The "Cambrian Explosion" of bilateral animals (and the Theory of Evolution)

> Karl J. Fryxell School of Systems Biology George Mason University

Osher Lifelong Learning Institute United Christian Parish, Reston VA October 14, 2013 The Geologic Time Scale

Time spans are eons, eras, periods and epochs. Ma = mya.

EC	DN	ERA PERIOD			EPOCH	Ма				
			Quaternary		Holocene	5.12.2				
					Distance	Late	-0.01 -			
					Pleistocene	Early	- 0.8 -			
				Paleogene Neogene		Late	- 1.8 -			
			Tertiary		Pliocene	Early	- 3.6 -			
		0			1-	Late	- 5.3 -			
		Cenozo			Miocene	Middle	-11.2 -			
						Early	-16.4 -			
					Olfersen	Late	-23.7 -			
					Oligocene	Early	-28.5 -			
					Eocene	Late	-33.7 -			
						Middle	-41.3 -			
						Early	-49.0 -			
					and a state of the state of the	Late	-54.8 -			
					Paleocene	Farly	-61.0 -			
		-			Late		-65.0 -			
17		U	Cretaceous		Early		-99.0 -			
i		6			Late	- 144 -				
	2	Ň	Iuraccie		Middle		- 159 -			
	υ	8	Julassic		Early	- 180 -				
-		Ö			Late		- 206 -			
-	LIC	Σ	Triassic		Middle	- 227 -				
0					Farly	- 242 -				
			Permian		Late	- 248 -				
					Early		- 256 -			
			Pennsylvan	ian		- 290 -				
			Mississippia	an	1	- 323 -				
			Madinterstationist of pitts		Late	- 354 -				
		U			Middle	Middle				
		ō			Larly	- 391 -				
		N	Silurian		Late	- 417 -				
		ĕ			Early	- 423 -				
		e	Ordovician		Late	- 443 -				
		۵.			Middle	- 458 -				
					Early	-4/0 -				
					D	- 490 -				
			Cambrian		C	- 500 -				
					B	- 512 -				
					A	- 520 -				
	12		-				- 543 -			
	oterozoio	Late								
=		The second					- 900 -			
ia		Made								
h										
E	à	(Carlot of the second s	Fairly							
ar		-								
ec.	Archean	Late								
P		Middle								
		Early								

2

The "Cambrian explosion" of animal phyla

- The Late Proterozoic is now divided into the Cryogenic (850-630 Mya, no animal fossils, several snowball earths) and the Ediacaran (630-542 Mya, a few soft-bodied animals sponges and Cnidarians (jellyfish, sea pens, etc)). The first convincing "trace fossils" (animal tracks) occur in Late Ediacaran (555-542 Mya).
- The very beginning of the Cambrian (542-485 Mya) is marked by a drop in C-13 deposition, and the appearance of radiolarians (protozoa with silica mineral exoskeletons), followed quickly by small shelled animals (molluscs) and trilobites (arthropods).
- Within the early Cambrian (542-509 Mya), a diverse array of animal shapes and sizes with hard body parts appear, suggesting a morphological "arms race" of predators and prey. Most authors agree that all 60 phyla of animals that have ever existed on earth were present in the early Cambrian.
- Atmospheric oxygen first reached modern levels in the Cambrian.

Soft-bodied animals from the Ediacaran (from South Australia and other places across the globe)



Strickberger (2000) Evolution (3rd edition) Jones & Bartlett, Sudbury, MA



Radiolarians are microscopic protozoa with silica exoskeletons They first appeared at the very beginning of the Cambrian, along with the first fossil shells!



www.radiolaria.org

Animals from the Middle Cambrian (Burgess Shale in Canada and other places across the globe)



FIGURE 10.15. Restoration of the Burgess Shale fauna. Key to the animals is as follows. Sponges: *Pirania* (1), *Vauxia* (2), *Wapkia* (3). Lobopods: *Aysheaia* (4), *Hallucigenia* (5). Anomalocaridids: *Anomalocaris* (6), *Laggania* (7). Arthropods: *Marrella* (8), *Odaraia* (9), trilobite *Olenoides* (10), *Sanctacaris* (11), *Sarotrocercus* (12). Priapulid: *Ottoia* (13). Polychaete annelid: *Canadia* (14). Chordate: *Pikaia* (15). Animals of disputed or uncertain affinity: *Amiskwia* (16), *Dinomischus* (17), *El-donia* (18), *Odontogriphus* (19), *Opabinia* (20), *Wiwaxia* (21).

10.15. modified from Briggs D., Am. Sci. 19: 130–141. © 1991 Sigma Xi. The Scientific Research Co.



Anomalocaris, a Cambrian-era predator



Cambrian fossils come in All shapes and sizes...



Cover of *Bioessays*, July 2009

Comparison of living vs. fossils from the Lower Cambrian (China), both adults and embryos

Chen (2009) *Int. J. Dev. Biol.* 53, 733-751.



Fig. 1. Animal groups with conserved body plan. (A) Optical image of late developing

Segmented worms with distinct head segments (Lower Cambrian, China)

Chen (2009) *Int. J. Dev. Biol.* 53, 733-751.



Fig. 3. The two-segmental head worm-like organisms from Lower Cambrian Maotianshan Shale deposits. (A-C) Photographs (A,C) and

Arthropods from the Lower Cambrian, showing that the first compound eyes were on eyestalks.



Fig. 4 (Left). Lower Cambrian arthropods. (A-B) Dorsal view of the anterior-displacement of the 2-segmental head and the anterior trunk portion

gu

Chen (2009) Int. J. Dev. Biol. 53, 733-751.



Fig.9. Early Cambrian lophophore-bearing animals. (A-E) The pelagic medusiform lophophore-bearing animals, Rotadiscus (A-C) and Eldonia (D-E). (F-G) Solitary sessile lophophore-bearing animals, Phlogetes (=Chenungkongella). Scale bars all are 1 cm.

Chen (2009) *Int. J. Dev. Biol.* 53, 733-751.

An early chordate from the Cambrian



Fig. 8. Anatomy of the Early Cambrian crest animal, Haikouella.

Chen (2009) Int. J. Dev. Biol. 53, 733-751.



Where did the sudden increase in morphological complexity and diversity come from?

- Increasing oxygen levels certainly did occur and made larger bodies and hard body parts possible.
- An early increase in complexity and diversity prior to hard body parts may have initiated an "arms race" of predators and prey, and allowed later "variations on a theme".
- The "trace fossils" of the late Ediacaran provide strong evidence for this hypothesis.
- Gene duplication and divergence clearly did play a major role in the Cambrian, particularly in defining new phyla. Gene duplication and divergence can occur very quickly when favored by selection (insecticide resistance, etc).

Trace fossils from late Ediacaran & early Cambrian are refuges of surface feeders, not sediment feeders



Dzik (2007) Geol. Soc. Lond. Special Publications 286, 405-414.

Trace fossils from late Ediacaran & early Cambrian are refuges of surface feeders, not sediment feeders



Dzik (2007) Geol. Soc. Lond. Special Publications 286, 405-414.

Where did the oxygen really come from?

Photosynthesis produces carbohydrates Light energy $CO_2 + 2H_2O$ --> $(CH_2O) + O_2 + H_2O$

Oxidative respiration produces ATP, which fuels metabolic processes ATP <-- $CO_2 + 2H_2O$ <-- $(CH_2O) + O_2 + H_2O$ % Carbon-13: upward peaks indicate more organic carbon burial (more erosion & oxygen production), valleys indicate less (glaciation etc)



Lane (2002) Oxygen: the molecule that made the world. Oxford Univ. Press.

Gene duplications and deletions occur at high frequencies



Hox genes are master regulator genes that evolved in tandem gene clusters



Foronda et al. (2009) Int. J. Dev. Biol 53, 1409-1419.

Duplication of entire Hox gene clusters in mammals



Garcia-Fernandez (2005) Nat. Rev. Genet. 6, 881-892.

Growth of the *Hox* gene cluster defines critical events in the Cambrian



Garcia-Fernandez (2005) *Nat. Rev. Genet.* 6, 881-892.

Hox gene clusters originally included *ParaHox* and *NK* gene clusters, also!



Garcia-Fernandez (2005) Nat. Rev. Genet. 6, 881-892.

Gene families arose during evolutionary history through a process of duplication of the ancestral gene followed by functional and structural specialization (divergence) of both copies¹⁻⁵. Gene families are widely accepted as a basis for classifying proteins, but their broader biological significance is less clear. Indeed, the existence of evolutionary homology between genes having dramatically different functional specificities is often dismissed as a curious but somewhat mysterious fact of life. Our knowledge of molecular genetics has now broadened so that it may be appropriate to re-examine this question and attempt to reach some general conclusions about the evolution of multigene families.

Gene duplication

Tandem genetic duplications in bacteria and bacteriophages occur spontaneously at a frequency of 10⁻³ to 10⁻⁵ per locus per generation, and can be of unlimited size⁶. Studies of the genetic basis of insecticide resistance suggest that spontaneous gene duplications occur in insects at a similar rate^{7,8}. Tandem duplications

The coevolution of gene family trees

KARL J. FRYXELL

Gene duplication mutants arise spontaneously at a bigb rate in bacteria, bacteriopbages, insects and mammalian cells, and are generally viable. Thus, the rate-limiting step in the evolutionary process of gene duplication and divergence was probably not gene duplication per se. Rather, it is likely that only a small fraction of all duplicated genes were retained, and were able to diverge into new specificities. Furthermore, gene duplications and functionally related gene families often show similarities in divergence dates, functional specificities, and phylogenetic tree topologies. These correlations suggest that the family trees of functionally related gene families coevolved because functionally complementary gene duplication and divergence events tended to be retained by natural selection. OK, but why chose the theory of evolution over rival ideas, such as "intelligent design" or special creation?

- A logical way to chose between theories is based on facts for which the theories make different predictions.
- Many such facts do support the theory of evolution, but can be obscured by nit-picking objections, and the volume of the evidence, and the vagueness of opposing ideas.
- It is easier and more satisfying for an author to bypass this unpleasantness by "preaching to the choir" with arguments that are compelling if you are already a "true believer" but do not address the concerns of the other side.
- Ironically, Darwin answered all of this in the beginning. He learned biology from Paley, an advocate of intelligent design, and tested this theory carefully before rejecting it.

What led Darwin to his conclusions?

- Darwin knew that the earth was more than 6,000 years old, and that different species had lived at different times. God was assumed to have created these animals at different times (Steno).
- Darwin was aware of the work of Malthus and the fact that animal population growth is limited by predation, starvation, and/or disease.
- Darwin was aware of the long history of artificial selection by English plant and animal breeders, which he studied in detail.
- But the key point was the natural geographical ranges of species, which was the focus of biology from Linnaeus to Paley. Darwin concluded that it was consistent with the origin of species by descent from common ancestors, but not with intelligent design as advocated by Paley.

Recent (100+ year old) lava flow, James Island



Landscape of Tower Island



Galapagos short-eared owl, Tower Island



Galapagos hawk, inside Volcan Alcedo



Galapagos penguin, James Island





Galapagos mockingbird, Tower Island



Galapagos ground finch, Tower Island



Galapagos tortoise, inside Volcan Alcedo

Pintail ducks, Tower Island

Flamingos, James Island

The geographical distribution of species on volcanic islands was predicted by "descent with modification"

- Species that can travel long distances over water (shore birds, wading birds) are common but did not form any new or unique species in the Galapagos.
- Species that travel less easily over water (large land birds) each formed one new & unique species in the Galapagos, but the same species is present on all islands of the Galapagos archipelago.
- Species with limited ability to travel over water (small land birds, reptiles, plants) formed different species on different islands.
- Species whose adults and eggs can not survive in seawater (amphibians, large mammals) did not reach the Galapagos, and were absent from the Galapagos.
- Species on the Galapagos are closely related to those on the coast of S. America (a very different climate), but completely different from those on the Cape Verde Islands (the same climate).
- Species that lived on particular islands or continents) were preceded in time by *closely related but distinct* fossil species that had the *same distribution*.

The geographical distribution of species on volcanic islands directly contradicts intelligent design (Paley)

- According to intelligent design (Paley), the species in the Galapagos should be similar or identical to those in Cape Verde Islands, but the opposite was observed.
- Intelligent design (Paley) predicts that non-native species would be less well-adapted than native species, but the opposite was observed.
- Intelligent design is unable to explain why there is an extremely close relationship between hundreds of species on the Galapagos & Peruvian coast, or between species on the Cape Verde islands & west Africa.
- Intelligent design is unable to explain why there are no native species of amphibians or large mammals on the Galapagos, or for that matter on any oceanic volcanic island on earth.
- Intelligent design is unable to explain why completely unsuitable species, such as the Galapagos penguin, were created uniquely in the Galapagos.
 "Descent with modification" easily explains this as the result of prevailing ocean currents.

Transposon families have star-shaped phylogenetic trees, which enables us to date each tree independently, by comparison to itself.

Transposon families shared between human & mouse are all older than the rodent-primate divergence, and are the same age in both genomes!

Table 6 Divergence levels of interspersed repeats predating the human-mouse speciation											
Interspersed repeat		Mouse				Human					
Family	Class	HÞ	Divergence	Range	JC	kb	Divergence	Range	JC	Substitution ratio	Adjusted ratio
L1MA6	LINE1	1,795	0.28	0.04	0.35	2,738	0.16	0.05	0.184	1.92	1.98
L1MA7	LINE1	789	0.28	0.04	0.35	3,502	0.16	0.04	0.181	1.92	1.96
L1MA8	LINE1	951	0.27	0.04	0.34	4,488	0.15	0.04	0.172	1.96	1.96
L1MA9	LINE1	1,032	0.28	0.04	0.35	6,468	0.18	0.05	0.201	1.74	1.96
L1MA10	LINE1	160	0.29	0.04	0.36	1,492	0.19	0.05	0.224	1.61	1.80
L1MB1	LINE1	627	0.29	0.04	0.36	2,947	0.18	0.05	0.211	1.71	1.87
L1MB2	LINET	725	0.28	0.04	0.35	3,309	0.18	0.06	0.201	1.75	1.87
L1MC1	LINE1	1,389	0.28	0.04	0.36	7,221	0.17	0.05	0.198	1.80	1.92
MLT1A	MaLR	984	0.31	0.04	0.39	2,203	0.21	0.04	0.242	1.62	1.73
MLT1A0	MaLR	1,794	0.30	0.04	0.38	5,424	0.19	0.04	0.219	1.74	1.80
MLT1A1	MaLR	539	0.29	0.04	0.37	1,705	0.19	0.04	0.214	1.74	1.78
MLT1B	MaLR	73	0.28	0.03	0.35	4,482	0.18	0.04	0.203	1.73	1.73
MLT1C	MaLR	2.071	0.30	0.04	0.37	5.511	0.21	0.04	0.245	1.53	1.64
Looper	DNA	33	0.28	0.04	0.34	48	0.18	0.03	0.211	1.62	1.69
MER20	DNA	435	0.29	0.05	0.37	2.205	0.19	0.05	0.222	1.65	1.76
MER33	DNA	232	0.27	0.05	0.33	1,207	0.18	0.04	0.211	1.57	1.63
MER53	DNA	82	0.26	0.05	0.31	524	0.17	0.05	0.191	1.63	1.63
Tigger6a	DNA	97	0.29	0.03	0.37	190	0.18	0.06	0.211	1.77	1.85

Shown are the number of kilobases matched by each subfamily (kb), the median divergence (mismatch) level of all copies from the consensus sequence, the interquartie range of these mismatch levels (range), and a Jukes-Gantor estimate of the substitution level to which the median divergence level corresponds (JC). Notice that RepeatMasker found, on average, four- to fivefold more copies in the human than in the mouse genome, as a result of the higher DNA loss in the rodent lineage as well a failure to identify many highly diverged copies. The two right columns contain the ratio of the JC substitution level in mouse over human, and an adjusted ratio (AR) of the mouse and human substitution level after subtraction from both of the approximate fraction accumulated in the common human-mouse ancestor. For this fraction we have taken the difference between the ancestral repeat average substitution level and least diverged ancestral repeat family (L1MA8). See the Supplementary information for a discussion of the origin of the variance in the human and mouse ratios.

Waterston et al. (2002) Nature 420, 520-562.