

Table S1. Non-NCBI databases

Kingdom Animalia databases	
<i>D. discoideum</i> , <i>E. histolytica</i> , <i>E. tenella</i> , <i>Leishmania</i> species, <i>Plasmodium</i> species, <i>Trypanosome</i> species	http://www.genedb.org/
<i>C. intestinalis</i> , <i>L. bicolor</i> , <i>M. graminicola</i> , <i>N. gruberi</i> , <i>N. haematococca</i> , <i>N. vectensis</i> , <i>O. tauri</i> , <i>P. stipitis</i> , <i>T. pseudonana</i> , <i>P. tricornutum</i>	http://genome.jgi-psf.org/Cioin2/Cioin2.home.html
<i>B. bovis</i> , <i>B. malayi</i> , <i>C. parvum</i> , <i>D. discoideum</i> , <i>E. histolytica</i> , <i>E. tenella</i> , <i>Leishmania</i> species, <i>N. caninum</i> , <i>Plasmodium</i> species, <i>S. neurona</i> , <i>S. mansoni</i> , <i>T. vaginalis</i> , <i>Trypanosoma</i> species, <i>T. gondii</i>	http://tigrblast.tigr.org/tgi/
<i>B. malayi</i> , <i>B. bigemina</i> , <i>D. discoideum</i> , <i>E. histolytica</i> , <i>E. tenella</i> , <i>H. contortus</i> , <i>Leishmania</i> species, <i>Plasmodium</i> species, <i>S. mansoni</i> , <i>S. scrofa</i> , <i>Trypanosome</i> species, <i>T. annulata</i> , <i>T. parva</i>	http://www.sanger.ac.uk/DataSearch/blast.shtml
<i>A. gambiae</i>	http://www.genoscope.cns.fr/externe/English/Projets/Projet_AK/organisme_AK.html
<i>C. merolae</i>	http://merolae.biol.s.u-tokyo.ac.jp/
<i>C. parvum</i>	http://cryptodb.org/cryptodb/
<i>D. discoideum</i>	http://dictybase.org/
Diatom ESTs	http://avesthagen.sznbowler.com/
<i>E. tenella</i>	http://www.genedb.org/genedb/etenella/
<i>G. lamblia</i>	http://gmod.mbl.edu/perl/site/giardia?page=intro
<i>M. domestica</i>	http://genome.ucsc.edu/cgi-bin/hgBlat
<i>P. tetraurelia</i>	http://www.genoscope.cns.fr/externe/Francais/Projets/Projet_FN/
<i>Plasmodium</i> species	http://www.plasmodb.org/plasmo/home.jsp
<i>T. thermophila</i>	http://www.ciliate.org/
<i>T. gondii</i>	http://www.toxodb.org/toxo-release4-0/home.jsp
<i>T. nigroviridis</i>	http://www.genoscope.cns.fr/externe/tetranew/
<i>T. rubripes</i>	http://www.fugu-sg.org/project/info.html
Cyanobacteria	http://www.kazusa.or.jp/cyano/cyano.html
Kingdom Plantae databases	
Plant species	http://www.sgn.cornell.edu/tools/blast/index.pl
Plant species	http://www.plantgdb.org/
<i>C. reinhardtii</i>	http://www.chlamy.org/
<i>O. tauri</i>	http://genome.jgi-psf.org/Ostta4/Ostta4.home.html
<i>O. tauri</i>	http://bioinformatics.psb.ugent.be/genomes.php
<i>Z. mays</i>	http://www.plantgdb.org/ZmGDB/

The non-NCBI databases used to search for laforin orthologues and the organism's genome in each database are listed. Many of the databases were found by performing a Google search of the organism's genus name and genome.

Table S2. Accession numbers for laforin orthologues

Organism	Database	Accession number
<i>C. merolae</i>	<i>C. merolae</i> Genome Project	CMT465C
<i>E. tenella</i>	GeneDB	Et_v1_Twnscn_Contig6817.tmp13
<i>G. gallus</i>	GenBank/EMBL/DDBJ	NP_001026240
<i>H. sapiens</i>	GenBank/EMBL/DDBJ	NP_005661
<i>P. tetraurelia</i>	Genoscope	GSPATT00028736001
<i>T. gondii</i>	ToxoDB	TgTwinScan_3925
<i>T. nigroviridis</i>	GenBank/EMBL/DDBJ	CAG03589
<i>T. thermophila</i>	GenBank/EMBL/DDBJ	EAR89845
<i>X. laevis</i>	GenBank/EMBL/DDBJ	AAH73202

Table S3. Small subunit ribosomal RNA accession numbers

Organism	Accession number
<i>A. thaliana</i>	X16077
<i>C. elegans</i>	AY268117
<i>C. merolae</i>	AB158483
<i>C. paradoxa</i>	NC_001675.1
<i>C. parvum</i>	AF093489
<i>C. reinhardtii</i>	M327083
<i>D. melanogaster</i>	M21017.1
<i>E. coli</i>	Z83205
<i>E. histolytica</i>	AF149911
<i>E. huxleyi</i>	X82156
<i>E. tenella</i>	U67121
<i>G. gallus</i>	AF173612
<i>G. theta</i>	NC_000926.1
<i>H. sapiens</i>	X03205
<i>Nostoc</i> species	NC_003272.1
<i>P. falciparum</i>	M19172
<i>P. tetraurelia</i>	X03772
<i>S. cerevisiae</i>	Z75578
<i>T. cruzi</i>	AF303660
<i>T. gondii</i>	X68523
<i>T. nigroviridis</i>	chrUn_random:61030224..62171431—Tetradon Genome Browser
<i>T. thermophila</i>	X56165
<i>X. laevis</i>	X04025

Listed on the left are the organisms from the phylogeny in Fig. 2, and on the right are the accession numbers for the small subunit ribosomal RNA genes. All of the accession numbers are from NCBI GenBank/EMBL/DDBJ unless otherwise noted.

Table S4. Genomes investigated for the presence of laforin

	<i>Coccidioides immitis</i>	Non-red algal descent
	<i>Coprinopsis cinerea</i>	Non-red algal descent
	<i>okayama</i>	
	<i>Cryptococcus neoformans</i> sp.	Non-red algal descent
	<i>Debaryomyces hansenii</i>	Non-red algal descent
	<i>Encephalitozoon cuniculi</i>	Non-red algal descent
	<i>Eremothecium gossypii</i>	Non-red algal descent
	<i>Gibberella moniliformis</i>	Non-red algal descent
	<i>Gibberella zea</i>	Non-red algal descent
	<i>Kluyveromyces lactis</i>	Non-red algal descent
	<i>Kluyveromyces waltii</i>	Non-red algal descent
	<i>Lodderomyces elongisporus</i>	Non-red algal descent
	<i>Magnaporthe grisea</i>	Non-red algal descent
	<i>Neosartorya fischeri</i>	Non-red algal descent
	<i>Neurospora crassa</i>	Non-red algal descent
	<i>Phaeosphaeria nodorum</i>	Non-red algal descent
	<i>Phanerochaete chrysosporium</i>	Non-red algal descent
	<i>Pichia guilliermondii</i>	Non-red algal descent
	<i>Pichia stipitis</i>	Non-red algal descent
	<i>Pneumocystis carni</i>	Non-red algal descent
	<i>Rhizopus oryzae</i>	Non-red algal descent
	<i>Saccharomyces bayanus</i>	Non-red algal descent
	<i>Saccharomyces castellii</i>	Non-red algal descent
	<i>Saccharomyces cerevisiae</i>	Non-red algal descent
	<i>Saccharomyces kluyveri</i>	Non-red algal descent
	<i>Saccharomyces kudriavzevii</i>	Non-red algal descent
	<i>Saccharomyces mikatae</i>	Non-red algal descent
	<i>Saccharomyces paradoxus</i>	Non-red algal descent
	<i>Schizosaccharomyces japonicus</i>	Non-red algal descent
	<i>Schizosaccharomyces pombe</i>	Non-red algal descent
	<i>Sclerotinia sclerotiorum</i>	Non-red algal descent
	<i>Sporobolomyces roseus</i>	Non-red algal descent
	<i>Trichoderma reesei</i>	Non-red algal descent
	<i>Uncinocarpus reesii</i>	Non-red algal descent
	<i>Ustilago maydis</i>	Non-red algal descent
	<i>Yarrowia lipolytica</i>	Non-red algal descent
Mesomycetozoa		
Choanomonada		
Metazoa		
	<i>Aedes aegypti</i>	Arthropod, lacks floridean starch/LBs
	<i>Anopheles gambiae</i>	Arthropod, lacks floridean starch/LBs
	<i>Apis mellifera</i>	Arthropod, lacks floridean starch/LBs
	<i>Aplysia californica</i>	Urochordate, lacks floridean starch/LBs
	<i>Bombyx mori</i>	Arthropod, lacks floridean starch/LBs
	<i>Bos taurus</i>	Mammal, has laforin
	<i>Caenorhabditis briggsae</i>	Nematode, lacks floridean starch/LBs
	<i>Caenorhabditis elegans</i>	Nematode, lacks floridean starch/LBs
	<i>Caenorhabditis remanei</i>	Nematode, lacks floridean starch/LBs
	<i>Canis familiaris</i>	Mammal, has laforin
	<i>Cavia porcellus</i>	Mammal, incomplete genome, has laforin
	<i>Ciona intestinalis</i>	Urochordate, lacks floridean starch/LBs
	<i>Ciona savignyi</i>	Urochordate, lacks floridean starch/LBs
	<i>Danio rerio</i>	Osteichthyes, has laforin
	<i>Dasyurus novemcinctus</i>	Mammal, incomplete genome, has laforin
	<i>Drosophila ananassae</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila erecta</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila grimshawi</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila mojavensis</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila persimilis</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila pseudoobscura</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila sechellia</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila simulans</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila virilis</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila willistoni</i>	Arthropod, lacks floridean starch/LBs
	<i>Drosophila yakuba</i>	Arthropod, lacks floridean starch/LBs
	<i>Echinops telfairi</i>	Mammal, incomplete genome, has laforin
	<i>Felis catus</i>	Mammal, has laforin
	<i>Glossina morsitans</i>	Arthropod, lacks floridean starch/LBs
	<i>Gallus gallus</i>	Aves, has laforin

		<i>Haemonchus contortus</i>	Roundworm, lacks floridean starch/LBs
		<i>Homo sapiens</i>	Mammal, has laforin
		<i>Loxodonta africana</i>	Mammal, incomplete genome, has laforin
		<i>Macaca mulatta</i>	Mammal, has laforin
		<i>Monodelphis domestica</i>	Mammal, incomplete genome, has laforin
		<i>Mus musculus</i>	Mammal, has laforin
		<i>Myotis lucifugus</i>	Mammal, incomplete genome, has laforin
		<i>Ornithorhynchus anatinus</i>	Mammal, incomplete genome, has laforin
		<i>Oryctolagus cuniculus</i>	Mammal, incomplete genome, has laforin
		<i>Oryzias latipes</i>	Osteichthyes, has laforin
		<i>Otolemur garnettii</i>	Mammal, incomplete genome, has laforin
		<i>Pan troglodytes</i>	Mammal, incomplete genome, has laforin
		<i>Pongo pygmaeus</i>	Mammal, incomplete genome, has laforin
		<i>Rattus norvegicus</i>	Mammal, has laforin
		<i>Schistosoma mansoni</i>	Flatworm, lacks floridean starch/LBs
		<i>Sorex araneus</i>	Mammal, incomplete genome, has laforin
		<i>Strongylocentrotus purpuratus</i>	Urochordate, lacks floridean starch/LBs
		<i>Sus scrofa</i>	Mammal, has laforin
		<i>Takifugu rubripes</i>	Osteichthyes, has laforin
		<i>Tetraodon nigroviridis</i>	Osteichthyes, has laforin
		<i>Tribolium castaneum</i>	Arthropod, lacks floridean starch/LBs
		<i>Xenopus laevis</i>	Amphibian, has laforin
Rhizaria	Cercozoa	<i>Phytophthora infestans</i>	Plant pathogen, lacks floridean starch
		<i>Phytophthora ramorum</i>	Plant pathogen, lacks floridean starch
		<i>Phytophthora sojae</i>	Plant pathogen, lacks floridean starch
Archaeplastida	Haplosporidia		
	Foraminifera		
	<i>Gromia</i>		
	Radiolaria		
	Glaucophyta	<i>Cyanidioschyzon merolae</i>	Has laforin
	Rhodophyceae	<i>Galdieria sulphuraria</i>	<u>Incomplete genome, likely has laforin</u>
	Chloroplastida	<i>Arabidopsis thaliana</i>	Land plant, has SEX4
		<i>Aquilegia sp.</i>	Land plant, has SEX4
		<i>Chlamydomonas reinhardtii</i>	Land plant, has SEX4
Chromalveolata	Cryptophyceae	<i>Citrus sinensis</i>	Land plant, has SEX4
		<i>Medicago truncatula</i>	Land plant, has SEX4
		<i>Oryza sativa</i>	Land plant, has SEX4
		<i>Ostreococcus tauri</i>	Land plant, has SEX4
		<i>Phaseolus vulgaris</i>	Land plant, has SEX4
		<i>Solanum lycopersicum</i>	Land plant, has SEX4
		<i>Solanum tuberosum</i>	Land plant, has SEX4
		<i>Sorghum bicolor</i>	Land plant, has SEX4
		<i>Triticum aestivum</i>	Land plant, has SEX4
		<i>Zea mays</i>	Land plant, has SEX4
		<i>Guillardia theta</i>	<u>Nucleomorph sequenced, nuclear genome not sequenced, likely has laforin</u>
	Haptophyta	<i>Emiliania huxleyi</i>	Phytoplankton, lacks floridean starch
	Stramenopiles	<i>Thalassiosira pseudonana</i>	Diatom, lacks floridean starch
	Alveolata	<i>Phaeodactylum tricornutum</i>	Diatom, lacks floridean starch
		<i>Babesia bovis</i>	Lacks floridean starch
		<i>Babesia bigemina</i>	Lacks floridean starch
		<i>Babesia malayi</i>	Lacks floridean starch
		<i>Cryptosporidium parvum</i>	Lacks mitochondrion
		<i>Cryptosporidium hominis</i>	Lacks mitochondrion
		<i>Eimeria tenella</i>	Has laforin
		<i>Neospora caninum</i>	<u>Incomplete genome, likely has laforin</u>
		<i>Paramecium tetraurelia</i>	Has laforin
		<i>Plasmodium berghei</i>	Lacks floridean starch
		<i>Plasmodium chabaudi</i>	Lacks floridean starch
		<i>Plasmodium falciparum</i>	Lacks floridean starch
		<i>Plasmodium gallinaceum</i>	Lacks floridean starch
		<i>Plasmodium knowlesi</i>	Lacks floridean starch
		<i>Plasmodium reichenowi</i>	Lacks floridean starch
		<i>Plasmodium vivax</i>	Lacks floridean starch
		<i>Plasmodium yoelii yoelii</i>	Lacks floridean starch
		<i>Sarcocystis neurona</i>	<u>Incomplete genome, likely has laforin</u>
		<i>Theileria annulata</i>	Lacks floridean starch
		<i>Theileria parva</i>	Lacks floridean starch

		Tetrahymena thermophila	Has laforin
		Toxoplasma gondii	Has laforin
Excavata	Fornicata <i>Malawimonas</i>	<i>Giardia lamblia</i>	Lacks mitochondrion, non-red algal descent
	Parabasalia	<i>Trichomonas vaginalis</i>	Lacks mitochondrion, non-red algal descent
	Preaxostyla		
	Jakobida		
	Heterolobosea		
	Euglenozoa	<i>Leishmania braziliensis</i> <i>Leishmania infantum</i> <i>Leishmania major</i> <i>Trypanosoma brucei</i> <i>Trypanosoma congolense</i> <i>Trypanosoma cruzi</i> <i>Trypanosoma gambiense</i>	Non-red algal descent Non-red algal descent Non-red algal descent Non-red algal descent Non-red algal descent Non-red algal descent Non-red algal descent
Prokaryotes - Archaea and Eubacteria	All 656 microbial genomes in NCBI		Lack mitochondrion, no floridean starch, non-red algal descent

The genome of each organism was searched for laforin using the appropriate database (Table S1). If laforin was absent, an extensive literature search was performed on the organism to determine which of the three criteria it lacked: red algal descent, mitochondrion, and/or floridean starch/LBs. If the organism lacked laforin, at least one of the three criteria that it lacks is presented beside its name. The organism names of genomes containing laforin are bold. The organism names of genomes that are nearing completion and that contain laforin based on our predictions are underlined. The phrase “incomplete genome, has laforin” refers to organisms with incomplete genomes but in which a partial CBM and DSP corresponding to laforin was found. Organism classification is based on Adl et al. (2005).

References

Adl S.M., A.G. Simpson, M.A. Farmer, R.A. Andersen, O.R. Anderson, J.R. Barta, S.S. Bowser, G. Brugerolle, R.A. Fensome, S. Fredericq, et al. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* 52:399–451.

Table S5. Accession numbers for AMPK β -GBD proteins and SEX4 orthologues

AMPK β -GBD proteins			
Gene	Organism	Database	Accession number
AMPK β 1-Hs	<i>H. sapiens</i>	GenBank/EMBL/DDBJ	NP_006244
AMPK β 1-At	<i>A. thaliana</i>	GenBank/EMBL/DDBJ	NP_197615
Gal83-Sc	<i>S. cerevisiae</i>	GenBank/EMBL/DDBJ	NP_010944
GBE-Ec	<i>E. coli</i>	GenBank/EMBL/DDBJ	NP_417890
	<i>A. thaliana</i>	GenBank/EMBL/DDBJ	AAN28817
	<i>Aquilegia</i> species	GenBank/EMBL/DDBJ	DT739859 and DT764798
	<i>C. reinhardtii</i>	JGI	149756
	<i>C. sinensis</i>	GenBank/EMBL/DDBJ	CV886681
	<i>M. truncatula</i>	GenBank/EMBL/DDBJ	BG581666 and AW689683
	<i>O. sativa</i>	GenBank/EMBL/DDBJ	ABF93554
	<i>P. vulgaris</i>	GenBank/EMBL/DDBJ	CV538569, CB540037, and CV534719
	<i>S. lycopersicum</i>	GenBank/EMBL/DDBJ	CAC44460
	<i>S. tuberosum</i>	GenBank/EMBL/DDBJ	ABB87109
	<i>Z. mays</i>	GenBank/EMBL/DDBJ	DT642676, CO452300, and CF051889

Some protein sequences were obtained from one or multiple cDNA sequences, and, in these cases, the cDNA accession numbers are listed.