hall N., Fung E., White O., Berriman M., Hyman R. W., Carlton J. M., Pain A., Nelson K. E., Bowman S., et al., (2002), Nature, 419, 498-511.
55. Carlton J. M., Adams J. H., Silva J. C., Bidwell S. L., Lorenzi H., Caler E., Crabtree J., Angiuoli S. V., Merino E. F., Amedeo P., et al., (2008), Nature, 455, 757-763.
56. Pain A., Böhme U., Berry A. E., Mungall K., Finn R. D., Jackson A. P., Mourier T., Mistry J., Pasini E. M., Aslett M. A., et al., (2008), Nature, 455, 799-803.
57. Carlton J. M., Angiuoli S. V., Suh B. B., Kooij T. W., Pertea M., Silva J. C., Ermolaeva M. D., Allen J. E., Selengut J. D., Koo H. L., et al., (2002), Nature, 419, 512-519.
58. Xu P., Widmer G., Wang Y., Ozaki L. S., Alves J. M., Serrano M. G., Puiu D., Manque P., Akiyoshi D., Mackey A. J., et al., (2004), Nature, 431, 1107-1112.
59. Abrahamsen M. S., Templeton T. J., Enomoto S., Abrahante J. E., Zhu G., Lancto C. A., Deng M., Liu C., Widmer G., Tzipori S., et al., (2004), Science, 304, 441-445.
60. Pain A., Renauld H., Berriman M., Murphy L., Yeats C. A., Weir W., Kerhornou A., Aslett M., Bishop R., Bouchier C., et al., (2005), Science, 309, 131-133.
61. Gardner M. J., Bishop R., Shah T., de Villiers E. P., Carlton J. M., Hall N., Ren Q., Paulsen I. T., Pain A., Berriman M., et al., (2005), Science, 309, 134-137.
62. Brayton K. A., Lau A. O., Herndon D. R., Hannick L., Kappmeyer L. S., Berens S. J., Bidwell S. L., Brown W. C., Crabtree J., Fadrosh D., et al., (2007), PLoS Pathog., 3, e148.
63. Endo T., Ikeo K., Gojobori T., (1996), Mol. Biol. Evol., 13, 685-690.

64. Graham S. P., Pellé R., Honda Y., Mwangi D. M., Tonukari N. J., Yamage M., Glew E. J., de Villiers E. P., Shah T., Bishop R., et al., (2006), Proc. Natl. Acad. Sci. USA., 103, 3286-3291.
65. Dobbelaere D. A., Küenzi P., (2004), Curr. Opin. Immunol., 16, 524-530.
66. Seeber F., Soldati-Favre D., (2010), Int. Rev. Cell. Mol. Biol., 281, 161-228.
67. Lau A. O., McElwain T. F., Brayton K. A., Knowles D. P., Roalson E. H., (2009), Exp. Parasitol., 123, 236-243.
68. Janouskovec J., Horák A., Oborník M., Lukes J., Keeling P. J., (2010), Proc. Natl. Acad. Sci. USA., 107, 10949-10954.
69. Tonkin C. J., Kalanon M., McFadden G. I., (2008), Traffic, 9, 166-175.
70. Huang J., Mullapudi N., Lancto C. A., Scott, M., Abrahamsen, M. S., Kissinger J. C., (2004), Genome Biol., 5, R88.
71. El-Sayed N. M., Myler P. J., Blandin G., Berriman M., Crabtree J., Aggarwal G., Caler E., Renauld H., Worthey E. A., Hertz-Fowler C., et al., (2005), Science, 309, 404-409.
72. Berriman M., Ghedin E., Hertz-Fowler C., Blandin G., Renauld H., Bartholomeu D. C., Lennard N. J., Caler E., Hamlin N. E., Haas B., et al., (2005), Science, 309, 416-422.
73. El-Sayed N. M., Myler P. J., Bartholomeu D. C., Nilsson D., Aggarwal G., Tran A. N., Ghedin E., Worthey E. A., Delcher A. L., Blandin G., et al., (2005), Science, 309, 409-415.
74. Peacock C. S., Seeger K., Harris D., Murphy L., Ruiz J. C., Quail M. A., Peters N., Adlem E., Tivey A., Aslett M. (2007), Nat. Genet., 39, 839-847.
75. Ivens A. C., Peacock C. S., Worthey E. A., Murphy L., Aggarwal G., Berriman M., Sisk E., Rajandream M. A., Adlem E., Aert R., et al., (2005), Science, 309, 436-442.
76. El-Sayed N., Ghedin E., Song J., MacLeod A., Bringaud F., Larkin C., Wanless D., Peterson J., Hou L., Taylor S., et al., (2003), Nucleic Acids Res., 31, 4856-4863.
77. Hall N., Berriman M., Lennard N., Harris B., Hertz-Fowler C., Bart-Delabesse E., Gerrard C., Atkin R., Barron A., Bowman S., et al., (2003), Nucleic Acids Res., 31, 4864-4873.
78. Worthey E., Martinez-Calvillo S., Schnaufer A., Aggarwal G., Cawthra J., Fazelinia G., Fong C., Fu G., Hassebrock M., Hixson G., (2003), Nucleic Acids Res., 31, 4201-4210.
79. Opperdoes F. R., Michels P. A., (2007), Trends Parasitol., 23, 470-476.
80. Leander B. S., (2004), Trends Microbiol., 12, 251-258.
81. Bodył A., Mackiewicz P., Milanowski R., (2010), J. Parasitol., 96, 465-475.
82. Haas B. J., Kamoun S., Zody M. C., Jiang R. H., Handsaker R. E., Cano L. M., Grabherr M., Kodira C. D., Raffaele S., Torto-Alalibo T., (2009), Nature, 461, 393-398.
83. Tyler B. M., Tripathy S., Zhang X., Dehal P., Jiang R. H., Aerts A., Arredondo F. D., Baxter L., Bensasson D., Beynon J. L., (2006), Science, 313, 1261-1266.
84. Cavalier-Smith T., (1999), J. Eukaryot. Microbiol., 46, 347-366.
85. Keeling P. J., (2009), J. Eukaryot. Microbiol., 56, 1-8.
86. Stiller J. W., Huang J., Ding Q., Tian J., Goodwillie C., (2009), BMC Genomics, 10, 484.
87. Lewis L. A., McCourt R. M., (2004), Am. J. Bot., 91, 1535-1556.
88. Behrenfeld M. J., Falkowski P. G., (1997), Limnol. Oceanogr., 42, 1-20.
89. Worden A. Z., Nolan J. K., Palenik B., (2004), Limnol. Oceanogr., 49, 168-179.
90. Palenik B., Grimwood J., Aerts A., Rouzé P., Salamov A., Putnam N., Dupont C., Jorgensen R., Derelle E., Rombauts S., (2007), Proc. Natl. Acad. Sci. USA., 104, 7705-7710.
91. Derelle E., Ferraz C., Rombauts S., Rouzé P., Worden A. Z., Robbens S., Partensky F., Degroeve S., Echeynié S., Cooke R., et al., (2006), Proc. Natl. Acad. Sci. USA., 103, 11647-11652.
92. Worden A. Z., Lee J. H., Mock T., Rouzé P., Simmons M. P., Aerts A. L., Allen A. E., Cuvelier M. L., Derelle E., Everett M. V., (2009), Science, 324, 268-272.
93. Bowler C., Allen A. E., Badger J. H., Grimwood J., Jabbari K., Kuo A., Maheswari U., Martens C., Maumus F., Otillar R. P., et al., (2008), Nature, 456, 239-244.
94. Armbrust E. V., Berges J. A., Bowler C., Green B. R., Martinez D., Putnam N. H., Zhou S., Allen A. E., Apt K. E., Bechner M., et al., (2004), Science, 306, 79-86.
95. Archibald J. M., (2009), Curr. Biol., 19, R81-R88.
96. Vesteg M., Vacula R., Krajcovic J., (2009), Folia Microbiol. (Praha), 54, 303-321.
97. Parkinson J., Gordon R., (1999), Trends Biotechnol., 17, 190-196.

98. Gordon R., Parkinson J., (2005), J. Nanosci. Nanotechnol., 5, 35-40.
99. Oudot-Le Secq M. P., Grimwood J., Shapiro H., Armbrust E. V., Bowler C., Green B. R., (2007), Mol. Genet. Genomics, 277, 427-439.
100. Bolte K., Bullmann L., Hempel F., Bozarth A., Zauner S., Maier U. G., (2009), J. Eukaryot. Microbiol., 56, 9-15.
101. Moustafa A., Beszteri B., Maier U. G., Bowler C., Valentin K., Bhattacharya D., (2009), Science, 324, 1724-1726.
102. Merchant S. S., Prochnik S. E., Vallon O., Harris E. H., Karpowicz S. J., Witman G. B., Terry A., Salamov A., Fritz-Laylin L. K., Maréchal-Drouard L., et al., (2007), Science, 318, 245-250.
103. Matsuzaki M., Misumi O., Shin-I T., Maruyama S., Takahara M., Miyagishima S. Y., Mori T., Nishida K., Yagisawa F., Nishida K., et al., (2004), Nature, 428, 653-657.
104. Nozaki H., Takano H., Misumi O., Terasawa K., Matsuzaki M., Maruyama S., Nishida K., Yagisawa F., Yoshida Y., Fujiwara T., (2007), BMC Biol., 5, 28.
105. Cavalier-Smith T., (2000), Trends Plant. Sci., 5, 174-182.
106. Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L. T., Wu X., Reith M., Cavalier-Smith T., Maier U. G., (2001), Nature, 410, 1091-1096.
107. Lane C. E., van den Heuvel K., Kozera C., Curtis B. A., Parsons B. J., Bowman S., Archibald J. M., (2007), Proc. Natl. Acad. Sci. USA., 104, 19908-19913.
108. Gilson P. R., Su V., Slamovits C. H., Reith M. E., Keeling P. J., McFadden G. I., (2006), Proc. Natl. Acad. Sci. USA., 103, 9566-95671.
109. Aury J. M., Jaillon O., Duret L., Noel B., Jubin C., Porcel B. M., Ségurens B., Daubin V., Anthouard V., Aiach N., et al., (2006), Nature, 444, 171-178.
110. Eisen J. A., Coyne R. S., Wu M., Wu D., Thiagarajan M., Wortman J. R., Badger J. H., Ren Q., Amedeo P., Jones K. M., et al., (2006), PLoS Biol., 4, e286.
111. Reyes-Prieto A., Moustafa A., Bhattacharya D., (2008), Curr. Biol., 18, 956-962.
112. Bodył A., Stiller J. W., Mackiewicz P., (2009), Trends Ecol. Evol., 24, 119-121.
113. Eichinger L., Pachebat J. A., Glöckner G., Rajandream M. A., Sucgang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q., et al., (2005), Nature, 435, 43-57.
114. King N., Westbrook M. J., Young S. L., Kuo A., Abedin M., Chapman J., Fairclough S., Hellsten U., Isogai Y., Letunic I., (2008), Nature, 451, 783-788.
115. Sun G., Yang Z., Ishwar A., Huang J., (2010), Mol. Biol. Evol., (w druku).
116. Vaulot D., Eikrem W., Viprey M., Moreau H., (2008), FEMS Microbiol. Rev., 32, 795-820.
117. Falkowski P. G., Barber R. T., Smetacek V., (1998), Science, 281, 200-206.
118. Field C. B., Behrenfeld M. J., Randerson J. T., Falkowski P., (1998), Science, 281, 237-240.
119. Nelson D. M., Treguer P., Brzezinski M. A., Leynaert A., Queguiner B., (1995), Global Biogeochem. Cycles, 9, 359-372.
120. Hefu Y., Kirst G. O., (1997), Polar Biol., 18, 402-409.
121. Stefels J., van Boekel W. H. M., (1993), Mar. Ecol. Prog. Ser., 97, 11-18.
122. Sellner K. G., Doucette G. J., Kirkpatrick G. J., (2003), J. Ind. Microbiol. Biotechnol., 30, 383-406.
123. Douglas A. E., (2003), Mar. Pollut. Bull., 46, 385-392.
124. Berkelmans R., De'ath G., Kininmonth S., Skirving W. J., (2004), Coral Reefs, 23, 74-83.
125. Hillen L. W., Pollard G., Wake L. V., White N., (1982), Biotechnol. Bioeng., 24, 193-205.
126. Metzger P., Largeau C., (2005), Appl. Microbiol. Biotechnol., 66, 486-496.
127. Kuyper M., Harhangi H. R., Stave A. K., Winkler A. A., Jetten M. S., de Laat W. T., den Ridder J. J., Op den Camp H. J., van Dijken J. P., Pronk J. T., (2003), FEMS Yeast Res., 4, 69-78.
128. van Maris A. J., Winkler A. A., Kuyper M., de Laat W. T., van Dijken J. P., Pronk J. T., (2007), Adv. Biochem. Eng. Biotechnol., 108, 179-204.
129. Rupprecht J., (2009), J. Biotechnol., 142, 10-20.
130. Chopin T., (1986), Can. Tech. Rep. Fish Aquat. Sci., 1514, 1-69.
131. Chopin T., Gallant T., Davison I., (1995), J. Phycol., 31, 283-293.
132. Mumford T. F., Miura A., (1988), in: *Algae and Human Affairs*, Eds. Lembi C. A., Waaland J. R., 87-117, Cambridge University Press, Cambridge.

133. Barbier G., Oesterhelt C., Larson M. D., Halgren R. G., Wilkerson C., Garavito R. M., Benning C., Weber A. P., (2005), Plant Physiol., 137, 460-474.
134. Ruiz-Trillo I., Burger G., Holland P. W., King N., Lang B. F., Roger A. J., Gray M. W., (2007), Trends Genet., 23, 113-118.
135. Cavalier-Smith T., Chao E. E., (1995), Proc. R. Soc. Lond. B, 261, 1-6.
136. Stechmann A., Cavalier-Smith T., (2003), Curr. Biol., 13, R665-R666.
137. Kim E., Simpson A. G., Graham L. E., (2006), Mol. Biol. Evol., 23, 2455-2466.
138. McCourt R. M., Delwiche C. F., Karol K. G., (2004), Trends Ecol. Evol., 19, 661-666.
139. Reyes-Prieto A., Weber A. P., Bhattacharya D., (2007), Annu. Rev. Genet., 41, 147-168.
140. Keeling P. J., (2009b), Methods Mol. Biol., 532, 501-515.
141. Delsuc F., Brinkmann H., Philippe H., (2005), Nat. Rev. Genet., 6, 361-375.
142. Ciccarelli F. D., Doerks T., von Mering C., Creevey C. J., Snel B., Bork P., (2006), Science, 311, 1283-1287, Erratum in: Science, (2006), 312, 697.
143. Leigh J. W., Susko E., Baumgartner M., Roger A. J., (2008), Syst. Biol., 57, 104-115.
144. Cavalier-Smith T., (1993), Microbiol. Rev., 57, 953-994.
145. Leipe D., Gunderson J. H., Nerad T. A., Sogin M. L., (1993), Mol. Biochem. Parasitol., 59, 41-48.
146. Vossbrinck C. R., Maddox J. V., Friedman S., Debrunner-Vossbrinck B. A., Woese C. R., (1987), Nature, 326, 411-414.
147. Sogin M. L., Gunderson J., Elwood H., Alonso R., Peattie D., (1989), Science, 243, 75-77.
148. Thomarat F., Vivares C. P., Gouy M., (2004), J. Mol. Evol., 59, 780-791.
149. Stiller J. W., Hall B. D., (1999), Mol. Biol. Evol., 16, 1270-1279.
150. Philippe H., (2000), Protist, 151, 307-316.
151. Brinkmann H., van der Giezen M., Zhou Y., Poncelin de Raucourt G., Philippe H., (2005), Syst. Biol., 54, 743-757.
152. Martin W., Müller M., (1998), Nature, 392, 37-41.
153. Embley T. M., Martin W., (2006), Nature, 440, 623-630.
154. Hackett J. D., Anderson D. M., Erdner D. L., Bhattacharya D., (2004), Am. J. Bot., 91, 1523-1534.
155. Hou Y., Lin S., (2009), PLoS One, 4, e6978.
156. Rizzo P. J., Noodén L. D., (1972), Science, 176, 796-797.
157. Wong J. T., New D. C., Wong J. C., Hung V. K., (2003), Eukaryot. Cell., 2, 646-650.
158. Hackett J. D., Scheetz T. E., Yoon H. S., Soares M. B., Bonaldo M. F., Casavant T. L., Bhattacharya D., (2005), BMC Genomics, 6, 80.
159. Livolant F., Bouligand Y., (1978), Chromosoma, 68, 21-44.
160. Costas E., Goyanes V., (2005), Genome Res., 109, 268-275.
161. Friz C. T., (1968), Comp. Biochem. Physiol., 26, 81-90.
162. Byers T. J., (1986), Int. Rev. Cytol., 99, 311-341.