Genome Analysis of Social Amoebae

Gernot Glöckner

Abstract Genomics is now an indispensible part of the biological sciences. Today a species description without genome information is incomplete. This chapter describes the current knowledge on the genome of the model species *Dictyostelium discoideum*. A comparison with other social amoebae genomes covering the whole breadth of this branch of evolution carves out driving forces of speciation and the common toolkit of all social amoebae. The vast evolutionary distance within this branch makes ortholog detection difficult. While the coding capacity of all social amoebae is largely conserved, species specific gene family expansions of proteins for environmental sensing, signaling, and secondary metabolites provide for diversification. The sequences of the functional chromosomal elements (telomeres and centromeres) are not conserved, rather they seem to have underwent severe modifications. Nucleosome patterns link the social amoebae to other, more sophisticated multicellular systems. Comparative curated databases make this wealth of genome information accessible and play an important role for the dissemination of the knowledge on this evolutionary branch.

1 Genomics

Genomics has revolutionized the biologist's view on taxonomy, species, and traits in the last few years. Molecular data provide an unambiguous data source for the analysis of these features and allow an unprecedented detailed view on the evolutionary history of organisms. Initially, only a handful of model organisms were selected for a full genome analysis since in the early years of genomics sequencing was expensive and standard methods for computational analysis were still in

G. Glöckner (✉)
Medical Faculty, Institute for Biochemistry I, University of Cologne,
D-50931 Cologne, Germany
e-mail: gernot.gloeckner@uni-koeln.de

M. Romeralo et al. (eds.), *Dictyostelids*,
DOI: 10.1007/978-3-642-38487-5_2,

development. Among this selection was *Dictyostelium discoideum*, a primary basic genetic model system for the analysis of development, signalling, and the cytoskeleton. It soon became evident however, that only one species as representative for whole eukaryote evolutionary branches is not enough to get insight into evolutionary events and trajectories in the eukaryote tree. To be able to discern species specific from branch specific inventions, genomes of several species must be deciphered and compared to each other. Thus, immediately after the completion of the *D. discoideum* genome in 2005 (Eichinger et al. 2005; Glöckner et al. 2002), several comparative genomics projects were launched. The previously calculated detailed phylogeny of the main groups of Dictyostelids (Schaap et al. 2006) enabled the selection of species from each of the four main groups of social amoebae for comparative purposes.

2 Nuclear Genomes of Social Amoebae

2.1 The Model Organism D. discoideum

Assembly of sequences ranging from only 100 bases to a little bit more than 1 kb normally yields only draft genomes, i.e. genomes represented by a large number of unordered contigs separated by gaps of unknown length. If higher order organizational features of a genome should be analyzed these contigs must be brought into their natural order and as many gaps as possible should be filled. For *D. discoideum* a genetic map of the genome already existed (Kuspa et al. 1992; Loomis et al. 1995) although routine crosses with sufficient progeny for linkage analysis are rarely successful (Francis 1998). This genetic map was based on mutants with a visible phenotype and therefore of low resolution. To get a marker dense map of the genome a physical mapping method [happy mapping (Dear and Cook 1993)], which does not rely on sexual reproduction, was employed with this species (Konfortov et al. 2000; Williams and Firtel 2000). Both the genetic and the "Happy" map generally agreed thus providing the basis for the reconstruction of whole chromosomes. The maps also showed that the 34 Mb of the genome are distributed over six chromosomes.

Dictyostelium discoideum has a genome highly biased towards A and T nucleotides. With a mean A/T content of 78 % (Table 1) it exhibits the second highest bias so far observed, only surpassed by *Plasmodium falciparum* (Gardner et al. 2002). Later comparisons between these highly biased genomes revealed, that they employ different strategies to achieve such high A/T contents. The A/T bias influences the codon usage more pronounced in *D. discoideum* than in *P. falciparum*, while *P. falciparum* extends its intergenic regions with long A/T stretches (Szafranski et al. 2005). Long stretches of A/T rich codons encoding the same amino acid (mainly asparagine) contribute significantly to the overall A/T content in *D. discoideum* and considerably alter also the amino acid sequence of

Table 1 Genome properties of social amoebae. Only completed genomes are shown

	DD	DP	DL	PP	DF
Contigs	226	1,213	54	52	33
Supercontigs	6	838	54	41	25
Total nucleotides (Mbp)	35	33	23	33	31
Average contig length (kbp)	155	27	433	634	1,064
Overall nucleotide frequency (A/T %)	77.6	75.4	70.2	68	66.2
Palindrome arm size (kb)	45	26	14	15	28
Mitochondrial genome size (kb)	55	52	47	48	56
Chromosome numbers	6	nd	8 (or more)	7	6
Repeat content(%)	~10	3.4	<1	<1	<1
Telomere repeat structure	Palindrome arm	Palindrome arm?	TAGGG + palindrome arm	TAAGGG	TTAGGG
Predicted coding sequences (CDS)	13,433	12,410	10,958	12,373	12,173
Average gene length	1,579	1,689	1,596	1,552	1,672
Gene density (CDS per Mb)	396	376	470	375	392
Nucleotide frequency in CDS (A/T %)	72.6	69.9	67.8	63.8	63.2
Predicted tRNAs	401	375	61	273	198

DD: *Dictyostelium discoideum*; DP: *D. purpureum* (both group 4); DL: *D. lacteum* (group 3); PP: *Polysphondilium pallidum* (group 2); DF: *D. fasciculatum* (group 1)

the encoded proteins. No functions so far could be assigned to those homopolymer runs and later genome wide comparisons revealed that the position of these runs within a protein is not conserved (Sucgang et al. 2011). Thus, the only purpose of the homopolymer runs is seemingly to drive the A/T content of the genome to higher extremes. The A/T bias also prevented the usage of large insert size vectors like fosmids or bacterial artificial clones (BAC) since recombination in the bacterial host frequently affects the inserts resulting in deletions in the inserts of these clones (Glöckner et al. 2002).

The analysis of the coding capacity of the *D. discoideum* genome brought a big surprise: It encodes more than twice the protein coding genes than the unicellular budding yeast *Saccharomyces cerevisiae* (Goffeau et al. 1996). Presumably, the higher complexity of the cytoskeleton, the amoeboid mobility, sophisticated signalling systems, complex environments, and the developmental cycle require this higher amount of genetic information in *D. discoideum*.

The *D. discoideum* genome harbours comparably many transposable elements (TEs). With a total of 10 % of the whole genome this is by far more than in other unicellular species observed (Glöckner et al. 2001). The TEs comprise retrotransposons and DNA transposons. Additionally, some smaller repeated sequences were found. The TEs exhibit a non-random distribution in the genome. Some TEs (the TREs) use tRNAs as target sequences for integration and thus are mainly found (with a few exceptions) in a defined vicinity of tRNAs. Dictyostelium

intermediate repeat sequence (DIRS) elements on the other hand are restricted to the tips of the chromosomes where they form extensive clusters interspersed with some DNA elements. Three of these clusters could be reconstructed by making use of read pair and positional information on single nucleotide polymorphisms (SNPs) (Glöckner and Heidel 2009). These clusters are up to 300 kb long and, according to the analysis of the distribution of SNPs in DIRS elements, exchange these elements or parts of clusters among each other either by homologous recombination or transposition. It remains enigmatic why a genome densely packed with protein coding genes enables the spreading of TEs. Possibly, the restriction of TEs to certain regions of the genome ameliorates the adverse effects of transposition events. On the other hand increased mortality of strains harbouring TEs at uncommon positions could be responsible for clearance of these strains from the population (Winckler et al. 2002, 2005).

Previous genetic analyses had shown that many protein components indispensible for multicellular life forms in the Opisthokontha branch of eukaryote life (comprising Metazoa, Amoebozoa, Fungi, and some unicellular flagellated species) are also present in *D. discoideum*. A genetic analysis using mutants however is hampered by the fact that functions or parts of functions are often encoded redundantly in genomes. The complete genome then gave additional insight into the composition and extent of gene families. Several gene families, myosins, actins, rho related proteins, rasGEFs, kinesins, etc. (Joseph et al. 2008; Kollmar 2006; Kollmar and Glöckner 2003; Rivero et al. 2001; Wilkins et al. 2005), were analysed separately yielding information on basic eukaryote family sets and species specific expansions.

As a surprise came the finding that transcription factors are rare in this genome and that basic helix-turn-helix motifs are missing. This indicates that transcription either plays a minor role in regulation of the life and development of *D. discoideum* or that specifically adapted not yet identified transcriptional regulators are encoded in this genome.

2.2 Comparative Sequencing of Other Social Amoebae

The development of next generation sequencing (NGS) technologies over the last few years made genome analysis much cheaper. This enabled the sequencing of additional species within the social amoebae. To date at least a genome of one species in each of the four main branches of social amoebae is available in draft or complete state (Heidel et al. 2011; Sucgang et al. 2011; Felder et al. 2013). Here, genomes are considered complete, if all or nearly all contigs could be ordered and oriented or gap closure procedures were employed to close as many remaining gaps as possible. This is only possibly if additionally to the NGS data large insert size clones are used to order and orient contigs in their natural order along the chromosomes. Since the A/T bias is not so extreme in the other social amoebae groups it was possible to construct fosmid libraries enabling long range mapping

of contigs (Heidel et al. 2011). Table 1 lists relevant features of the available 5 genomes. Intriguingly, the total genome size ranges from 23 Mb for *D. lacteum* from group 3 to 34 Mb for *D. discoideum* in group 4. Part of the larger size of *D. discoideum* compared to all other social amoebae is attributable to the exceptional amount of TEs in *D. discoideum*, but another part is due to more extensive gene family expansions and larger intergenic regions. Thus, all genomes have nearly the same coding capacity, emphasizing the establishment of a life style early in evolution with only a few modifications required for the differentiation into lineages.

Orthologs between the species can be identified, if the similarity is not decayed below a certain threshold. Over the whole range of species in the social amoebae only roughly half of the genes have clear orthologs (Heidel et al. 2011). If no similar protein can be detected in another species, but an encoded protein contains an identifiable domain, this domain might be a hint that functional relationships exist between otherwise unrelated proteins. For example only a minority of transcription factors has identifiable direct orthologs, but the transcription factor domains are equally present in all genomes. Possibly, the lack of similarity is here due to the adaptation of these factors to a specific genomic environment.

Alignments of coding DNA between species are possible, if the coding region in question represents a highly conserved protein and the protein sequence is used as a guide for the alignment (Fig. 1). The third position of each codon can be altered often without changing the amino acid sequence in the protein. The alignment shows that this is the case at almost all third positions, which indicates a saturation with mutations. This fact emphasizes the large evolutionary distance between the species analysed. Synteny is scarce between groups, only a few genes have the same neighbours in species of different groups. Moreover, the location on a certain chromosome as it is the case in Drosophila species (Heger and Ponting 2007) is also not conserved. This indicates a frequent reshuffling of the genetic material irrespective of the chromosomal position.

The analysis of gene families revealed that only some specific families are affected by differential family expansions, others have nearly the same number of members in each genome with clear orthology relationships among the members. One of the most prominent examples of species specific differences is the polyketide synthase family (PKS). This family has more than 15 members in each species. A phylogenetic analysis revealed that only a few PKS family members had orthologs in other species, the majority of the members cluster together in species specific clades. Thus, while this family has a common origin, species specific amplifications contribute to species diversification.

Taken together, families encoding proteins required for environmental sensing, defence, or production of secondary metabolites seem to be driving forces for speciation, while the basic repertoire for the vegetative and developmental cycle remained remarkably stable.

Interestingly, chromosomal structures are variable within the social amoebae. The DIRS element is present as a cluster at each tip of the chromosomes of *D. disoideum* and here constitutes the centromeres (Dubin et al. 2010). This might

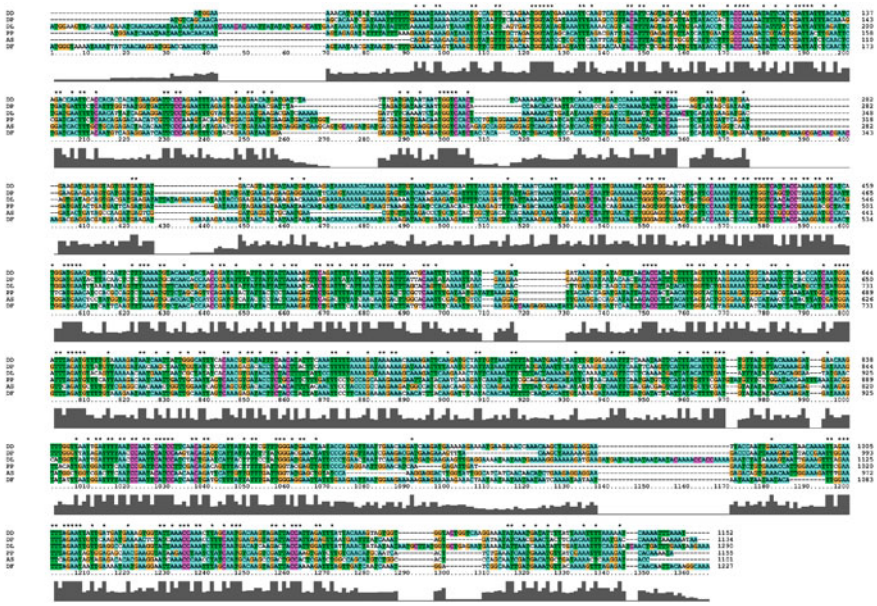

Fig. 1 DNA alignment of gene cdc123 between six social amoebae of which the genomic sequence is known. Coding sequences were retrieved from the respective databases and aligned in MEGA5 (Tamura et al. 2011) based on the amino acid sequence. Thus the alignment was built from three-base blocks (codons). Identical residues are marked with an *asterisk* above the alignment. The overall conservation is depicted as *blocks* below the alignment. DD *Dictyostelium discoideum*; DP *D. purpureum* (both group 4); DL *D. lacteum* (group 3); PP *Polysphondilium pallidum*, AS *Acytostelium subglobosum* (both group 2); DF *D. fasciculatum* (group 1)

also be true for *D. fasciculatum*, where also, albeit smaller, DIRS clusters were found at chromosomal tips (Fig. 2). All other genomes as the nearly complete assemblies of *P. pallidum*, *D. lacteum*, and *D. fasciculatum* show, do not contain enough DIRS sequences to endow each chromosome with at least one DIRS element. Instead, in case of *D. lacteum* a short species specific sequence is located at chromosome tips, whereas *P. pallidum* has a complex pattern of sequences at each end of the chromosome (Fig. 2). It is not clear, which sequences fulfil the role of centromeres in these species since no functional analysis was done so far. We only can state that obviously the functional elements of centromeres are different from those in *D. discoideum*. Likely, DIRS elements were already the centromeres in the last common ancestor of social amoebae, since the most divergent groups 1 and 4, represented by *D. fasciculatum* and *D. discoideum*, have such centromeres. The other species analysed have engaged other sequences to fulfil these functions. The telomere sequences also differ between species (Fig. 2). While group 1–3 species have normal eukaryote telomere structures with differing telomere repeat sequences at chromosomes (Table 1 and Fig. 2), group 4 species seem to have replaced these with long sequences stemming from the rDNA palindrome (see below). The rDNA palindromes normally have the same structures at the ends as

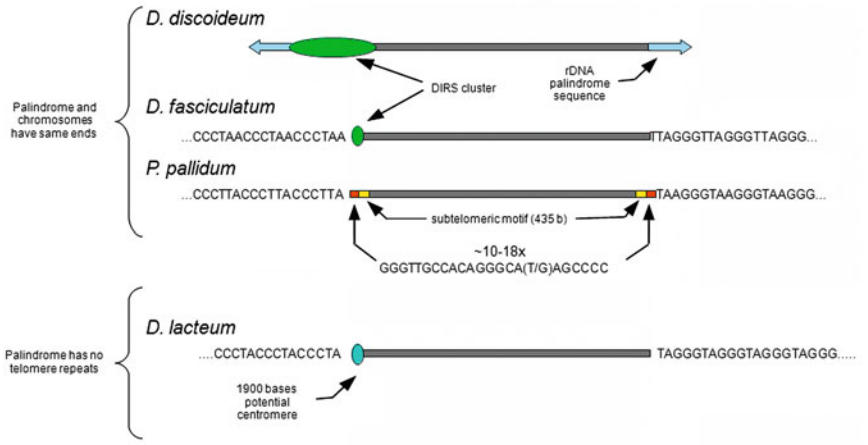

Fig. 2 Schematic representation of chromosome structures and palindromes. Only species for which the structures were investigated in detail are shown

the chromosomes of the species. Only in case of *D. lacteum* these structures differ between chromosomes and rDNA palindrome. While the chromosomes are endowed with normal telomere repeat sequences, the palindrome has structures resembling those in *D. discoideum*. This situation could be interpreted as an intermediate stage between that with the basic eukaryote telomere sequences seen in groups 1 and 2 and the complete replacement of those sequences by a specific sequence stretch as in *D. discoideum* and presumably also in *D. purpureum*, since no eukaryote telomere sequences could be found in this assembly (own analysis).

Taken together, the original centromeres and telomeres seem to have been exchanged several times in the social amoebae clade so that each species has its unique outfit with such features.

3 Plasmids, rDNA Palindrome, and Mitochondrial Genomes

Besides the nuclear genome in social amoebae there exist extrachromosomal elements and plasmids in the nucleus and organellar genomes in mitochondria.

3.1 Plasmids

It is unknown which role plasmids play in the life of social amoebae. No function besides for plasmid maintenance itself could so far be assigned to one of the

protein coding genes of the deciphered plasmid sequences. Plasmids seem also to be dispensable since some strains harbour them and some not without any obvious differences in fitness or other features (Rieben et al. 1998).

3.2 rDNA Palindromes

Extrachromosomal elements encode the rRNA genes on an inverted repeat with a small centre region of only a few bases. Such an organisation is also called a palindrome. The purpose of the palindrome is to provide the organism with an amplified number of rRNA genes just like the repeated arrays of rRNA genes in chromosomes of other organisms (Pendas et al. 1993). Besides the genes the palindrome contains repeated sequence stretches, which are species specific. Thus, the palindromes cannot be aligned to each other outside the genic regions. The extent of the repeats of the non-coding part of the palindrome arms is responsible for the varying length of the palindrome in the different species from a mere 14 kb in *D. lacteum* to more than 80 kb for one arm in *D. discoideum* (Heidel et al. 2011; Sucgang et al. 2003). Since the palindromes are presumably being replicated like chromosomes they also harbour species specific chromosomal end structures (Fig. 2).

3.3 Mitochondria

The mitochondrial genomes are around 45 kb in length (Table 1). They all encode the same genes mainly in the same order. Only one segmental shift occurring in group 4 could be observed (Heidel and Glöckner 2008). Comparison to the *Acanthamoeba castellanii* mitochondrial genome revealed extensive syntenic regions also to this distantly related genome within the Amoebozoa (Glöckner and Noegel 2013). The mitochondrial genomes are thus stable in terms of coding capacity and gene order, yet the nucleotide sequence is so divergent that alignments outside the coding regions are impossible. The coding parts of the mitochondrial genomes were also used to calculate a molecular phylogeny. The tree obtained differed from the previously calculated rDNA and some protein coding genes based phylogeny of social amoebae, but agreed with the second best possible tree (Schaap et al. 2006). Later calculations based on large concatenated data sets from complete genomes confirmed the correctness of the mitochondrial tree. Thus, the revised social amoebae phylogeny shows a bifurcation of groups 1/2 and 3/4 rather than a sequential emergence of the major branches.

4 Regulation of Transcription on a Global Scale

Efforts are underway to characterize the whole transcriptomes during the developmental cycle of all social amoebae of which the genomes are currently known. So far, only data on the transcriptomes of *D. discoideum* and *D. purpureum* are available. Initially, for *D. discoideum* microarrays were used to produce expression profiles of all at that time available genes (Booth et al. 2005; Sasik et al. 2002). Later, RNASeq (Oshlack et al. 2010) was also employed, which is a more sensitive method for transcriptome analysis.

A comparison of transcription of *D. discoideum* and *D. purpureum* during the developmental cycle revealed that this program is largely conserved between the two species (Parikh et al. 2010). Given the large evolutionary distance of the two species, which is estimated to be between 300 and 400 million years, this is remarkable. It remains to be shown that this grade of conservation is also present in the more distantly related groups. Yet, at least the core components remained presumably unchanged since the establishment of the whole clade some 600 million years ago.

Only in the last few years it became apparent that besides proteins also small RNAs play a pivotal role in regulation of expression. Some species of these RNAs play roles in silencing transposons. Indeed, a large number of such RNAs were found to regulate DIRS and skipper elements in *D. discoideum* (Hinas et al. 2007). Furthermore, also some developmental genes seem to be regulated by such RNAs. For other social amoebae species currently are no data on small RNAs available.

Methylation of DNA or histones is another way to regulate transcription. The genomes of social amoebae only harbour DNA-methyltransferase, which methylates Cs in a context independent manner. These methylases are involved in silencing of some, but not all transposon species (Kuhlmann et al. 2005). Thus, regulation of transposon activity is either achieved by siRNAs (DIRS and skipper) or DNA methylation (TRE elements). Histone methylation was also proven to play a role in developmental timing in *D. discoideum* (Chubb et al. 2006).

Gene transcription is also influenced by the packaging state of the DNA around a certain locus (Farkas et al. 2000; Ito 2003). Thus, the analysis of nucleosome positions around genes is an indispensible step towards the complete understanding of gene regulation in an organism. Global analysis of nucleosomes can be done using next generation sequencing technology. In short, DNA is cross-linked with the proteins attached to it and is subsequently digested using endonucleases. This way, only DNA covered by proteins remains intact and can be, after selective enrichment and reversal of the cross-link, sequenced. Such analyses in different eukaryote model organisms revealed that non- coding portions of a gene are normally depleted of nucleosomes. Not surprisingly, this same pattern seemingly common to all eukaryotes was found in *D. discoideum*. In *D. discoideum* the transcripts have poly-T tracts at their 5' end and poly A-tracts at their 3' end. These tracts are associated with nucleosome free regions. The nucleosome border regions are precisely marked with homo- and heteropolymeric tracts of A and T

nucleotides (Chang et al. 2012). The nucleosomes do not change their position upon entry into the developmental cycle as revealed by the comparative analysis of nucleosome positions between the vegetative state and 4 h into the developmental cycle. One interesting finding was that the transcriptional start site is marked by nucleosomes at the +1 position indicative of a paused RNA polymerase in the same way as in most multicellular systems in contrast to unicellular species. Thus, a multicellular stage seems to depend not only on coding capabilities of an organism but also on DNA occupancy properties. Since currently nucleosome data are available only for a handful of species this finding must remain preliminary until more species have been examined.

5 Curation of Genomes and Databases

Genome analyses should not end with the completion of a first draft annotation of a genome. Rather, annotation must proceed and ameliorated by integrating literature, the latest genome sequence, new data like expression analysis, and information on mutation phenotypes. Fortunately, in the early days of the *D. disoideum* genome project a database system was set up to fulfil all these purposes (Fey et al. 2009). Now, dictybase (http://dictybase.org) has curated all *D. discoideum* predicted genes and made them publicly available. Other genomes are being integrated so that this database will give a comprehensive overview on all available genomes of social amoebae. Another database has been used to initially annotate the genomes of three social amoebae (*D. lacteum*, *D. fasciculatum*, and *P. pallidum*; http://sacgb.fli-leibniz.de) (Felder et al. 2013). This database also provides alignments of all genes to their best bidirectional hits and tools for the analysis of synteny between genomes. A further database provides information on the draft and yet incomplete genome of *Acytostelium subglobosum* (http://acytodb.biol.tsukuba.ac.jp/).

6 Future Directions of Genomics in Social Amoebae

Social amoebae likely diverged at least 600 million years ago (Heidel et al. 2011). This is a larger time span than the evolution of vascular plants needed, and only a little bit less than that for the evolution of Metazoa. Given the big differences in genome size and coding capacity observed in these evolutionary branches, it is surprising that these features remained stable in the social amoebae. Since we now have genomes from all major social amoebae branches in hand we can extrapolate that this is true for most, if not all social amoebae. The small genome size thus makes it easy to sequence additional genomes at very low costs. However, this approach only makes sense if species specific inventions are under investigation, since the common genetic outfit of social amoebae has been described already in

detail (Heidel et al. 2011). The genome sequence of new species must be also endowed with transcriptome data. Of great interest will be to investigate the regulation and timing of transcription in diverse social amoebae taxa. This will foster our understanding of the evolution of regulatory circuits. Systems biology approaches then can elucidate why the developmental cycle remained stable and whether this stability is an intrinsic feature of the sophistication of the cycle.

References

Booth EO, Van Driessche N, Zhuchenko O, Kuspa A, Shaulsky G (2005) Microarray phenotyping in Dictyostelium reveals a regulon of chemotaxis genes. Bioinformatics 21(24):4371–4377

Chang GS, Noegel AA, Mavrich TN, Muller R, Tomsho L, Ward E, Felder M, Jiang C, Eichinger L, Glöckner G, Schuster SC, Pugh BF (2012) Unusual combinatorial involvement of poly-A/T tracts in organizing genes and chromatin in Dictyostelium. Genome Res 22(6):1098–1106

Chubb JR, Bloomfield G, Xu Q, Kaller M, Ivens A, Skelton J, Turner BM, Nellen W, Shaulsky G, Kay RR, Bickmore WA, Singer RH (2006) Developmental timing in Dictyostelium is regulated by the Set1 histone methyltransferase. Dev Biol 292(2):519–532

Dear PH, Cook PR (1993) Happy mapping: linkage mapping using a physical analogue of meiosis. Nucleic Acids Res 21(1):13–20

Dubin M, Fuchs J, Graf R, Schubert I, Nellen W (2010) Dynamics of a novel centromeric histone variant CenH3 reveals the evolutionary ancestral timing of centromere biogenesis. Nucleic Acids Res 38(21):7526–7537

Eichinger L, Pachebat JA, Glöckner G, Rajandream MA, Sucgang R, Berriman M, Song J, Olsen R, Szafranski K, Xu Q, Tunggal B, Kummerfeld S, Madera M, Konfortov BA, Rivero F, Bankier AT, Lehmann R, Hamlin N, Davies R, Gaudet P, Fey P, Pilcher K, Chen G, Saunders D, Sodergren E, Davis P, Kerhornou A, Nie X, Hall N, Anjard C, Hemphill L, Bason N, Farbrother P, Desany B, Just E, Morio T, Rost R, Churcher C, Cooper J, Haydock S, van Driessche N, Cronin A, Goodhead I, Muzny D, Mourier T, Pain A, Lu M, Harper D, Lindsay R, Hauser H, James K, Quiles M, Madan Babu M, Saito T, Buchrieser C, Wardroper A, Felder M, Thangavelu M, Johnson D, Knights A, Loulseged H, Mungall K, Oliver K, Price C, Quail MA, Urushihara H, Hernandez J, Rabbinowitsch E, Steffen D, Sanders M, Ma J, Kohara Y, Sharp S, Simmonds M, Spiegler S, Tivey A, Sugano S, White B, Walker D, Woodward J, Winckler T, Tanaka Y, Shaulsky G, Schleicher M, Weinstock G, Rosenthal A, Cox EC, Chisholm RL, Gibbs R, Loomis WF, Platzer M, Kay RR, Williams J, Dear PH, Noegel AA, Barrell B, Kuspa A (2005) The genome of the social amoeba *Dictyostelium discoideum*. Nature 435(7038):43–57

Farkas G, Leibovitch BA, Elgin SC (2000) Chromatin organization and transcriptional control of gene expression in Drosophila. Gene 253(2):117–136

Felder M, Romualdi A, Petzold A, Platzer M, Suhnel J, Glöckner G (2013) GenColors-based comparative genome databases for small eukaryotic genomes. Nucleic Acids Res 41(D1):D692–D699

Fey P, Gaudet P, Curk T, Zupan B, Just EM, Basu S, Merchant SN, Bushmanova YA, Shaulsky G, Kibbe WA, Chisholm RL (2009) DictyBase-a *Dictyostelium* bioinformatics resource update. Nucleic Acids Res 37(Database issue):D515–519

Francis D (1998) High frequency recombination during the sexual cycle of *Dictyostelium discoideum*. Genetics 148(4):1829–1832

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J,

Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419(6906):498–511

Glöckner G, Eichinger L, Szafranski K, Pachebat JA, Bankier AT, Dear PH, Lehmann R, Baumgart C, Parra G, Abril JF, Guigo R, Kumpf K, Tunggal B, Cox E, Quail MA, Platzer M, Rosenthal A, Noegel AA (2002) Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*. Nature 418(6893):79–85

Glöckner G, Heidel AJ (2009) Centromere sequence and dynamics in *Dictyostelium discoideum*. Nucleic Acids Res 37(6):1809–1816

Glöckner G, Noegel AA (2013) Comparative genomics in the Amoebozoa clade. Biol Rev Camb Philos Soc 88(1):215–225

Glöckner G, Szafranski K, Winckler T, Dingermann T, Quail MA, Cox E, Eichinger L, Noegel AA, Rosenthal A (2001) The complex repeats of *Dictyostelium discoideum*. Genome Res 11(4):585–594

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. Science 274(5287):546, 563–567

Heger A, Ponting CP (2007) Evolutionary rate analyses of orthologs and paralogs from 12 Drosophila genomes. Genome Res 17(12):1837–1849

Heidel AJ, Glöckner G (2008) Mitochondrial genome evolution in the social amoebae. Mol Biol Evol 25(7):1440–1450

Heidel AJ, Lawal HM, Felder M, Schilde C, Helps NR, Tunggal B, Rivero F, John U, Schleicher M, Eichinger L, Platzer M, Noegel AA, Schaap P, Glöckner G (2011) Phylogeny-wide analysis of social amoeba genomes highlights ancient origins for complex intercellular communication. Genome Res 21(11):1882–1891

Hinas A, Reimegard J, Wagner EG, Nellen W, Ambros VR, Soderbom F (2007) The small RNA repertoire of *Dictyostelium discoideum* and its regulation by components of the RNAi pathway. Nucleic Acids Res 35(20):6714–6726

Ito T (2003) Nucleosome assembly and remodeling. Curr Top Microbiol Immunol 274:1–22

Joseph JM, Fey P, Ramalingam N, Liu XI, Rohlfs M, Noegel AA, Muller-Taubenberger A, Glöckner G, Schleicher M (2008) The actinome of *Dictyostelium discoideum* in comparison to actins and actin-related proteins from other organisms. PLoS ONE 3(7):e2654

Kollmar M (2006) Thirteen is enough: the myosins of *Dictyostelium discoideum* and their light chains. BMC Genomics 7:183

Kollmar M, Glöckner G (2003) Identification and phylogenetic analysis of *Dictyostelium discoideum* kinesin proteins. BMC Genomics 4(1):47

Konfortov BA, Cohen HM, Bankier AT, Dear PH (2000) A high-resolution HAPPY map of *Dictyostelium discoideum* chromosome 6. Genome Res 10(11):1737–1742

Kuhlmann M, Borisova BE, Kaller M, Larsson P, Stach D, Na J, Eichinger L, Lyko F, Ambros V, Soderbom F, Hammann C, Nellen W (2005) Silencing of retrotransposons in Dictyostelium by DNA methylation and RNAi. Nucleic Acids Res 33(19):6405–6417

Kuspa A, Maghakian D, Bergesch P, Loomis WF (1992) Physical mapping of genes to specific chromosomes in *Dictyostelium discoideum*. Genomics 13(1):49–61

Loomis WF, Welker D, Hughes J, Maghakian D, Kuspa A (1995) Integrated maps of the chromosomes in *Dictyostelium discoideum*. Genetics 141(1):147–157

Oshlack A, Robinson MD, Young MD (2010) From RNA-seq reads to differential expression results. Genome Biol 11(12):220

Parikh A, Miranda ER, Katoh-Kurasawa M, Fuller D, Rot G, Zagar L, Curk T, Sucgang R, Chen R, Zupan B, Loomis WF, Kuspa A, Shaulsky G (2010) Conserved developmental transcriptomes in evolutionarily divergent species. Genome Biol 11(3):R35

Pendas AM, Moran P, Garcia-Vazquez E (1993) Ribosomal RNA genes are interspersed throughout a heterochromatic chromosome arm in Atlantic salmon. Cytogenet Cell Genet 63(2):128–130

Rieben WK Jr, Gonzales CM, Gonzales ST, Pilkington KJ, Kiyosawa H, Hughes JE, Welker DL (1998) *Dictyostelium discoideum* nuclear plasmid Ddp5 is a chimera related to the Ddp1 and Ddp2 plasmid families. Genetics 148(3):1117–1125

Rivero F, Dislich H, Glöckner G, Noegel AA (2001) The *Dictyostelium discoideum* family of Rho-related proteins. Nucleic Acids Res 29(5):1068–1079

Sasik R, Iranfar N, Hwa T, Loomis WF (2002) Extracting transcriptional events from temporal gene expression patterns during Dictyostelium development. Bioinformatics 18(1):61–66

Schaap P, Winckler T, Nelson M, Alvarez-Curto E, Elgie B, Hagiwara H, Cavender J, Milano-Curto A, Rozen DE, Dingermann T, Mutzel R, Baldauf SL (2006) Molecular phylogeny and evolution of morphology in the social amoebas. Science 314(5799):661–663

Sucgang R, Chen G, Liu W, Lindsay R, Lu J, Muzny D, Shaulsky G, Loomis W, Gibbs R, Kuspa A (2003) Sequence and structure of the extrachromosomal palindrome encoding the ribosomal RNA genes in *Dictyostelium*. Nucleic Acids Res 31(9):2361–2368

Sucgang R, Kuo A, Tian X, Salerno W, Parikh A, Feasley CL, Dalin E, Tu H, Huang E, Barry K, Lindquist E, Shapiro H, Bruce D, Schmutz J, Salamov A, Fey P, Gaudet P, Anjard C, Babu MM, Basu S, Bushmanova Y, van der Wel H, Katoh-Kurasawa M, Dinh C, Coutinho PM, Saito T, Elias M, Schaap P, Kay RR, Henrissat B, Eichinger L, Rivero F, Putnam NH, West CM, Loomis WF, Chisholm RL, Shaulsky G, Strassmann JE, Queller DC, Kuspa A, Grigoriev IV (2011) Comparative genomics of the social amoebae *Dictyostelium discoideum* and *Dictyostelium purpureum*. Genome Biol 12(2):R20

Szafranski K, Lehmann R, Parra G, Guigo R, Glöckner G (2005) Gene organization features in A/T-rich organisms. J Mol Evol 60(1):90–98

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28(10):2731–2739

Wilkins A, Szafranski K, Fraser DJ, Bakthavatsalam D, Muller R, Fisher PR, Glöckner G, Eichinger L, Noegel AA, Insall RH (2005) The Dictyostelium genome encodes numerous RasGEFs with multiple biological roles. Genome Biol 6(8):R68

Williams JG, Firtel RA (2000) HAPPY days for the Dictyostelium genome project. Genome Res 10(11):1658–1659

Winckler T, Dingermann T, Glöckner G (2002) Dictyostelium mobile elements: strategies to amplify in a compact genome. Cell Mol Life Sci 59(12):2097–2111

Winckler T, Szafranski K, Glöckner G (2005) Transfer RNA gene-targeted integration: an adaptation of retrotransposable elements to survive in the compact *Dictyostelium discoideum* genome. Cytogenet Genome Res 110(1–4):288–298

http://www.springer.com/978-3-642-38486-8

Dictyostelids
Evolution, Genomics and Cell Biology
Romeralo, M.; Baldauf, S.; Escalante, R. (Eds.)
2013, VIII, 255 p., Hardcover
ISBN: 978-3-642-38486-8