

I

**Table I Metrics of *Hypsibius dujardini* genome assemblies.**

Data Source	This work	Edinburgh	UNC
Sequencing technologies	Illumina & PacBio	Illumina	Illumina & PacBio
Genome version	nHd.3.0	nHd.2.3	tg
Scaffold number	1,421	13,202	16,175
Total Scaffold Length (bp)	104,155,103	134,961,902	212,302,995
Average Scaffold Length (bp)	73,297	10,222	13,125
Longest Scaffold Length (bp)	2,115,976	594,143	1,208,507*
Shortest Scaffold Length (bp)	1,000	500	2,002
N50 (bp) (no. scaffs in N50)	342,180 (#85)	50,531 (#701)	17,496 (#3,422)
N90 (bp) (no. scaffs in N90)	65,573 (#343)	6,194 (#3,280)	6,637 (#11,175)
CEGMA genes found (partial)	237 (240)	220 (241)	221 (235)
CEGMA gene duplication ratio	1.17 (1.23)	1.35 (1.56)	3.26 (3.53)

\* The longest scaffolds in the tg assembly are derived from bacterial contaminants.

**Table 2 Comparison of the genomes of *H. dujardini* 3.0 and *R. varieornatus* 1.1**

Assembly	<i>H. dujardini</i> 3.0		<i>R. varieornatus</i> 1.1		Difference	
GENOME	Mb	%	bp	%	Mb	%
<b>Total span</b>	104.16	-	55.83	-	48.33	
<b>Genic</b>	59.03	56.67%	31.94	57.21%	27.09	56.06%
exon span	25.25	24.24%	19.56	35.03%	5.69	11.78%
intron span	33.78	32.43%	12.38	22.17%	21.40	44.28%
<b>Intergenic</b>	45.13	43.33%	23.89	42.79%	21.23	43.94%
repeat	27.11	26.03%	10.11	18.12%	17.00	35.17%
GENES	# families	# genes	# families	# genes	#families	#genes
number of genes	11,705	19,901	9,029	13,917	2,676	5,984
number of proteins (including isoforms)		20,815		14,538		6,277
species-specific singletons	4,364	4,364	1,995	1,995	2,369	2,369
species-specific gene families	45	258	20	123	25	135
shared gene families	7,296	15,279	7,014	11,799	258	3,480
uniquely retained ancestral genes *	471	999	189	311	282	688
genes with BLAST matches to SwissProt		8,337		6,978		
genes with BLAST matches to TrEMBL		10,202		8,359		
genes with InterPro domain matches		11,227		8,633		
genes with Gene Ontology terms		7,804		6,030		
All genes	mean	median	mean	median	ratio of means	ratio of medians
gene length (bp)	2966	2131	2295	1641	1.29	1.30
exon span (bp)	1269	978	1405	1074	0.90	0.91
exon count (#)	5.94	4	6.02	4	0.99	1.00
intron span (bp)	1697	1109	889	520	1.91	2.13
intron count (#)	4.94	3	5.02	3	0.98	1.00
Single-copy Orthologues **	mean	median	mean	median	ratio of means	ratio of medians
gene length (bp)	3716	2776	2579	1929	1.44	1.44
exon span (bp)	1615	1278	1581	1253	1.02	1.02
exon count (#)	7.64	6	6.96	6	1.10	1.00
intron span (bp)	2101	1475	998	635	2.11	2.32
intron count (#)	3716	2776	2579	1929	1.44	1.44

\* Uniquely retained ancestral genes include genes shared by only one Tardigrade and at least one non-Tardigrade taxon.

\*\* Single-copy Orthologues: orthologues with CDS lengths differing by more than 20% were not considered.

**Table 3 Gene family births that support different relationships of Tardigrada**

Protein Family*	Number of proteins	Proportion of proteomes represented				Domain annotations**
		all	Nematoda (n=9)	Arthropoda (n=15)	Tardigrada (n=2)	
Synapomorphies with membership ≥ 0.7 under the Panarthropoda (Tardigrada + Arthropoda) hypothesis						
OG0000436	104	1.00	0.00	1.00	1.00	Serine proteases, trypsin domain (IPR001254)
OG0001236	54	1.00	0.00	1.00	1.00	Major facilitator superfamily associated domain (IPR024989)
OG0002592	36	1.00	0.00	1.00	1.00	Spätzle (IPR032104)
OG0006538	19	1.00	0.00	1.00	1.00	Leucine-rich repeat (IPR001611)
OG0006541	19	1.00	0.00	1.00	1.00	None
OG0006869	17	1.00	0.00	1.00	1.00	Thioredoxin domain (IPR013766)
OG0005117	27	0.88	0.00	0.93	0.50	BTB/POZ domain (IPR000210)
OG0005941	22	0.77	0.00	0.73	1.00	None
OG0006662	18	0.82	0.00	0.80	1.00	None
OG0006889	17	0.71	0.00	0.73	0.50	None
OG0006940	17	0.82	0.00	0.87	0.50	EGF-like domain (IPR000742), Laminin G domain (IPR001791)
OG0006941	17	0.71	0.00	0.67	1.00	EF-hand domain (IPR002048)
OG0006951	17	0.71	0.00	0.67	1.00	Adipokinetic hormone (IPR010475)
OG0007141	16	0.82	0.00	0.80	1.00	None
OG0007285	15	0.71	0.00	0.67	1.00	GPCR, family 2, secretin-like (IPR000832)
OG0007290	15	0.82	0.00	0.80	1.00	Allatostatin (IPR010276)
OG0007298	15	0.88	0.00	0.87	1.00	None
OG0007328	15	0.71	0.00	0.67	1.00	Sulfakinin (IPR013259)
OG0007463	14	0.77	0.00	0.73	1.00	Peptidase S1A, nudel (IPR015420), Serine proteases, trypsin domain (IPR001254), Low-density lipoprotein (LDL) receptor class A repeat (IPR002172)
OG0007689	13	0.71	0.00	0.67	1.00	Marvel domain (IPR008253)
Synapomorphies with membership ≥ 0.7 under the Tardigrada + Nematoda hypothesis						
OG0005423	26	0.82	0.89	0.00	0.50	Amidinotransferase (PF02274)
OG0006414	20	0.82	0.78	0.00	1.00	Proteolipid membrane potential modulator (IPR000612)
OG0007199	16	0.91	1.00	0.00	0.50	Zona pellucida domain (IPR001507)
OG0007812	13	0.82	0.78	0.00	1.00	None
OG0008368	11	0.82	0.78	0.00	1.00	RUN domain (IPR004012)

\* Protein families from OrthoFinder clustering at inflation value 1.5

\* Domain annotations are reported where proteins from more than one third of the proteomes in the family had that annotation. IPR = InterPro domain identifier; PF = PFam identifier.