The *Dictyostelium* Genome

William F. Loomis

Cell and Developmental Biology, Division of Biology, University of California San Diego, La Jolla, CA 92093, USA

Abstract

The 34 Mb genome of *Dictyostelium discoideum* is carried on 6 chromosomes and has been fully sequenced by an international consortium. The sequence was assembled on the classical and physical maps that had been built up over the years and refined by HAPPY mapping. Annotation of the sequence predicted about 12,000 genes for proteins of at least 50 amino acids in length. The total number of amino acids encoded (the proteome) is more than double that in yeast and rivals that of metazoans. The genome sequence shows all the proteins available to *Dictyostelium* as well as definitively showing which domains have been lost since *Dictyostelium* diverged from the line leading to metazoans. Genomics opens the door to determining the expression patterns of all the genes during growth and development using microarrays. This approach has already uncovered a wealth of new markers for the stages of development and the various cell types. Transcription factors and their *cis*-regulatory sites that account for the surprising complexity of *Dictyostelium* development can be analyzed much more easily now that we have the complete sequence.

Introduction

Completion of the 34 Mb sequence of the *Dictyostelium* genome in 2004 brought the resolution of the genetic maps to the nucleotide level. Many years of effort to characterize genetic markers and map them to the 6 chromosomes paid off in the successful assembly of over 600,000 sequencing reads into less than 300 contigs that were positioned and oriented on the basis of ordered markers. The chromosomal position of any specific gene gives very little biological insight because the gene has to be somewhere in the chromosomes, but the relative position of related genes sheds light on the genetic history and can lead to novel approaches to generate complex genotypes. Perhaps the most important consequence of having the complete genomic sequence is that we now know all the proteins that *Dictyostelium* has at its disposal.

Early days

Seventy years ago Ken Raper isolated *Dictyostelium discoideum* from the forest detritus at Little Butts Gap, North Carolina (Raper, 1935). This species still abounds in the area (Strassman *et al.*, 2000). Raper took samples into his laboratory where he found he could grow the amoebae in petri dishes filled with nutrient agar that were spread with a lawn of bacteria for the cells to eat. On such plates a spore will germinate, engulf bacteria and multiply. After a few days at 22º, translucent plaques can be seen in the bacterial lawn where clonal populations derived from a single cell or spore have multiplied and cleared the area. Over the next few days, cells within the plaque will aggregate into hemispherical mounds each containing up to 10^5 cells. A tip is formed at the top, which leads the slug-shaped organism as it migrates towards light. During culmination cells at the anterior form a stalk that penetrates the underlying mass of posterior cells. Prestalk cells, which make up the anterior 20% of the slug, enter the stalk, vacuolize and lay down thick cellulose walls such that the stalk rises several millimeters upwards and can support the mass of posterior cells that move up the stalk as it extends. A disc of vacuolized, cellulose encased cells is also formed at the base to anchor the fruiting body (Fig. 1). It is this structure that gave the species its name (*discoideum*). When prespore cells that had been at the back of the slug reach the top of the stalk, they receive signals to encapsulate into spores. The resulting fruiting bodies have smoothly tapering stalks with symmetric spheres of spores at the top. Since each plaque has a genetically identical clonal population, all the fruiting bodies that develop within a single plaque are similar. Any strain carrying a mutation that affects morphogenesis can be easily recognized by comparison of the fruiting bodies among plaques. As many as 100 plaques can be screened on a 9 cm petri dish allowing rare mutations to be isolated. It is this aspect of the life cycle that has attracted geneticists to the study of *D. discoideum.*

Over 50 years ago Maurice Sussman isolated strains that were unable to aggregate, others that aggregated rapidly (FR-17) or formed tiny fruiting bodies (*fruity*), as well as many other mutants with clear phenotypes (Sussman and Sussman, 1953; Sussman, 1954; 1955; Sonneborn *et al.*, 1963). Analyses of these strains led to many insights on cell-cell signaling but could not elucidate specific molecular pathways until cloning techniques were developed in the 1970s. Although we now know that the mutation in strain FR-17 inactivated a gene, *rdeA*, on chromosome 4 that encodes a signal transduction component affecting the cytoplasmic cAMP phosphodiesterase, most of the other mutations generated in the pre-cloning era are still uncharacterized. Nevertheless, they could be used to map the chromosomes.

With the introduction of chemical mutagens, such as nitrosoguanidine, the isolation of many classes of mutations became routine (Yanagisawa *et al.,* 1967). One class that was particularly useful for the isolation of diploids results in temperature sensitive growth. By picking cells from a few hundred plaques and replica plating them in grids on two plates, one incubated at 18º and the other at 28º, temperature sensitive (ts) mutants can be routinely isolated. When two independently isolated *ts* strains are

For correspondence: wloomis@ucsd.edu

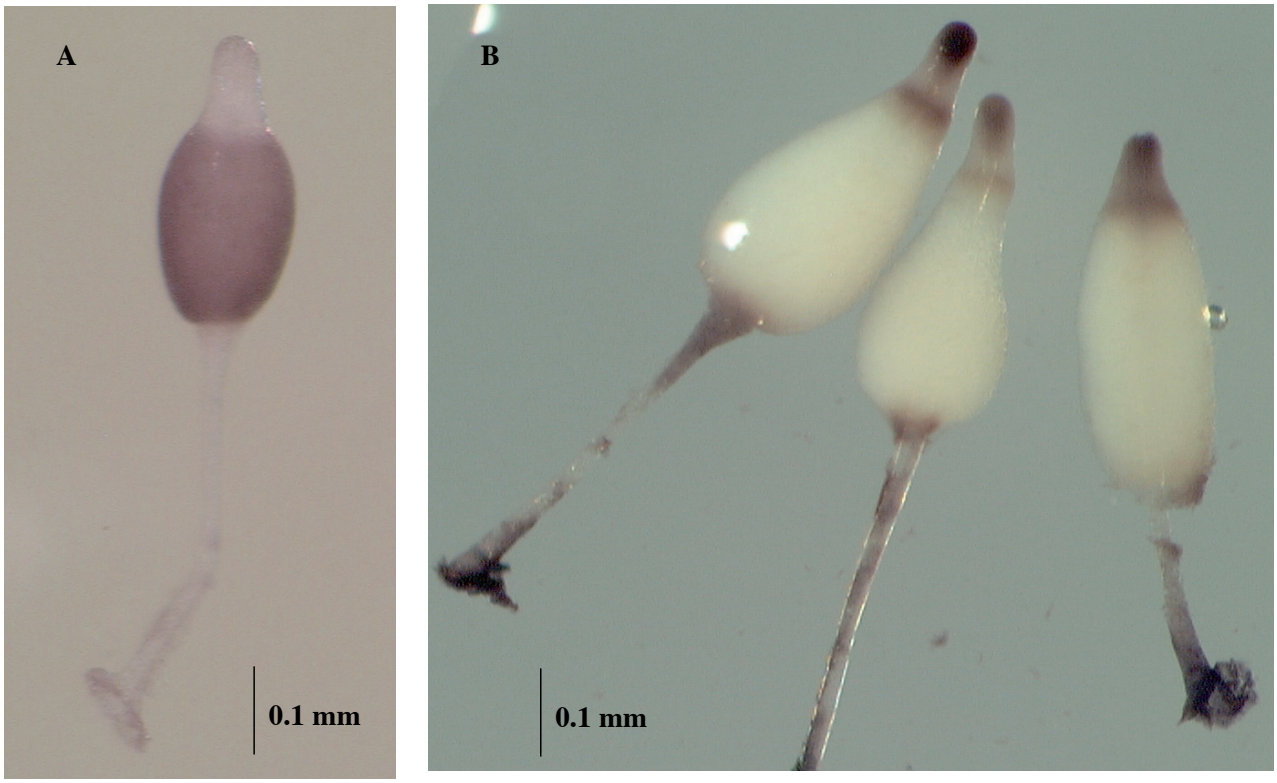


Fig. 1. Culminants. A basal disc can be seen at end of the stalk of each fruiting body. Prespore cells are still rising up the stalk at this stage and form a pear-shaped mass. A) *in situ* hybridization of *wacA* mRNA stains the prespore cells B) *in situ* hybridization of *ecmA* mRNAs stains the prestalk cells at the apex as well as cells within the stalk and in the basal disc.

developed together, rare diploids are formed and can be isolated by plating at 28º. Analyses of such diploid strains made it clear that the original NC-4 strain is haploid and that diploids arise at a frequency of about 10^{-5} (Loomis, 1969). Diploids strains were found to spontaneously re-express recessive markers at a low frequency which could be increased by disrupting microtubule structures with compounds such as ben late (Loomis, 1969; Williams and Barrand, 1978). The segregation patterns of genetic markers indicated that diploids randomly segregate chromosomes to generate haploids (Sinha and Ashworth, 1969). This parasexual cycle opened the door to a variety of mapping techniques (Loomis, 1987).

Katz and Sussman (1972) found the first sets of linked markers that defined linkage groups I and II. Other linkage groups were rapidly defined by other markers until a total of 6 linkage groups were well established (Newell *et al.* 1993). Each linkage group appeared to define a chromosome since chromosomal stains of haploid strains indicated the presence of 7 chromosomes (more about the seventh “chromosome” later) (Sussman, 1961; Robson and Williams, 1977; Sucgang *et al.*, 2003). At this point renewed efforts were made to better understand the sexual cycle so that the high frequency of recombination that is usually associated with meiosis could be harnessed to order the markers along the chromosomes. When certain strains which had been independently isolated from nature were allowed to develop together submerged in the dark, they were found to form thick walled dark spheres, termed macrocysts, that were not seen with cells from a single isolate (Erdos *et al.*, 1973). This phenomenon had the trappings of sex. In fact, large binucleated cells were seen to form within the macrocyst that subsequently engulfed surrounding cells. When plated on bacterial lawns, a few of the macrocysts gave rise to plaques. However, when macrocysts were generated with multiply marked strains, the progeny of a single macrocyst were almost always found to be genetically identical. Progeny from different macrocysts showed different assortments of markers, but there were no apparent Mendelian patterns that could be used for mapping on the basis of recombinational frequencies. Nevertheless, both parental and non-parental di-types could be found among linked loci in the progeny of some macrocysts (Wallace and Raper, 1979). This was good evidence for meiotic segregation but the efficiency of macrocyst germination was too low for large scale mapping. Only when physical mapping had progressed to the point where a variety of markers were known to be within 200 kb could meiotic recombination frequency be estimated. Using PCR based assays for the segregants it was found that the recombinational frequency was very high such that crossing over occurred 20% of the time between some genes that were only about 100 kb apart (Francis, 1998). This frequency of recombination is comparable to that in yeast. In retrospect, such high frequency meiotic recombination would have confounded attempts to recognize linkage groups because it would have required pairwise comparison of several thousand

markers before chromosomal maps could be built up. Luckily, parasexual segregation was an efficient way to define the chromosomes.

Physical mapping

Molecular cloning could assign specific genes to a linkage group on the basis of variations in flanking restriction enzyme sites. The technique of restriction fragment length polymorphism (RFLP) mapping depends on independent isolates from nature that have the same long range order of markers but fine scale differences in sequence. When a large number of restriction enzymes are used to characterize the flanking regions, alleles can be distinguished on the basis of RFLP. The pattern of hybridization of a gene marker on the gels is the phenotype. Parasexual mapping with marked strains together with RFLP put some of the first *Dictyostelium* cloned genes into linkage groups (Welker *et al.*, 1986).

To extend this approach to the coverage of complete chromosomes required the development of pulsed field gel electrophoresis that could reliably separate fragments up to several megabases (Mb) in length. Such techniques were used for whole chromosomes from *Dictyostelium* and showed that there were 6 bands in the range of 3 to 9 Mb (Cox *et al.*, 1990). Many of the bands could be assigned to specific linkage groups on the basis of hybridization to Southern blots of the whole-chromosome pulsed field gels. However, single copy genes were found to give weak signals. As the number of cloned genes accumulated, it became possible to generate maps *de novo* that could be assigned to the linkage groups. A library of 147 independent strains with random insertions of a plasmid carrying rare restriction sites such as Not 1 and Apa 1 was generated using Restriction Enzyme Mediated Insertion (REMI) (Kuspa and Loomis, 1994). DNA from these strains could be digested with a variety of rare cutting restriction enzymes and the fragments separated on pulsed field gels. Blots of the gels were then probed with hundreds of cloned genes and cases where sets of genes were all found in a smaller fragment in one of the REMI-RFLP strain but not in others were taken as evidence for linkage. Since the size of the native fragment for each probe was known from single restriction enzyme digests as well as from multiple enzyme digests, a long range map could be built up with the position of each insertion site well defined and tethered to flanking genes (Loomis *et al.*, 1995). In some cases the evidence for linkage of a gene was weak or equivocal. Probes from these genes were RFLP mapped using classical parasexual genetics allowing them to be assigned to a specific linkage group. With this information the difficult genes could all be used to define the emerging maps of the 6 chromosomes (Loomis *et al.*, 1995). All told, over 150 genes were ordered over the 34 Mb genome. The position of each gene was given within a few hundred kb and provided the template for further high resolution mapping and sequencing. Only multicopy genes caused problems. While other approaches could have been used to sort out the errors and refine the map, it was decided that the work and expense this would entail was not worth the effort and that it made more sense to make the corrections and increase the map resolution by sequencing the whole genome. Moreover, the sequence would provide system-wide information of enormous value concerning the genetic repetoire of *Dictyostelium*. With the physical maps of the 6 chromosomes in hand, plans were made for an international effort to sequence the 34 Mb genome.

Sequencing

When sequencing began in 1998 there were two major approaches being used for other large genomes: either segmental sequencing using a set of Bacterial Artificial Chromosome (BAC) clones with ~ 200 kb inserts or whole genome shot-gun sequencing with reliance on paired-reads from both ends of small inserts carried in plasmids. Each approach had its advantages and its weaknesses. The BAC-by-BAC approach makes assembly of individual reads much easier but requires establishing a tiling set of BACs along each chromosome. Whole genome shot-gun sequencing uses convenient plasmid clones as templates but requires ten fold redundancy in reads to be able to robustly assemble on the basis of sequence alone. The assembly of plasmid reads can be confounded by repetitive elements such as retrotransposons as well as by homopolymer runs which are known to be scattered throughout the *Dictyostelium* genome. For this and other resasons, considerable effort was made to generate large-insert BAC libraries of the *Dictyostelium* genome. However, it was found that the overall high A/T content of the genome (87%), precluded maintenance of large-insert BAC clones in bacteria (Souza and Kuspa, unpublished). Partial deletions of the inserts were observed as soon as the BACs were grown up in bacteria. Such random deletions would negate the advantages of BACs for assembly. Cathy Zheng and Peter de Jong made circular YAC/BAC clones in yeast that were used in a limited way in the early stages of assembly of Chromosome 2, but this bottom up approach quickly gave way to a chromosome-by-chromosome approach (Glockner *et al*., 2001). The consortium decided to sequence plasmid libraries generated from individual chromosomes that had been separated by pulsed field electrophoresis in the laboratory of Ted Cox at Princeton, N.J.. This chromosome-by-chromosome approach facilitated the distribution of efforts by assigning chromosomes 1, 2 and 3 to the team at the Institute of Molecular Biotechnology in Jena, and assigning chromosomes 4, 5 and 6 to the team at Baylor College of Medicine, Houston, Texas and the Sanger Centre, Hinxton, UK.

If difficulties in assembly from whole chromosomes reads were encountered, the sequencing teams knew they could turn to a set of large insert clones carried in Yeast Artificial Chromosomes (YACs). Unlike BACs with large inserts of *Dictyostelium* DNA which are unstable in bacteria, YACs with similar inserts are stable when carried in yeast. Prior to the start of the sequencing effort a set of 5,000 YAC clones with inserts that averaged 250 kb were prepared and probed with 400 cloned genes many of which had been ordered by REMI-RFLP (Kuspa and Loomis, 1996). A tiling set that covered each of the 6 chromosomes was found which could be thinly sequenced and used to bin whole chromosome reads. YACs were only sequenced when necessary to resolve

a given chromosome assembly because it is difficult to isolate insert DNA free of contamination with yeast DNA and the yields are low due to the low copy number of YACs.

Funding from the Deutsche Forschungsgemeinschaft was received for sequencing at Jena in 1998 followed shortly thereafter by funding from the Institute for Child Health and Human Development of the NIH was received for sequencing in Houston and funding from the European Union and later from the Wellcome Trust was recieved for sequencing at the Sanger Centre. Genome sequencing is expensive and more than 8 million dollars was required before the *Dictyostelium* genome was finished. Within a few years over 400,000 reads had been made providing about 5-fold coverage. Since the pulsed field fractions were contaminated with about 50% DNA from each of the other chromosomes, data from the various sequencing centers was pooled to generate contigs which provided a first look at about two thirds of the genome. Hidden Markov Model (HMM) programs were trained with over 400 genes with known exons and introns to establish parameters for automatic gene finding. When these programs were applied to the contigs, they found over 8,000 genes. A Preliminary Directory was established on a public web site in 2001 giving the proteins and domain structures. This site was a significant resource to the field for several years before the genome was completed. Genes could be discovered with the click of a mouse rather than laborious screening of genomic libraries or PCR amplification with degenerate primers. In 2002 the Preliminary Directory was updated with contigs prepared at the Sanger Centre from 8-fold coverage of the genome. Over 12,000 potential genes were found and annotated.

HAPPY mapping

Shortly after sequencing began in earnest, Paul Dear's lab at the MRC Cambridge was enrolled to map the chromosomes to 10 kb resolution using the HAPPY mapping technique they had invented (Dear and Cook, 1993). This technique depends on PCR with nested primers designed from random sequencing reads to generate fragments of a few hundred bases. Automatic scanning of electrophoretic gels is used to determine whether a cognate template is present in a sample or not. Genomic DNA is sheared to an average size of 100 kb and deposited in a large number of multitest wells such that each well receives only a portion of the genome, usually about 50%. Genes that are closely linked will almost always be found in the same 100 kb fragments and both will be amplified in the same well. Genes that are far apart on the chromosomes will be amplified from samples that overlap less frequently. Since the break points of the 100 kb fragments are random, one can connect sequences over a complete chromosome when enough have been HAPPY mapped. The frequency of co-segregation is used just as it is in meiotic recombination mapping to estimate the relative distance between markers. There was no shortage of sequences that could be used for HAPPY mapping since the sequencing centers were rapidly turning out thousands of reads a week.

Attention was first turned to Chromosome 6, the smallest chromosome. Using probes designed from reads of the chromosome 6 enriched libraries as well as YACs previously mapped to this chromosome, HAPPY mapping was able to position 300 sequence-tagged sites over the 3.6 Mb chromosome (Konfortov *et al*., 2000). Similar approaches were used to HAPPY map Chromosome 2, the largest chromosome. Long range restriction mapping had shown that there is a 1.5 Mb segment containing an inverted duplication about a third of the way along this chromosome in the strain being sequenced, AX4 (Loomis *et al*., 1995). The duplication defied HAPPY mapping because of the size of the HAPPY fragments but could be bypassed on the way to the ends. With this one exception a complete HAPPY map was generated that covered this 8.5 Mb chromosome (Glockner *et al*., 2002).

HAPPY maps were then generated and positioned along the other four chromosomes on the basis of the established genetic markers. In most cases the HAPPY map confirmed the order and placement of these markers although it was based on completely independent data. Having map data for 3902 markers with 10 kb resolution turned out to be essential for generating sequence assemblies for chromsomes 1, 2 and 3 and verified the assemblies for chromosomes 4, 5 and 6. It had been recognized at the outset that assembly above the 50 kb range was a mapping problem, but it was only fully appreciated how valuable the HAPPY map would be as the chromosomes were all put together.

The genome

Ten-fold coverage of the complete genome was reached in 2003 and the reads were assembled into contigs that averaged 50 kb. They could be positioned on the HAPPY maps and gaps recognized. Pair reads provided evidence for cloned gaps that could be filled by further sequencing of the appropriate plasmids. Moreover, many of the gaps were closed when adjacent sequences were carefully inspected and found to overlap. Other gaps were closed by designing PCR primers on abutting contigs and sequencing the PCR products. Some of the remaining gaps resulted from long segments containing nothing but repetitive elements that could not be distinguished from other such elements in the genome. However, YAC skims could be used in some cases to resolve repeat gaps. These finishing steps closed all but 300 gaps genome-wide. Only the telomeres remained poorly defined as is the case for almost all eukaryotic genome sequences. Since the size and position of the gaps was known, it was possible to state authoritatively that >95% of the genome sequence was covered by 33.85 Mb of assembled sequence (Eichinger *et al*., 2005). Moreover, it was possible to estimate that >99% of all genes were represented by high quality sequence. The genome was finished and ready for inspection.

The chromosomes

Centromeres. In Fig. 2 the chromosomes are presented with the ends adjacent to a cluster of DIRS repetitive elements at the left. Long range restriction mapping had determined that there were 6 regions with multiple DIRS elements tightly packed together with other repetitive elements. Since each chromosome had a single such

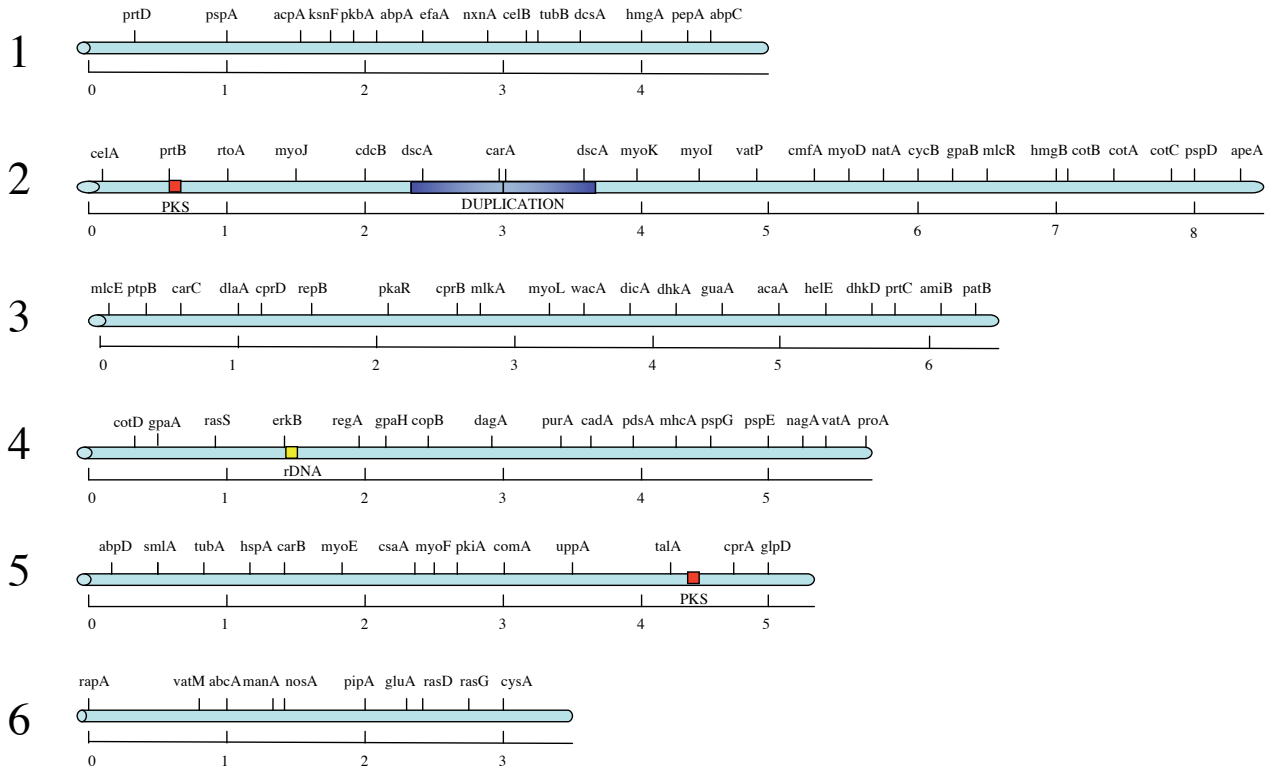


Fig. 2. Sequence based maps of the 6 chromosomes. Genes which had been physically mapped are positioned along the assembled chromosomes which are marked out in megabases. Individual loci are defined in Loomis *et al.*, (1994), Kuspa and Loomis (1995), Loomis and Kuspa (1997). The position of the 750 kb duplication on chromosome 2 is marked by a blue box. The master copy of the extrachromosomal palindrome on chromosome 4 is marked by a yellow box. The region of extraordinary duplications of genes encoding polyketide synthase on chromosome 5 is marked by a red box.

region near one end and the chromosomes were known to be telocentric, it was suggested that the DIRS clusters might be related to centromeric functions (Loomis *et al*, 1995). In fact, *in situ* hybridization of chromosomes during mitotic segregation with a fluorescently labeled DIRS probe has shown that each chromosome has a single labeled end and that they circle together during mitosis. The "seventh" chromosome seen by DAPI staining was shown to lack a DIRS cluster and to result from the condensation of the approximately 100 extrachromosomal copies of the 88 Kb palindromic DNA that carries the genes encoding ribosomal RNA (Sucgang *et al.* 2003).

The fine structure of the DIRS region on chromosome 1 has been solved (Eichinger *et al*., 2005). There are at least 10 complete copies of this 4.6 kb LTR transposon interspersed with the related Skipper LTR transposon. Partial DIRS sequences are found inserted into other DIRS elements and portions of Skipper are found associated with full length Skipper. Several non-LTR transposons are found interspersed among the LTRs. There is very little single copy DNA sequence over a 187 Kb region at the end of chromosome 1 and no apparent long range order. DIRS-like elements are found dispersed over the chromosomes of other organisms, suggesting that the clustering in *Dictyostelium* has been under purifying selection, perhaps to maintain a single centromeric function. However, it will be necessary to construct DACs (*Dictyostelium* Artificial Chromosomes) with a DIRS region linked to a selectable marker and show stable segregation during mitoses to be sure of the function.

Telomeres. The extrachromosomal palindromic rDNA is an interesting genetic element which appears to be generated from a master copy near *erkB* on chromosome 4 (Fig. 2) and then replicated free in the nucleus as the cells grow. Each 88 kb element ends in telomeres of four nearly perfect tandem repeats of 29 bp, 3 copies of AG_7, and a variable number of repeats of the form $AG_{1–8}$ (Emery and Weiner, 1981; Sucgang *et al*., 2003). Contigs have been found that contain junctions between sequences in the rDNA and repeat regions on chromosome 4 (Eichinger *et al*., 2005). The master copy contains the central 42 bp region of asymmetry flanked by 5 kb and 40 kb palindromic arms. If a single strand of DNA is replicated and excised from this portion of the genome, it could form a hairpin with the short arm hybridized to its palindromic copy on the long arm (Fig. 3). The 5 kb arm ends in a simple GC-rich repeat region that is also found on the long arm, 5 kb away from the central region of asymmetry. Extension of the short arm would complete the hairpin. When opened up and replicated, it would form an 80 kb double-stranded palindrome that could have telomeres added at the ends to protect them during further replication. Such a mechanism may also occur at each end of the chromosomes.

Although the telomeres are not well defined for most of the chromosomes, they all appear to have repetive elements immediately followed by short portions of the rDNA palindrome (Eichinger *et al*., 2005). The portions of the rDNA element differ at the ends of the chromosomes, but they may each provide signals for addition of telomeric sequences. In any case, it is more than a coincidence

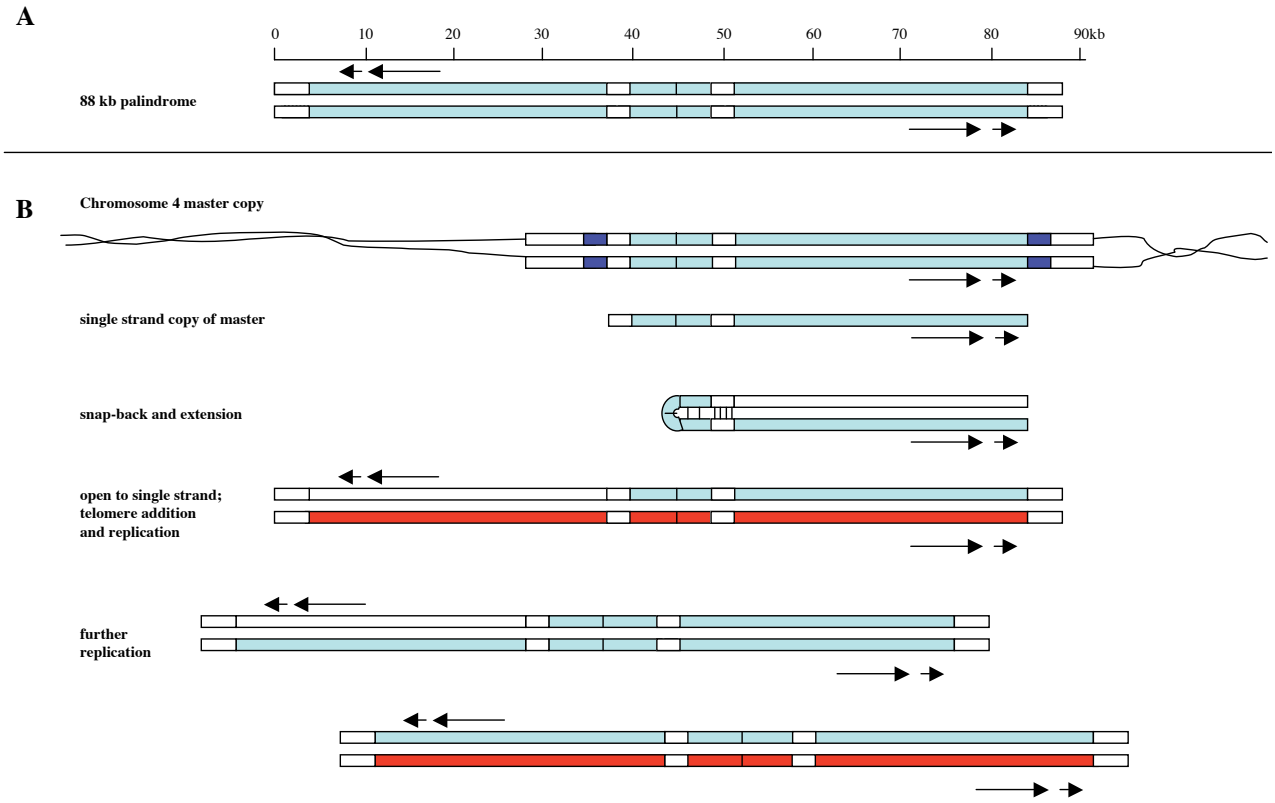


Fig. 3. The extrachromosomal rDNA palindrone. There are about 100 copies of the 88 Kb palindromic DNA in each nucleus that encodes the precursor to 18S and 28S ribosomal RNA and 5S ribosomal RNA (long and short arrows, respectively). No other genes are found on these extrachromosomal elements. About 5 kb from the central 42 bp region of asymmetry (vertical bar) there are G/C rich direct and inverted repeats that may facilitate asssociation of the arms (light boxes). The extrachromosomal elements terminate in telomeres with 4 repeats of a 29bp and multiple copies of the simple sequence AG_{1-8} (stippled boxes) (Sucgang *et al.,* 2003). A master copy resides near the *erkb* locus on chromsome 4 which carries the central region of asymmetry and a complete arm carrying the rDNA genes. The other arm stops just after the repeat region. Telomeric sequences are not found in the chromosomal copy but the whole region is bounded by complex repeats (blue boxes) which may serve as guides for replicating a single strand copy. This copy could snap back on itself since it is palindromic. Extension of the short arm would generate the full 88 kb sequence and rectify the arms. Unfolding, telomere addition and replication could then form a double stranded extrachromosomal element which could be further replicated to maintain the element at 100 copies per nucleus [adapted from Eichinger *et al*., 2005].

that both ends of each of the 6 chromosomes have these short portions of the palindrome adjacent to a repetitive element.

Duplications. Long range restriction mapping of genes on chromosome 2 uncovered a 750 kb inverted duplication that was present in the axenic strain AX3 and its direct descendent, strain AX4, but not in the parental strain NC-4 or another axenic strain, AX2, that was isolated at about the same time as AX3 from strain NC-4 (Loomis *et al.*, 1995). It appears that when the population of strain NC4 was mutagenized with nitrosoguanidine and selected for axenic growth in HL5 medium, a process similar to the "breakage-fusion-bridge" events described by Barbara McClintock in maize (McClintock, 1951) occurred in *Dictyostelium*. A chromosomal break occurred about 3 Mb along chromosome 2 near the *carA* gene and resulted in a duplication of 750 kb containing about 280 genes (Fig. 2). A portion of the rDNA palindrome became attached to the end of the duplication perhaps in an effort to create a telomere (Eichinger *et al*., 2005). The ends then fused to reform an enlarged chromosome 2.

There are no other duplications in this size range in the 6 chromosomes. On the other hand, there are many cases of duplication of several kb of sequence. When the coding regions were analyzed, 299 groups with two or more genes were found to be ≥ 90% identical at the nucleotide level. Most of the copies are within 10 kb of each other indicating that they resulted from tandem duplications. For example, within the 750 kb duplication on chromosome 2 there are 11 copies of a 2 kb gene encoding acetyl-CoA synthetase all within 33 kb that resulted from 10 adjacent duplications. These copies and others that occur on chromosome 2 may have been recently derived from a mitochondrial gene. Another dramatic case has been the expansion of the polyketide synthase family of genes on chromosome 5. There are 6 members of this 9 kb gene within a 85 kb region on chromosome 5 (Fig. 2). The region is flanked by pairs of polyketide synthase genes about 400 kb on either side. The nucleic acid sequences of these genes are all >90% identical indicating that they arose fairly recently in a series of at least seven duplications from a common ancestral copy. Most of the 6 introns are conserved in each member of this family ruling out insertion of retrocopies generated from mRNA. There are partial copies of the gene interspersed among the copies as well as a few unrelated genes. Many of the genes are only a kb apart and are transcribed from the same strand. Unequal crossing-over after the generation of an initial pair of genes by adjacent duplication appears

to be the most likely mechanism to generate such an array.

Considering that ~4% of the proteins encoded by the *Dictyostelium* genome are ≥98% identical in amino acid sequence over their full length with at least one other gene, the total number of gene duplications is surprisingly high. On the other hand the amount of non-coding sequence in the genome is surprisingly low. It appears that the rate of duplication is high in this organism but that the rate of deletion is even higher. Genome size is determined to a considerable extent by the ratio of duplication to deletion (Loomis and Gilpin, 1986). In the long run, *Dictyostelium* ends up with a compact genome with a considerable number of genes that can diverge and take on new functions. Genes that are adjacent to each other on the chromosomes can be inspected for evidence of duplication from a common ancestoral gene.

Many genes encoding tRNAs also appear to have undergone recent tandem duplications resulting in a total of 390 genes encoding tRNAs. There are 98 pairs of genes encoding tRNAs with the same anticodon and often the same exact sequence. The members of each pair are found within 20 kb of each other, often within 5 kb. Duplication of these genes may be the consequences of their being flanked by the non-LTR repetitive elements TRE-5 and TRE-3 which preferentially insert adjacent to tRNAs (Glockner *et al.*, 2001).

The actin family has been known for a long time to be extensive (Romans and Firtel, 1985). There are over 25 genes encoding actins that are over 80% identical in amino acid sequence. There is a region of 670 kb on chromosome 2 with 8 actin genes encoding exactly the same protein. There are 6 more copies of this gene in a 1 Mb region of chromosome 5. While there is more variability at the nucleotide level among these genes, this large family has apparently expanded significantly in relatively recent times.

At the fine scale, there have been many cases of triplet-repeat expansion resulting in homopolymer tracts of more than 25 asparagines or glutamines in many proteins. The presence of such long poly-Q tracts is surprising because they lead to the devastating Huntington's disease in humans. *Dictyostelium* seems to have found a way to make them benign. If we knew how these cells tolerate such poly-Q tracts, we might be able to harness it for treatment of afflicted humans.

The proteome

The chromosomes are all tightly packed with genes. On average, there is a gene every 2.5 kb with a coding region of 1.77 kb. That does not leave much room for regulatory regions, and in fact, most regulatory elements have been found within 500 bp upstream of the start codons. Most genes (69%) carry at least one intron and the average is 1.9 introns per gene (Eichinger *et al.*, 2005). However, the introns are small, being 146 bp on average, such that the mean size of proteins is 525 amino acids in length. The total number of amino acids encoded by the *Dictyostelium* genome is 7,086,000 which can be compared to 2,471,000 amino acids encoded by *Saccharomyces cerevisiae* and 9,838,000 amino acids encoded by the human genome. The protein complexity of *Dictyostelium* is closer to that of humans than yeast.

There are 11,945 proteins of greater than 100 amino acids encoded by the *Dictyostelium* genome, however, some of these are pseudogenes. Amost all the *Dictyostelium* genes are descendent from genes present in the common ancestor of plants and animals. However, 26 genes encode proteins with domains found in bacteria but absent in eukaryotes (Eichinger *et al.*, 2005). These are good candidates for horizontal gene transfer from bacteria in the last 500 million years. One of these genes, *thy1*, encodes a thymidylate synthetase that is completely unrelated to the thymidylate synthetase that is found highly conserved in most bacteria as well as all other eukaryotes. There is little question that the standard thymidylate synthetase was displaced by a gene that was acquired from a bacterium after *Dictyostelium* diverged from the line leading to fungi and animals (Myllykallio *et al.*, 2002). *Dictyostelium* feeds on bacteria and so is continuously exposed to bacterial genes. Since the feeding cells are also the ones that propagate the species, perhaps it is surprising that there are not more cases of gene replacement with a bacterial homolog.

Many genes inherited from the common ancestor of plants and animals duplicated and diverged to generate new genes responsible for different functions but still retained sufficient similarity in one or more domains to be recognizable using Pfam. *Dictyostelium* shares 597 protein domains with plants, fungi and animals which are not seen in any bacterial species. These appear to be eukaryotic specific protein sequences. Another 114 eukaryotic specific domains were lost in *Dictyostelium* either because they were not needed or because they were replaced by other domains. There are 4 domains present only in *Dictyostelium* and plants that appear to have been lost in fungi and metazoans. Fifty-three domains are found in *Dictyostelium*, fungi and metazoans but absent in plants. These domains either arose after the divergence of plants or were lost in the plants for which we have genome sequence information. These include 7 components of RAS/Rac/Rho signaling, 3 G-protein signaling components and 3 domains involved in clathrin function.

One of the families of interest, the histidine kinases, is found throughout the bacterial kingdom but is sparsely represented in plants. Only a single gene encoding a histidine kinase, sln1, is found in the genome of the budding yeast, *Saccharomyces cerevisiae*. There are none in metazoans. *Dictyostelium* has 15 genes encoding histidine kinases spread out over 4 chromosomes (Anjard and Loomis, 2002). In bacteria the histidine kinases function in two-component systems where they often control transcriptional regulators by phosphorelay. In yeast, Sln1p controls a MAP kinase cascade that is involved in protection from osmotic stress. In *Dictyostelium* the histidine kinases play several developmental roles. The best studied histidine kinase, DhkA, provides the paradigm for the family. It autophosphorylates a histidine in a highly conserved catalytic domain using ATP as substrate. The phosphate is relayed to an aspartate moiety near the carboxy terminal receptor domain that is also highly conserved across diverse species (Wang *et al.*, 1996). From the aspartate the phosphate is relayed to a histidine in the small H2 protein encoded by *rdeA*. There appears to only a single gene encoding an H2

protein in *Dictyostelium* and, since phosphorelay from histidine kinases always passes through an H2, all the other histidine kinases must impinge on RdeA. From there, the phosphate can be relayed to aspartate moieties found in the response regulator domains of a variety of proteins. One of the major targets for phosphorelay is the cAMP phosphodiesterase RegA that controls levels of internal cAMP (Shaulsky *et al.*, 1996; 1998; Thomason *et al.*, 1998; 1999). When an aspartate in its response regulator domain near the N-terminus is phosphorylated, the phosphodiesterase activity is stimulated up to 5 fold.

DhkA is a membrane imbedded protein with an extracellular domain that binds a small peptide, SDF-2, generated from the *acbA* product. In the absence of ligand binding, DhkA acts as a constitutive kinase in a heterologous yeast assay (Anjard and Loomis, 2005). When SDF-2 is bound to DhkA, the kinase activity is inhibited and phosphate is no longer passed down the relay pathway and DhkA becomes an active phosphatase removing phosphates from RegA. Similar conversion of constitutive histidine kinase activity to phosphatase activity upon ligand binding has been observed in the Lux system of the marine bacterium *Vibrio harveyi* (Freeman and Bassler, 1999). Secretion of SDF-2 into the extracellular space during *Dictyostelium* culmination results in DhkA drawing phosphate off RegA and a rapid reduction in phosphodiesterase activity. The resulting rise in cAMP and PKA activity subsequently leads to encapsulation of spores. Another member of the histidine kinase family, DhkB, acts synergistically with DhkA to keep the spores dormant (Wang *et al.*, 1999). A third histidine kinase, DhkC, appears to act early during development to stimulate RegA activity since strains carrying null mutations in *dhkC* show precocious aggregation (Singleton *et al.*, 1998).

Protection from osmotic stress is mediated by another member of the histidine kinase family, *dokA* (Schuster *et al.*, 1996). This enzyme acts predominantly as a phosphatase drawing phosphate out of the phosphorelay system. As expected the level of cAMP rises rapidly upon osmotic shock as the phosphates are drawn off RegA. This effect is not seen in *dokA*$^-$ null mutants.

The physiological roles of the other 11 members of this family remain to be determined, however, the facile molecular genetics and robust biochemistry of the *Dictyostelium* system make such studies attractive.

Transcription

Knowing the sequence of the genome makes it possible to carry out genome-wide expression analyses. Portions of each gene can be robotically positioned on microarrays for hybridization to probes made from mRNA isolated at various stages in the life cycle. In fact, even before the genome was completed, it became possible to profile a large number of expressed genes by using cDNAs sequenced by the Japanese EST Project (Morio *et al.*, 1998). About 4,000 genes were represented by 6,000 targets (Van Driesche *et al.*, 2002; Iranfar *et al.*, 2003). Most of the genes expressed in growing cells were found to be turned off or to stay constant during development. There is no increase in mass and very little cell division during development which proceeds in the absence of nutrients. After 4 hours of development, almost all of the genes encoding ribosomal proteins are repressed and their mRNAs decay exponentially with a half life of 4 hours. Very little ribosomal protein mRNA is left by the end of development at 24 hours. On the other hand, about 300 genes are expressed at high rates at one stage or another of development. Expression of all of these genes is dependent on the DNA sequence specific binding protein CbfA which is essential for early transcription of the gene encoding adenylyl cyclase, *acaA* (Winkler *et al.*, 2004). Expression of later developmental genes is controlled by intercellular cAMP signaling. One of the genes expressed during aggregation, *gbfA*, encodes a DNA binding protein that is necessary for transcription of later genes including *srfA* that makes a MADS-box transcription factor essential for expression of many late prespore genes (Escalante *et al.*, 2004). The dependent sequence that gates and regulates progression through the stages of development can be recognized in the patterns of expression of the developmental genes in wild type and selected mutant strains.

The genome sequence has significantly expanded the number of developmental genes that can be used to define the specific stages. Moreover, the genome sequence has uncovered over 150 genes that are expressed exclusively in one cell type but not others. *In situ* hybridization with these cell-type specific markers have defined several new prestalk cell types (Maeda *et al.* 2003; Maruo *et al.*, 2004).

There are 107 genes which have the domain structures indicating that they may act as transcription factors. These include members of the Myb, GATA, MADS-box and homeodomain families. However, considering that there are at least a dozen developmental stages and up to 6 distinct patterns of expression of cell-type specific genes, the total number of transcription factors recognizable in the genome appears to be rather low. It is possible that there are other DNA binding proteins that have diverged so much that they cannot be found by primary sequence comparisons alone. On the other hand, there may be only a small number of transcription factors specific to developmental genes. Post-translation modifications, such as phosphorylation, together with combinatorial use of multiple transcription factors may provide sufficient specificity to generate the complexity of a fruiting body. This point of view is supported by the paucity of developmental mutations affecting DNA binding proteins that have been found in saturation screens for morphological aberrations. Transcription factors that play essential roles in growing cells and then are recruited to specific developmental roles would escape detection in the mutagenic screen that have been carried out to date.

Developmental genes are found scattered over all 6 chromosomes with little clustering that might result from long range chromosomal regulation. There are a few clusters of prespore or prestalk specific genes but they are interspersed with vegetative and non-cell type specific genes. It would appear that genes are generally regulated on an individual basis. The *cis*-regulatory regions of a variety of developmental genes have been delimited by molecular genetic manipulations and almost all have been found to lie within 500 bp of the start of transcription. However, the intergenic regions are mostly

dull sequences made up of runs of As and Ts that appear frequently throughout the genome. There are a few scattered Gs and Cs that might provide the necessary sequence specificity for 8 base sites recognizable by transcription complexes, but the rules are not yet clear. It may be necessary to empirically determine the degrees of freedom in the regulatory sequences for co-regulated genes to get an idea of how the cells are able to accurately restrict expression of sets of genes to a unique developmental stage or a specific cell type. Once such information is available for a representative set of genes, the genome sequence can be scanned for occurence of potential *cis*-regulatory sites adjacent to all the rest of the genes. It is now easy to download the upstream sequence for each gene in the genome, a task that was laborious and frustrating before the sequence of the genome was completed.

Conclusions

Just as geographical maps give the lay of the land, genetic maps give the hereditary space of the organism. Knowing the complete genomic sequence for an organism brings it into the genomic arena where it can join other organisms with fully sequenced genomes for comparative studies. Detailed evolutionary analyses gain immeasurably from whole genome data sets. Absence of a set of genes, such as those involved in cilia formation in organisms lacking cilia, can only be determined when the complete genome is available. Genes unique to an organism often have much to tell about the organism's life style and physiology. However, at present about a third of the genes in every phylum with a fully sequence genome are unique to that phylum and most are of unknown function. There is still much to do at the bench and in the field before we will know the selective advantages that result from having these genes.

A powerful approach that has been used for the budding yeast is to sequence the genomes of relatively closely related species. If they have been separated by about 10 million years, many dispensible sequences will have randomly mutated while essential sequences will have been conserved. Aligning three or four related genomes can draw attention to genes and their regulatory regions and avoid distraction with irrelevant sequences (Kellis *et al.*, 2003; 2004). It will be some time before this will be possible for *Dictyostelium* due to the size of the genomes of related species (around 35 Mb) and the difficulty of sequencing high A/T genomes, a property that they share with *Dictyostelium*. The most closely related species with an on-going genome project, *Entamoeba histolytica*, is too distantly related to provide much detailed comparative information. However, as sequencing technology improves, it may be possible to sequence dictyostelids such as *Dictyostelium purpureum*, *Dictyostelium mucoroides* and *Polysphondylium pallidum*. Comparisons among these genomes will undoubtedly uncover surprises and solidify what are now mostly conjectures.

Although we now have the sequence of the 34 Mb *Dictyostelium* genome, as with any genome, there are mis-assemblies, missing genes and errors in the sequence. They will be corrected over time as they become apparent or significant. Refinement is greatly facilitated by the establishment of a central clearing house that is provided by the team at dictyBase. Their efforts have already made the genome accessible for inspection and their dedication to curration and genomic analyses has taken the information to the next level of accuracy. It can be easily accessed at dictybase.org.

Even as it stands, the *Dictyostelium* genome sequence is a most valuable resource for those who work with this system. They no longer need to sequence each gene as it is encountered in genetic or biochemical studies. Construction of complex genotypes for specific questions is greatly facilitated by knowing the chromosomal position of each gene. We can expect the pace of experiments to accelerate and lead to new insights on fundamental processses that are often best studied in *Dictyostelium*.

References

Anjard, C., and Loomis, W. F. (2002). Histidine kinases of *Dictyostelium*. In Histidine Kinases in Signal Transduction, R. D. a. M. Inouye, ed. (San Diego, CA, Academic Press), pp. 421–438.

Anjard, C. and Loomis, W.F. (2005) Peptide signaling during terminal differentiation of *Dictyostelium.* Proc. Natl. Acad. Sci. *102*: 7607–7611.

Cox, E. C., Vocke, C. D., Walter, S., Gregg, K. Y., and Bain, E. S. (1990). Electrophoretic karyotype for *Dictyostelium discoideum*. Proc. Natl. Acad. Sci. USA *87*, 8247–8251.

Dear, P., and Cook, P. (1993). Happy mapping: linkage mapping using a physical analogue of meiosis. Nucleic Acids Res. *21*, 113–120.

Eichinger, L., Pachebat, J., Glockner, G., Rajandream, M-A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., Tunggal, B., Kummerfeld, S., Madera, M., Konfortov, A., Rivero, F., Bankier, A., Lehmann, R., Hamlin, N., Davies, R. Gaudet, P., Fey, P., Pilcher, K., Chen, G., Saunders, D., Sodergen, E., Davis, P.,Kehornou, 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., Mohan, M., Saito, T., Buchrieser, C., Wardroper, A., Felder, Thangavelu, M., Johnson, D., Knights, A., Loulseged, H., Mungall, K., M., Oliver, K., Price, C., Quail, M., Urushihara, H., Hernadez, 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, E., Chisholm, R., Gibbs, R., Loomis, W. F., Platzer, M., Kay, R.R., Williams, J., Dear, P., Noegel, A.A., Barrell, B., and Kuspa, A. (2005) The genome of the social amoeba *Dictyostelium discoideum.* Nature *435*: 43–57.

Emery, H. S., and Weiner, A. M. (1981). An irregular satellite sequence is found at the termini of the linear extrachromosomal rDNA in *Dictyostelium discoideum*. Cell *26*, 411–419.

Erdos, G. W., Raper, K. B., and Vogen, L. K. (1973). Mating types and macrocyst formation in *Dictyostelium discoideum*. Proc. Natl. Acad. Sci. USA *70*, 1828–1830.

Escalante, R., Iranfar, N., Sastre, L., and Loomis, W. F. (2004). Identification of genes dependent on the MADS-box transcription factor SrfA in *Dictyostelium* development. Euk. Cell *3*, 564–566.

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

Freeman, J., and Bassler, B. (1999). Sequence and function of LuxU: a two-component phosphorelay protein that regulates quorum sensing in *Vibrio harveyi*. J. Bacteriol. *181*, 899–906.

Glockner, G., Eichinger, L., Szafranski, K., Pachebat, J. A., Bankier, A. T., Dear, P. H., Lehmann, D., Baumgart, C., Parra, G., Abril, J. F*., et al.* (2002). Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*. Nature *418*, 79–85.

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

Iranfar, N., Fuller, D., and Loomis, W. F. (2003). Genome-wide expression analyses of gene regulation during early development of *Dictyostelium discoideum*. Euk. Cell *2*, 664–670.

Katz, E. R., and Sussman, M. (1972). Parasexual recombination in *Dictyostelium discoideum*: selection of stable diploid heterozygotes and stable haploid segregants. Proc. Natl. Acad. Sci. USA *69*, 495–498.

Kellis, M., Birren, B., and Lander, E. (2004). Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature *428*, 617–624.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature *423*, 241–254.

Konfortov, B. A., Cohen, H. M., Bankier, A. T., and Dear, P. H. (2000). A high-resolution HAPPY map of *Dictyostelium discoideum* chromosome 6. Genome Res. *10*, 1737–1742.

Kuspa, A., and Loomis, W. F. (1994). REMI-RFLP mapping in the *Dictyostelium* genome. Genetics *138*, 665–674.

Kuspa, A., and Loomis, W. F. (1996). Ordered yeast artificial chromosome clones representing the *Dictyostelium discoideum* genome. Proc. Natl. Acad. Sci. USA *93*, 5562–5566.

Loomis, W. F. (1969). Temperature-sensitive mutants of *Dictyostelium discoideum*. J. Bacteriol. *99*, 65–69.

Loomis, W. F. (1987). Genetic tools for *Dictyostelium discoideum*. In Methods in Cell Biology, J. A. Spudich, ed. (Orlando, FL, Academic Press), pp. 31–65.

Loomis, W. F., and Gilpin, M. (1986). Multigene families and vestigial DNA. Proc. Natl. Acad. Sci. USA *83*, 2143–2147.

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

Loomis, W.F. and Kuspa, A. (1997) The genome of *Dictyostelium discoideum*. In *Dictyostelium*: a Model System for Cell and Developmental Biology, Y. Maeda, K. Inouye and I Takeuchi, eds, (University Academy Press, Tokyo). pp. 15–30.

Maeda, M., Sakamoto, H., Iranfar, N., Fuller, D., Maruo, T., Ogihara, S., Morio, T., Urushihara, H., Tanaka, Y., and Loomis, W. F. (2003). Changing patterns of gene expression in *Dictyostelium* prestalk cell subtypes recognized by in situ hybridization with genes from microarray analyses. Euk. Cell. *2*, 627–637.

Maruo, T., Sakamoto, H., Iranfar, N., Fuller, D., Morio, T., Urushihara, H., Tanaka, Y., Maeda, M. and Loomis, W. F. (2004) Control of cell type proportioning in *Dictyostelium discoideum* by Differentiation-Inducing Factor as determined by *in situ* hybridization. Euk. Cell *3*: 1241–1248.

McClintock, B. (1951). Chromosomal organization and genic expression. Cold Spring Harb. Symp. Quant. Biol. *16*, 13–47.

Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshino, R., Mitra, B. N., Pi, M., Sato, T*., et al.* (1998). The *Dictyostelium* developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development. DNA Res. *5*, 335–340.

Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P., and Liebl, U. (2002). An alternative flavin-dependent mechanism for thymidylate synthesis. Science *297*, 61–62.

Newell, P. C., Williams, K. L., Kuspa, A., and Loomis, W. F. (1993). Genetic map of *Dictyostelium*. In Genetic Maps: Locus Maps of Complex Genomes (6th edition), S. J. O'Brien, ed. (Cold Spring Harbor, CSH Laboratory Press), pp. 3.1–3.10.

Raper, K. B. (1935). *Dictyostelium discoideum*, a new species of slime mold from decaying forest leaves. J. Agr. Res. *50*, 135–147.

Robson, G. E., and Williams, K. L. (1977). The mitotic chromosomes of the cellular slime mould *Dictyostelium discoideum*: a karyotype based on Giemsa banding. J. Gen. Microbiol. *99*, 191–200.

Romans, P., and Firtel, R. A. (1985). Organization of the actin multigene family of *Dictyostelium discoideum* and analysis of variability in the protein coding regions. J. Mol. Biol. *186*, 321–335.

Schuster, S. C., Noegel, A. A., Oehme, F., Gerisch, G., and Simon, M. I. (1996). The hybrid histidine kinase DokA is part of the osmotic response system of *Dictyostelium*. EMBO J. *15*, 3880–3889.

Shaulsky, G., Escalante, R., and Loomis, W. F. (1996). Developmental signal transduction pathways uncovered by genetic suppressors. Proc. Natl. Acad. Sci. USA *93*, 15260–15265.

Shaulsky, G., Fuller, D., and Loomis, W. F. (1998). A cAMP-phosphodiesterase controls PKA-dependent differentiation. Development *125*, 691–699.

Singleton, C. K., Zinda, M. J., Mykytka, B., and Yang, P. (1998). The histidine kinase dhkC regulates the choice between migrating slugs and terminal differentiation in *Dictyostelium discoideum*. Dev. Biol. *203*, 345–357.

Sinha, U., and Ashworth, J. M. (1969). Evidence for the existence of elements of a para-sexual cycle in the cellular slime mould, *Dictyostelium discoideum*. Proc. R. Soc. Lond. B *173*, 531–540.

Sonneborn, D. R., White, G. J., and Sussman, M. (1963). A mutation affecting both rate and pattern of morphogenesis in *Dictyostelium discoideum*. Dev. Biol. *7*, 79–93.

Strassmann, J. E., Zhu, Y., and Queller, D. C. (2000). Altruism and social cheating in the social amoeba *Dictyostelium discoideum*. Nature *408*, 965–967.

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

Sussman, M. (1954). Synergistic and antagonistic interactions between morphogenetically deficient variants of the slime mould *Dictyostelium discoideum*. J. Gen. Microbiol. *10*, 110–120.

Sussman, M. (1955). 'Fruity' and other mutants of the cellular slime mould, *Dictyostelium discoideum*: a study of developmental aberrations. J. Gen. Microbiol. *13*, 295–309.

Sussman, R. R. (1961). A method for staining the chromosomes of *Dictyostelium discoideum* myxamoebae in the vegetative stage. Exp. Cell Res. *24*, 154–155.

Sussman, R. R., and Sussman, M. (1953). Cellular differentiation in Dictyosteliaceae: heritable modifications of the developmental pattern. Ann. N. Y. Acad. Sci. *56*, 949–960.

Thomason, P. A., Traynor, D., Cavet, G., Chang, W. T., Harwood, A. J., and Kay, R. R. (1998). An intersection of the cAMP/PKA and two-component signal transduction systems in *Dictyostelium*. EMBO J. *17*, 2838–2845.

Thomason, P. A., Traynor, D., Stock, J. B., and Kay, R. R. (1999). The RdeA-RegA system, a eukaryotic phospho-relay controlling cAMP breakdown. J. Biol. Chem. *274*, 27379–27384.

VanDriessche, N., Shaw, C., Katoh, M., Morio, T., Sucgang, R., Ibarra, M., Kuwayama, H., Saito, T., Urushihara, H., Maeda, M*., et al.* (2002). A transcriptional profile of multicellular development in *Dictyostelium discoideum*. Development *129*, 1543–1552.

Wallace, M. A., and Raper, K. B. (1979). Genetic exchanges in the macrocysts of *Dictyostelium discoideum*. J. Gen. Microbiol. *113*, 327–337.

Wang, N., Shaulsky, G., Escalante, R., and Loomis, W. F. (1996). A two-component histidine kinase gene that functions in *Dictyostelium* development. EMBO J. *15*, 3890–3898.

Wang, N., Soderbom, F., Anjard, C., Shaulsky, G., and Loomis, W. F. (1999). SDF-2 induction of terminal differentiation in *Dictyostelium discoideum* is mediated by the membrane-spanning sensor kinase DhkA. Mol. Cell Biol. *19*, 4750–4756.

Welker, D. L., Hirth, K. P., Romans, P., Noegel, A., Firtel, R. A., and Williams, K. L. (1986). The use of restriction fragment length polymorphisms and DNA duplications to study the organinzation of the actin multigene family in *Dictyostelium discoideum*. Genetics *112*, 27–42.

Williams, K. L., and Barrand, P. (1978). Parasexual genetics in the cellular slime mould *Dictyostelium discoideum*: haploidisation of diploid strains using ben late. FEMS Microbiol. Lett. *4*, 155–159.

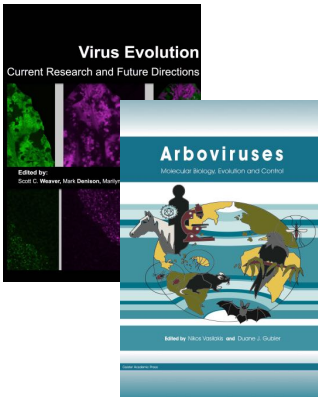
Winckler, T., Iranfar, N., Beck, P., Jennes, I., Baik, U., Loomis, W. F., and Dingermann, T. (2004). CbfA, the C-module DNA binding factor, plays an essential role in the initiation of development of *Dictyostelium*. Eukary. Cell *3*: 1349–1358.

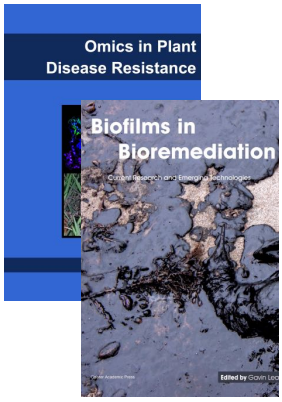
Yanagisawa, K., Loomis, W. F., and Sussman, M. (1967). Developmental regulation of the enzyme UDP-galactose polysaccharide transferase. Exp. Cell Res. *46*, 328–334.

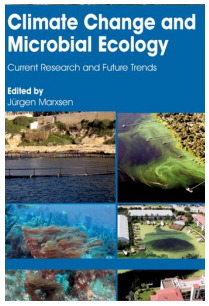
Further Reading

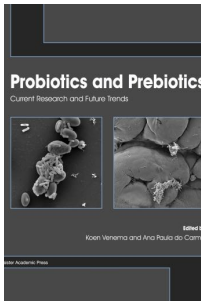
Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at **caister.com**

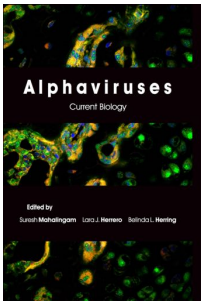
- **MALDI-TOF Mass Spectrometry in Microbiology**
  Edited by: M Kostrzewa, S Schubert (2016)
  www.caister.com/malditof
- ***Aspergillus* and *Penicillium* in the Post-genomic Era**
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  www.caister.com/aspergillus2
- **The Bacteriocins: Current Knowledge and Future Prospects**
  Edited by: RL Dorit, SM Roy, MA Riley (2016)
  www.caister.com/bacteriocins
- **Omics in Plant Disease Resistance**
  Edited by: V Bhadauria (2016)
  www.caister.com/opdr
- **Acidophiles: Life in Extremely Acidic Environments**
  Edited by: R Quatrini, DB Johnson (2016)
  www.caister.com/acidophiles
- **Climate Change and Microbial Ecology: Current Research and Future Trends**
  Edited by: J Marxsen (2016)
  www.caister.com/climate
- **Biofilms in Bioremediation: Current Research and Emerging Technologies**
  Edited by: G Lear (2016)
  www.caister.com/biorem
- **Microalgae: Current Research and Applications**
  Edited by: MN Tsaloglou (2016)
  www.caister.com/microalgae
- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**
  Edited by: H Shintani, A Sakudo (2016)
  www.caister.com/gasplasma
- **Virus Evolution: Current Research and Future Directions**
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  www.caister.com/virusevol
- **Arboviruses: Molecular Biology, Evolution and Control**
  Edited by: N Vasilakis, DJ Gubler (2016)
  www.caister.com/arbo
- ***Shigella*: Molecular and Cellular Biology**
  Edited by: WD Picking, WL Picking (2016)
  www.caister.com/shigella
- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**
  Edited by: AM Romaní, H Guasch, MD Balaguer (2016)
  www.caister.com/aquaticbiofilms
- **Alphaviruses: Current Biology**
  Edited by: S Mahalingam, L Herrero, B Herring (2016)
  www.caister.com/alpha
- **Thermophilic Microorganisms**
  Edited by: F Li (2015)
  www.caister.com/thermophile

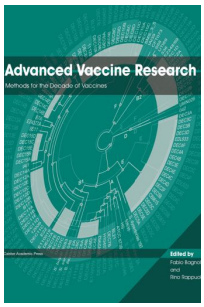


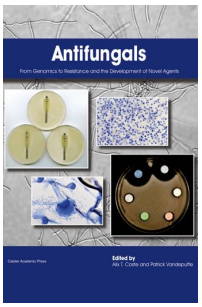












- **Flow Cytometry in Microbiology: Technology and Applications**
  Edited by: MG Wilkinson (2015)
  www.caister.com/flow
- **Probiotics and Prebiotics: Current Research and Future Trends**
  Edited by: K Venema, AP Carmo (2015)
  www.caister.com/probiotics
- **Epigenetics: Current Research and Emerging Trends**
  Edited by: BP Chadwick (2015)
  www.caister.com/epigenetics2015
- ***Corynebacterium glutamicum*: From Systems Biology to Biotechnological Applications**
  Edited by: A Burkovski (2015)
  www.caister.com/cory2
- **Advanced Vaccine Research Methods for the Decade of Vaccines**
  Edited by: F Bagnoli, R Rappuoli (2015)
  www.caister.com/vaccines
- **Antifungals: From Genomics to Resistance and the Development of Novel Agents**
  Edited by: AT Coste, P Vandeputte (2015)
  www.caister.com/antifungals
- **Bacteria-Plant Interactions: Advanced Research and Future Trends**
  Edited by: J Murillo, BA Vinatzer, RW Jackson, et al. (2015)
  www.caister.com/bacteria-plant
- ***Aeromonas***
  Edited by: J Graf (2015)
  www.caister.com/aeromonas
- **Antibiotics: Current Innovations and Future Trends**
  Edited by: S Sánchez, AL Demain (2015)
  www.caister.com/antibiotics
- ***Leishmania*: Current Biology and Control**
  Edited by: S Adak, R Datta (2015)
  www.caister.com/leish2
- ***Acanthamoeba*: Biology and Pathogenesis (2nd edition)**
  Author: NA Khan (2015)
  www.caister.com/acanthamoeba2
- **Microarrays: Current Technology, Innovations and Applications**
  Edited by: Z He (2014)
  www.caister.com/microarrays2
- **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**
  Edited by: D Marco (2014)
  www.caister.com/n2

Order from caister.com/order