Phylogenetic assessment of the evolution of the fictional races of Tolkien across multiple fictional universes

Jermaine Mahguib
EEOB 563, ISU Spring 2019

INTRODUCTION

The high fantasy races of J. R. R. Tolkien, as he described them, became the basis for races that inhabit numerous universes in modern fiction. Here the term 'race' is used freely to refer to fictional, reproductively independent species of intelligent human-like peoples inhabiting fantastical worlds; the term is used ubiquitously in fantasy and science-fantasy based literature, and commonly employed in open discussions on these intellectual properties (IPs).

Permutations of Tolkien's elves, dwarves, humans and orcs are found across a broad range of written literature and published video game series, each adding new physical features, behavioral ranges and story roles, while retaining others from the original descriptions. For example, in Tolkien's fictional works, the elves fill the archetypal role of the 'old man,' the ancient first race to inhabit the world. In the contemporary timelines of Tolkien, the elves are a race whose time of prominence has come and gone, and they serve as a fount of wisdom on matters of history and ancient arcana, and are often thrust into the role of providing some form of guidance for younger races. This archetypal role is often maintained in derivatives of the race across fictional universes.

Other examples of retained traits include the dwarves' affinity for occupying subterranean domains, the humans' tendency to expand their empires, and the orcs' green skin and 'monstrous' visages. Conversely, a number of traits have seen some drift from Tolkien, such as the role that the orc race plays in stories. While Tolkien's orcs were portrayed as vile and heinous, and this persisted across other IPs, some universes established their orc peoples as being more on par with their other 'noble' races (e.g. humans, elves and dwarves).

In the present study, a suite of traits that describe the core races of several fictional universes was established and a character state matrix was created to provide a dataset for phylogenetic analysis in order to assess how these races have evolved from Tolkien's original incarnations. Character traits were categorized into three classifications: physiological traits, behavioral traits, and archetypal traits. For example, comparative physical *strength* of typical members of a given group was categorized as a physiological trait, while *aggression* level was categorized under behavioral traits.

Character state matrix data was used in a set of maximum likelihood (ML) analyses in an attempt to address two primary questions: (1) do various incarnations of Tolkien's fictional races from different IPs form statistically supported phylogenetic clades; (2) in what categorical respects (physiological, behavioral or archetypal) have the original high-fantasy races of Tolkien diverged the most across multiple incarnations?

The former was approached via a ML analysis using character state sequence data from all three trait categories concatenated. If the tested IPs exhibit minimal alteration to character traits describing their fictional races relative to the type races of Tolkien, then we can expect to see, for example, all elven derivatives forming a supported clade in the tree topology. If, however, the fictional peoples of fantasy IPs have substantially deviated in our established character traits across universes, then we can expect the topology of the tree to suggest few if any supported clades, and instead observe a tree characterized by polytomies.

To address the second question, phylogenetic trees were generated using categorical sequence alignments separately, and compared to one another; in other words, trees were estimated using aligned character state sequences from the physiological traits category, others from the behavioral traits and archetypal traits categories. If a category of character traits has experienced less deviation from the original Tolkien descriptions, then we would expect to see distinct clades in the tree topology, were all elves form a group, all dwarves, etc. If, on the other hand, some incarnations of Tolkien's fictional races have experienced substantial categorical deviation, then we might expect to see more polytomies or a single large polytomy in the tree topologies.

METHODS

Taxa and Multistate Sequence Data

A total of 70 traits (32 physiological, 20 behavioral and 18 archetypal) were considered for this study, each assigned a set of numerically coded states (see Tables 1 – 3). The number of different states per trait varied from 2 to as many as 7, though the majority of traits were coded as having between 3 to 5 states. The 70 character traits were scored for a total of 32 taxa hailing from one of seven fictional universes: 4 taxa from J. R. R. Tolkien's *The Lord of the Rings*; 5 from Ed Greenwood's Dungeons & Dragons-inspired *The Forgotten Realms* series of novels and games, produced by various authors and publishers; 5 from Bethesda Studios' *The Elder Scrolls* series; 5 from Blizzard Entertainment's *Warcraft* series; 5 taxa each from Games Workshop's *Warhammer* and *Warhammer 40K* series of novels and games, produced by varies authors and publishers; and 3 taxa from CD Projekt Red's *The Witcher* series, adapted from the novels by Andrzej Sapkowski.

Four additional taxa were selected and scored for use as outgroups in phylogenetic inference: 2 taxa from Blizzard Entertainment's *Starcraft* series; an additional taxon from Games Workshop's *Warhammer 40K* series; and 1 taxon from film series, novels and graphic novels based on Dan O'Bannon and Ronald Shusett's *Alien*. All taxa, the respective fictional universes to which they belong, and the numerical sequence of trait scores extracted from the character state matrix are shown in Table 4. All character states in a given category, for a given race, were determined relative to other members of the same universe; cross-universe comparisons were not considered when determining character states for taxa.

state matrix. The scores of the matrix were used as sequence alignments for phylogenetic inference.

3 = average human	4 = slightly taller than average human	5 = 1 ft or more taller than average human	6 = varies among members of group	
3 = stocky and muscular	4 = bulky and muscular	5 = varies among members of group		
3 = green or grey skin tones	4 = purple or blue or grey skin tones	5 = golden skin tones	6 = black	7 = variable
3 = varies among members of group				
3 = varies among members of group				
3 = varies among members of group				
3 = varies among members of group				
3 = rolling	4 = varies among members of group			
3 = no eyes				
3 = no teeth				
3 = no ears				
3 = elongated	4 = very elongated	5 = no ears		
3 = characteristic				
3 = sexually monomorphic				
3 = highly attractive	4 = monstrous			
3 = high strength	4 = very high strength			
3 = high agility	4 = very high agility			
3 = high intelligence	4 = very high intelligence			
3 = high mental fortitude	4 = very high mental fortitude			
3 = high hardiness	4 = very high hardiness			
3 = high perception	4 = very high perception			
3 = very long-lived (thousands of years)	4 = functionally immortal	5 = varies among members of group		
3 = moderate magical aptitude	4 = moderate to high magical aptitude	5 = low to high magical aptitude		
3 = moderate psychic aptitude	4 = moderate to high psychic aptitude	5 = low to high psychic aptitude		

Table 1. List of physiological character traits and numerically-coded associated states used to score the character-

Physiological Traits	**Scores**	
Height	1 = 1 ft or more shorter than average human	2 = slightly shorter than average human
Body Type	1 = lean and lithe	2 = lean and muscular
Skin Tone	1 = human fair skin tones	2 = full range of human skin tones
Skeletal Structure	1 = endoskeleton	2 = exoskeleton
Cranium	1 = globose human-like	2 = elongated
Type of Forelimbs	1 = terminate in digits	2 = terminate in claws/scythes
Number of Forelimb Digits	1 = 5 digits per limb	2 = 4 digits per limb
Opposable Thumbs Present	1 = yes	2 = no
Type of Legs	1 = plantigrade	2 = digitigrade
Type of Locamotion	1 = legged	2 = limbless
Type of Eyes	1 = physical, human-like eyes	2 = glowing, energy-infused eyes
Mouth Present	1 = yes	2 = no
Type of Teeth	1 = flat teeth	2 = sharpened teeth
Tusks Present	1 = yes	2 = no
Ears Present	1 = yes	2 = no
Ear Shape	1 = lobed like humans	2 = pointy
Ear Length	1 = human-lengthed	2 = slightly elongated
Hair Growth on Head	1 = yes	2 = no
Facial/Body Hair	1 = none	2 = minimal
Nose Present	1 = yes	2 = no
Reproduction	1 = sexual	2 = asexual
Sexual Appearance	1 = androgynous	2 = sexually dimorphic
Attractiveness	1 = unattractive	2 = moderate to highly attractive
Strength	1 = low to moderate strength	2 = moderate to high strength
Agility	1 = low to moderate agility	2 = moderate to high agility
Intelligence	1 = low to moderate intelligence	2 = moderate to high intelligence
Mental Fortitude	1 = low to moderate mental fortitude	2 = moderate to high mental fortitude
Hardiness	1 = low to moderate hardiness	2 = moderate to high hardiness
Perception	1 = low to moderate perception	2 = moderate to high perception
Longevity	1 = short-lived (human lifespan or less)	2 = long-lived (hundreds of years)
Magical Aptitude	1 = none	2 = low to moderate magical aptitude
Psychic Aptitude	1 = none	2 = low to moderate psychic aptitude

Trait Categorical Descriptions

The 70 character traits used in this study were divided up into three categories, *physiological* traits, *behavioral* traits and *archetypal* traits, each of which warrants some additional elaboration here. The physiological category consists of traits that characterize physical descriptions of typical members of a given fictional race/taxon. Examples include physical characteristics such as the typical height of taxon members, the type of physical locomotion they utilize to move around in 3D space, the presence and type of, or absence, of physical features such as ears and hair, traits associated with physical or mental performance such as strength and intelligence of typical taxon members, as well as traits related to fictionally established physiology such as longevity and aptitude for supernatural expression.

The behavioral category consists of traits that characterize some generalized behaviors associated each fictional race as a whole. Some examples include behavioral adaptability of typical taxon members, referring to the ability of individuals to adapt or adjust their general thinking and/or behavior to different circumstances, a race's typical attitude toward matters of foreign policy, the general

associated states used to score character-state matrix. Scores of matrix were used as sequence alignments for phylogenetic inference.

2 = somewhat adaptable	3 = adaptable	4 = highly adaptable	
2 = adaptable	3 = highly adaptable	4 = extremely adaptable	
2 = low tolerance	3 = moderate or grudging tolerance	4 = tolerant	
2 = low to moderate aggression	3 = moderate to high aggression	4 = highly aggressive	
2 = provoked into aggression on threat	3 = easily provoked into aggression		
2 = moderate to high	3 = high to very high	4 = exquisite	5 = feral
2 = moderate engagement	3 = high engagement		
2 = moderate exploitation	3 = high exploitation	4 = severe exploitation	
2 = moderate grade	3 = high grade	4 = highest grade	
2 = trade and commerce	3 = trade, commerce and conquest	4 = war and conquest	5 = invasive
2 = low to moderate levels	3 = moderate to high levels	4 = prescient	
2 = low to moderate impulsivity	3 = moderate to high impulsivity		
2 = lean toward social dependence	3 = socially interconnected	4 = hive mind	
2 = display some spirituality	3 = display broad spirituality	4 = display all-consuming spirituality	
2 = low	3 = low to moderate	4 = moderate to high	5 = very high
2 = form strong familial bonds	3 = form no familial bonds		
2 = moderately so	3 = overly so	4 = excessively so	5 = not expressed
2 = moderately so	3 = overly so	4 = excessively so	5 = not expressed
2 = moderate	3 = overt	4 = excessive	5 = not expressed
2 = openly expressed	3 = overtly expressed	4 = none expressed	

Table 2. List of behavioral character traits and numerically-coded

Behavioral Traits	**Scores**
Behavioral Adaptability	1 = rigid
Environmental Adaptability	1 = somewhat adaptable
Tolerance of Other Races	1 = near zero tolerance
Aggression	1 = low aggression
Incitable Aggression	1 = not easily provoked into aggression
Martial Prowess	1 = basic to moderate
Explorative	1 = low ingagement
Environmentally Exploitative	1 = low exploitation
Craftsmenship	1 = low grade
Foreign Policy	1 = isolationists
Prescience	1 = none
Impulsivity	1 = low impulsivity
Sociality	1 = lean toward individual independence
Spirituality	1 = none
Scholarship	1 = none
Familial Bonds	1 = form loose familial bonds
Hedonistic	1 = minimally
Sadistic	1 = minimally
Curiosity	1 = minimal
Joviality	1 = subtly expressed

level of expression of individualistic behaviors such as joviality, impulsivity and sadism, general level of engagement in sociality and the formation of familial bonds, and engagement in practical behaviors such as martial training, crafts and predicting future events.

The archetypal category encompasses character traits and states that describe how a race fits into a particular universe in terms of the stories that are told within them. For example, in Tolkien's stories the orcs fill the role of an irredeemably evil race of vile killers who serve as a persistent villainous army that threatens the noble races of the world. In another universe, the Warcraft universe, the orcs are initially presented as a race occupying a similarly villainous story role, but are later revealed to actually be a noble race themselves who are only pitted against the protagonist races circumstantially. These adjustments to the original archetypes are part of the evolution of these races across fiction, and thus an attempt has been made here to include these considerations in the proposed analysis. Some examples of archetypal traits featured in the present study include racial age, or how long the group has inhabited the fictional world, contemporary demographic prevalence, role in storytelling, and racial origin/creation.

The states and scoring for these traits, across all three categories, was accomplished with the understanding that individuals of any given fictional group can and do vary dramatically from what their collective race may have been scored as here for a given trait; scoring for these traits was meant to reflect broad generalities that are often used to describe these races in either their source material or in supplemental secondary material found on the internet.

matrix. The scores of the matrix were used as sequence alignments for phylogenetic inference.

3 = antagonistic	4 = evil	5 = "force of nature"	
3 = young; "upstart"			
3 = uncommon	4 = fading	5 = rare or extinct	
3 = highly influential	4 = influential to security		
3 = high connection to nature	4 = deep connection to nature		
3 = moderate to high tech	4 = highly advanced teched	5 = hyper advanced tech	
3 = highly evolved or created	4 = hyper evolved or created		
3 = susceptible	4 = highly susceptible		
3 = primary focus (villainous)	4 = secondary focus (villainous)	5 = background lore	
3 = often entangled	4 = often associated with		
3 = martial (technological & supernatural)	4 = martial (biological)		
3 = created by another race via tech or magic	4 = unknown		
3 = seek to consume their environment			
3 = excessively motivated	4 = all-consuming		
3 = seek to defeat and subjugate enemies	4 = seek to defeat and assimilate enemies		
3 = advanced	4 = highly advanced	5 = hyper advanced	6 = biological
3 = grand	4 = sweepingly grand		
3 = underground or within mountains	4 = unassociated with particular habitat		

Table 3. List of archetypal character traits and numerically-coded associated states used to score the character-state

Archetypal Traits	**Scores**	
Conflict Role	1 = noble	2 = pragmatic
Racial Age	1 = ancient; "first born"	2 = middle aged; "long established"
Prevalence	1 = common	2 = moderately common
Culturally Influential	1 = minimally	2 = in some respects
Connection to "Nature"	1 = none	2 = some connection to nature
Technological Advancement	1 = low tech	2 = low to moderate tech
Biological Advancement	1 = primitively evolved or created	2 = moderately evolved or created
Corruptability	1 = insusceptible	2 = somehwat susceptible
Story Role	1 = primary focus (protagonists)	2 = secondary focus (protagonists)
Prophecy Entanglement	1 = rarely entangled	2 = sometimes entangled
Threat Type	1 = martial (technological)	2 = martial (supernatural)
Creation	1 = divinely created or manipulated	2 = product of evolution
Broad Motivation	1 = seek harmony with their environment	2 = seek to control their environment
Economic Motivation	1 = minimally motivated	2 = moderate to highly motivated
Domination Motivation	1 = seek to defaet and hault enemies	2 = seek to defeat and exterminate enemies
Architectural Design	1 = basic	2 = moderate
Architectural Scope	1 = minimalist	2 = expansive
Associated Homeland Habitat	1 = in the open world	2 = in forests and woods

Maximum Likelihood Analyses

A set of ML analyses were conducted using the program *RaxML* in order to approach addressing two questions; how much have iterations of Tolkien's fantasy races changed across different IPs, and have they changed, if at all, more or less so with respect to one of the trait categories over the others? To get at the first question, character state sequences extracted from the trait scoring matrix from all three trait categories were concatenated for each taxa. Then, depending on outgroup(s) chosen, 33 or 36 concatenated sequences were aligned for use in three

Table 4. Taxa and character state sequence data used in this study. Taxa are grouped according to the fictional universes they are from. Reading from left to right along the sequences, positions 1 – 32 correspond to physiological traits from Table 1, in order from *Height* to *Psychic Aptitude*; positions 33 – 52 correspond to behavioral traits from Table 2, in order from *Behavioral Adaptability* to *Joviality*; positions 53 – 70 correspond to archetypal traits from Table 3, in order from *Conflict Role* to *Associated Homeland Habitat*

Intellectual Property (IP)	Taxa	Character State Sequences
Lord of the Rings (LotR)	Humans_LotR	32211111111111211113112222212212132432222323132242213213131224131122122 1
	Elves_LotR	41111111111112122211111313321344111221311412113511121114143322431111312
	Dwarves_LotR	13211111111112111131111322332231223312323212224221321232233121111313 33
	Orcs_LotR	12311111111122122121114321132221131432141413211113143224111441112121 21
The Forgotten Realms (FR)	Humans_FR	32211111111111211113112222222221513233222323222342223213132323123222122 1
	Elves_FR	21111111111112122211111313221325122321222322323412133113232331231221211
	Drow_FR	21411111111111212221111221332132512123332233231341332131322333423123332 3
	Dwarves_FR	13211111111111211113112132233322122332222232222232213211322332123123133 3
	Orcs_FR	34311111111112112213111432113112112233212141321111311322411144212213111
The Elder Scrolls (ES)	Humans_ES	33211111111111211113112222222221513233222223222342222223131223123222122 1
	Altmer_ES	42511111111111212212111323322335122322211222223412132213223321232221311
	Dunmer_ES	32411111111111212221211221322133512123321123231341222131322333123222121 1
	Dwemer_ES	32411111111111212221311221242122512222212214121225211411151243251322114 33
	Orsimer_ES	43311111111112112213112132233225132333312332313332222222221223123223111 11
Warcraft (WC)	Humans_WC	33211111111111211113112222222221513243222323232342223213131323123222132 1
	Night_Elves_WC	52411111112112124131123233224341213213112131124111211132423212311112 12
	High_Elves_WC	31111111112112123121123133213341213222113232224122411222223323213221141 2
	Dwarves_WC	13211111111111211113112132233225132432232422222422143122323322131131433
	Orcs_WC	44311111111112112213112132223215132333323233323322222131322331232221221
Warhammer (WH)	Humans_WH	32211111111111211113112222212212132232222323232342223223131224133322122 1
	High_Elves_WH	41111111111111212221111131332133412123231131322351112111312332243321132 1
	Dark_Elves_WH	41411111111111212221111121332133412123333233433123344334124133442332333 21
	Dwarves_WH	13211111111111211113112132233222132232323224122224221311122233221331433
	Orcs_WH	44311111111112112221123442113132213143513141332135314521411314133312111
Warhammer 40K (40K)	Humans_40K	32211111111111211113112211321211522133124342324312231331314241333232444
	Eldar_40K	41111111111111212221111132343243141113241141422251112131411532243321254 4
	Dark_Eldar_40K	41411111111111212221111122343243141114342244431243444341341534423323354 4
	Squats_40K	13211111111111211113112122322221522132223322223422131235114225133232433
	Orks_40K	44311111111112112221123442113132213143513241332135314521413314133312114
The Witcher (W)	Humans_W	32211111111111211113112222222221513233222323122242322223131323133222322 1
	Elves_W	21111111111111212221111131322132513232122132122231312221322223243222111 2
	Dwarves_W	13211111111111211113112121222222213243222232122242212221322223213222121 1
Outgroups (OG)	Protoss_OG	51412121212232235212122334223314222323124132325111311114153213332215 44
	Zerg_OG	65722332342122235212234441343412141435241513411355545214114333433446 24
	Tyranids_OG	65722333342122235212234441343512141435241513411355545214114144443446 24
	Xenomorphs_OG	51622131213122235212234441234511131435141513311355545214113131433446 24

sets of analyses conducted in *RaxML*. For each set of analyses, a search of the tree space of possible topologies for the data and identification of the best tree was performed, and bootstrap support values for tree nodes were then calculated.

A total of 1000 likelihood trees were searched (--tree pars{500},rand{500}) using a multistate model with equal rates of change for all possible state changes (--model MULTI#_MK). The 'MULTI#' model used depended on how many different states appeared in each alignment that was assessed, which depended on which taxa were designated as outgroups and whether or not some of the intended outgroup taxa were excluded from an alignment (see Table 5 for a summary of all analyses conducted and options used). Multistate models must be set to a maximum number of states that is greater than the total number of unique states that appear in the dataset by at least one. The '_MK' designates the model as having equal rates of change for all possible state changes. This was used because the assumption was made that in reality creators are essentially free to make whatever changes they want to Tolkien's archetypes for use in their own IPs; nothing technically restricts them from doing so.

Tree node support values were calculated by a non-parametric bootstrap analysis performed with 1000 replicates (--bs-trees 1000), using the same outgroup assignments and multistate models for each corresponding tree search analysis. Convergence of bootstraps was checked post-analyses using the --bsconverge command in *RaxML*. The default cut-off value for MRE-based bootstopping tests performed by the program during convergence testing was used.

To attempt to gain an idea of how the taxa used in this study have changed comparatively between the three trait categories (physiological, behavioral and archetypal), matrix score sequences for traits from each category were aligned separately and analyzed in a similar manner as described for the concatenated alignments above, with appropriate adjustments to the multistate model utilized depending on the alignment. Additionally, the outgroups that were included were adjusted as well. *RaxML* was unable to root trees during tree searches for any of the single category alignments when any combination of the four outgroup taxa were designated as such. The program was only able to root trees searched when a single taxon was designated as the outgroup for each alignment. As a result, ML analyses were conducted on two sets of single-category alignments, each set with a different designated single taxon outgroup.

Of the four fictional races chosen for this study to serve as outgroups, the Zerg, Tyranids and Xenomorphs taxa are all very similar to one another across a majority of the character traits scored. Therefore, one of these taxa (Zerg) was chosen to be the sole outgroup for the first set of single-category alignments (Tyranids_OG, Xenomorphs_OG and Protoss_OG sequences were removed from these alignments entirely). For the second set of single-category alignments, the Protoss_OG taxon was designated as the outgroup and sequences from the other three outgroup taxa were removed from the alignments. This replication was done in part because it was difficult to decide which of the four taxa would work best as the single outgroup for the rest of the dataset, and in part simply to see how using two different outgroups would affect the resulting tree topologies. Together with

Table 5. Summary of options used for running analyses in *RaxML* per alignment/outgroup combination assessed. The model of evolution used, number of likelihood trees searched, and number of bootstrap replicates performed are indicated for each dataset. Figure # is also indicated for datasets that produced a tree shown as a figure.

Figure	Analysis Dataset	Outgroup(s) Used	Model Used	# of Trees Searched	# of Bootstrap Replicates
1	Concatenated Alignment	Zerg_OG Tyranids_OG Xenomorphs_OG	MULTI8_MK	1000	1000
2	Concatenated Alignment	Protoss_OG Zerg_OG Tyranids_OG Xenomorphs_OG	MULTI8_MK	1000	1000
3	Concatenated Alignment	Protoss_OG	MULTI6_MK	1000	1000
--	Physiological Alignment	Zerg_OG	MULTI8_MK	400	1000
--	Behavioral Alignment	Zerg_OG	MULTI6_MK	400	1000
4	Archetypal Alignment	Zerg_OG	MULTI7_MK	400	1000
--	Physiological Alignment	Protoss_OG	MULTI8_MK	400	1000
--	Behavioral Alignment	Protoss_OG	MULTI6_MK	400	1000
4	Archetypal Alignment	Protoss_OG	MULTI7_MK	400	1000

the concatenated alignment runs, a total of 9 ML analyses were conducted on datasets for this study.

RESULTS

Analyses of Concatenated Alignment

Maximum likelihood trees, with mapped-on bootstrap support values, generated from the concatenated character state alignment are shown in Figures 1 – 3. Figure 1 shows the ML tree resulting from the analysis of the concatenated alignment where 3 taxa (Zerg_OG, Tyranids_OG and Xenomorphs_OG) were designated as outgroups. The tree topology recovered all of the elven taxa as monophyletic. Of the 7 IPs that were included in this study, 4 of them utilize the relatively common trope in fantasy and science-fantasy of incorporating the presence of 'dark' elves in the fictional universe; dark elves are often portrayed as a

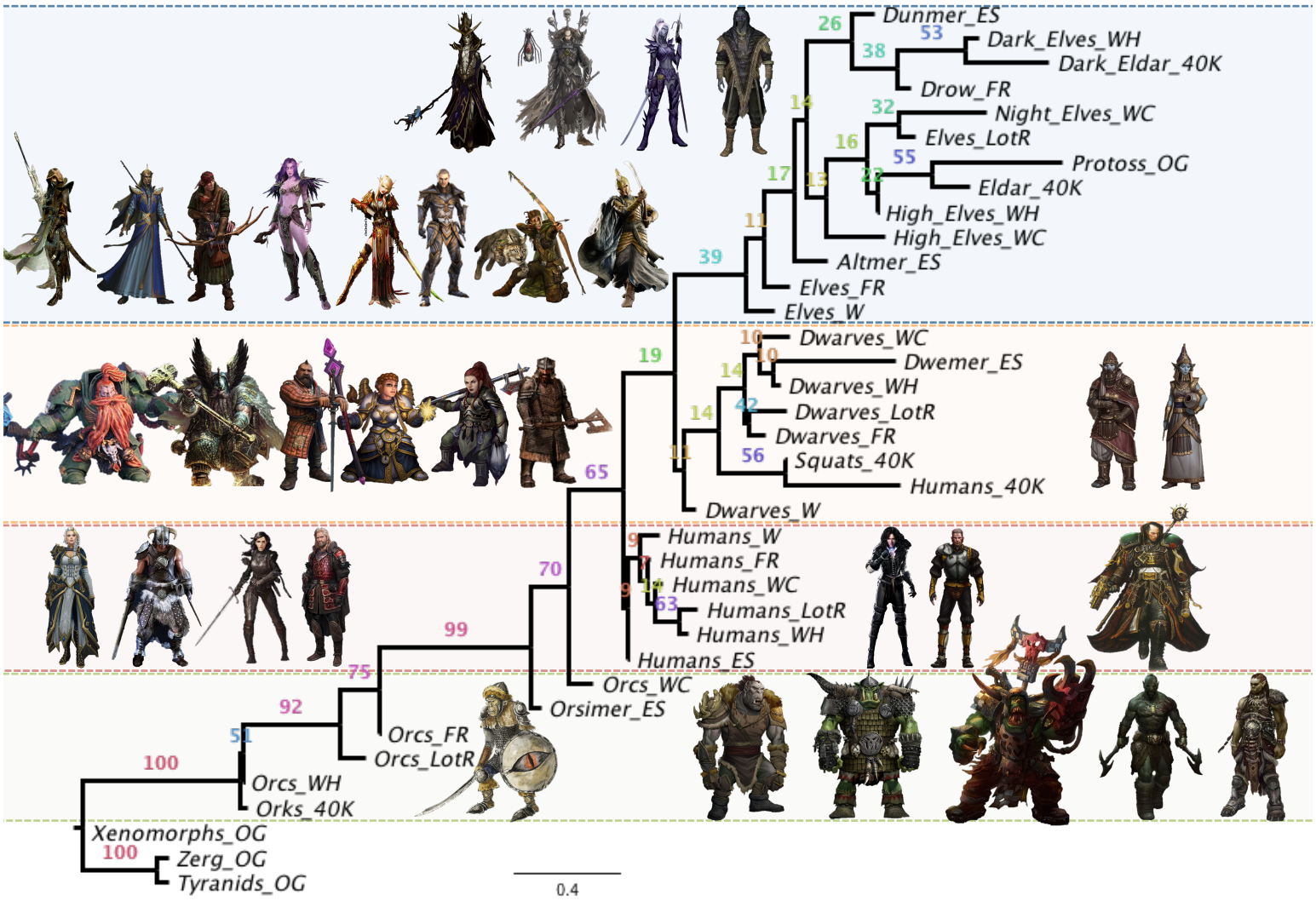

Figure 1. Maximum likelihood phylogeny from concatenated character trait alignment with the Zerg, Tyranid and Xenomorph taxa set as outgroups.

villainous group, the opposite of their virtuous elven counterparts. In the Figure 1 phylogeny, all the dark elven taxa formed a monophyletic clade. While all elves grouped together, the remaining, non-dark elven taxa did not form a monophyletic clade. Interestingly, the Protoss_OG taxon, though originally intended as an outgroup and although not a race of elves, claded within that group. Another Tolkien-derived fictional race that grouped together were the dwarves. Dwarves from across all IPs formed a paraphyletic clade, and this clade was sister to the elven clade in the tree topology. The elf-dwarf clade came out as sister to the polyphyletic humans. All human taxa except one (Humans_40K) nested within a clade sister to the elves and dwarves; the Humans_40K taxon from the *Warhammer 40K* universe ended up being nested within the dwarven clade, sister to the Squats (the dwarves of the same universe). The orc taxa, like the dwarves, were also paraphyletic, although spread out on the tree topology rather than forming an inclusive clade.

Figure 2 shows the ML tree that resulted when all four intended outgroup taxa (Protoss_OG, Zerg_OG, Tyranids_OG and Xenomorphs_OG) were designated as outgroups in the tree search and bootstrap analyses. With respect to the human, dwarven and elven taxa, the tree topology was the same as resulted from the previous analysis described above; dark elves and elves as a whole were monophyletic and sister to the paraphyletic dwarves, and these two clades together were sister to the polyphyletic humans. Where additionally designating the Protoss_OG taxon as an outgroup changed the tree topology (at least at a glance) as compared to the previous analysis was in the placement of the orcish taxa; rather than clading with the other three sets of Tolkien-derived taxa, the orc taxa claded with the other three outgroup taxa (Zerg_OG, Tyranids_OG and Xenomorphs_OG) and were collectively sister to the former. Even though the Protoss_OG taxon was assigned as an outgroup here, it still ended up being nested within the elven clade as it was in the previous analysis; since the outgroups did not form a monophyletic group, *RaxML* was unable to root the tree based on the outgroups. Given that the tree in Figure 2 is unrooted, the relationships suggested by the tree topology are actually the same as the tree shown in Figure 1.

What *did* change the relationships depicted by the tree topology resulting from the concatenated dataset was removing the Zerg_OG, Tyranids_OG and Xenomorphs_OG taxa from the alignment and designating the Protoss_OG taxon as the sole outgroup in the analysis. Figure 3 shows the resulting tree; the similarities in character states between the Protoss_OG taxon and the elven taxa caused the monophyly of the latter to be broken up, and dragged the previous elven clade out of its position as sister to the dwarves. In this topology, the elves are paraphyletic and basal to both the dwarves and humans. The change in outgroup also shifted the placement of the orcs; rather than being basal as in the previous trees, the orcs came out here as more derived, and sister to the humans.

Bootstrap support values across all three of the trees described above were overall low except in a few cases. The relationships suggested among the orc taxa were well supported in these analyses, with nodes being supported between 70 and 100 percent of the time across all bootstrap replicates. In the two instances they were utilized, the relationships among the Zerg_OG, Tyranids_OG and

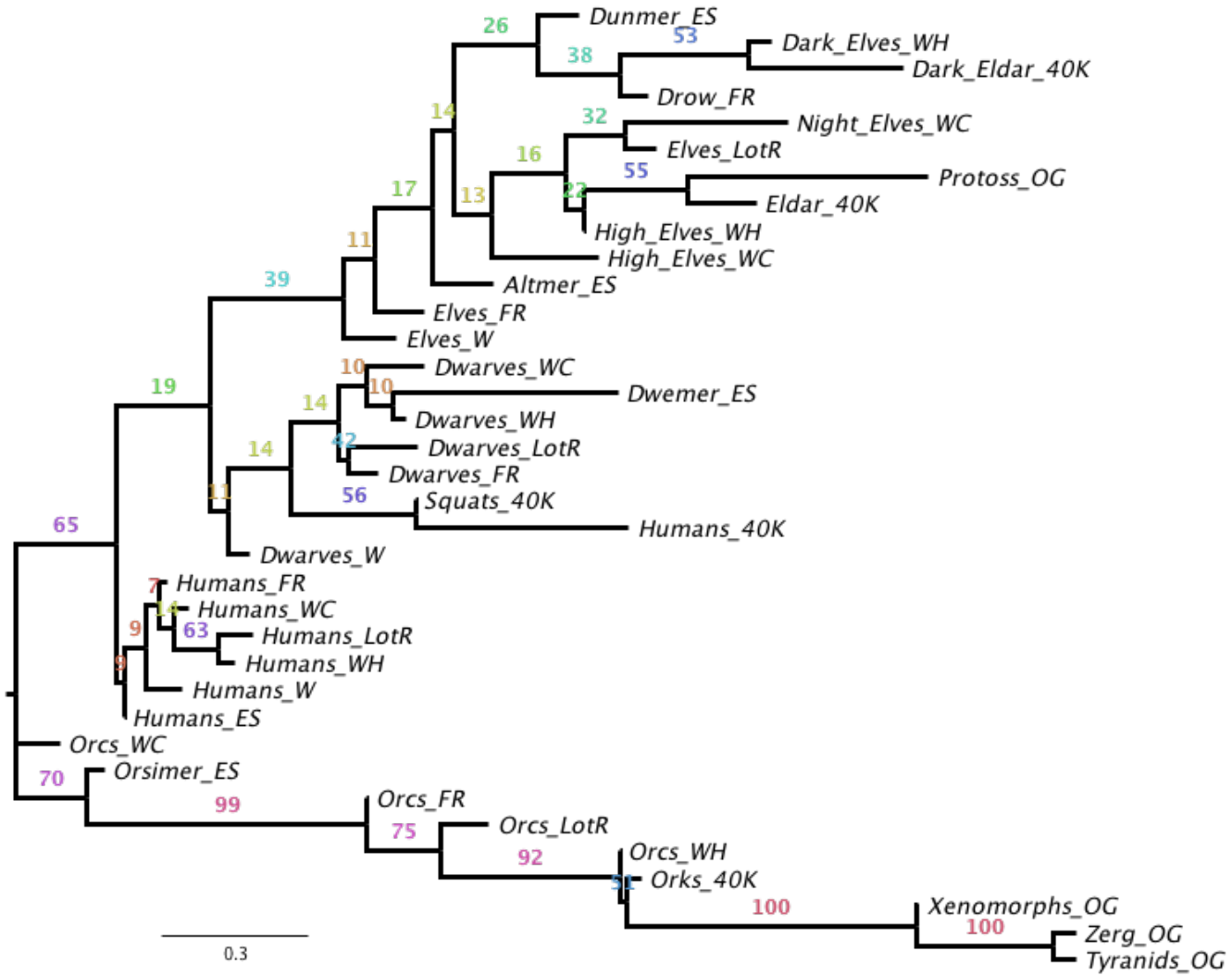

Figure 2. Maximum likelihood phylogeny from concatenated character trait alignment with the Protoss, Zerg, Tyranid and Xenomorph taxa set as outgroups.

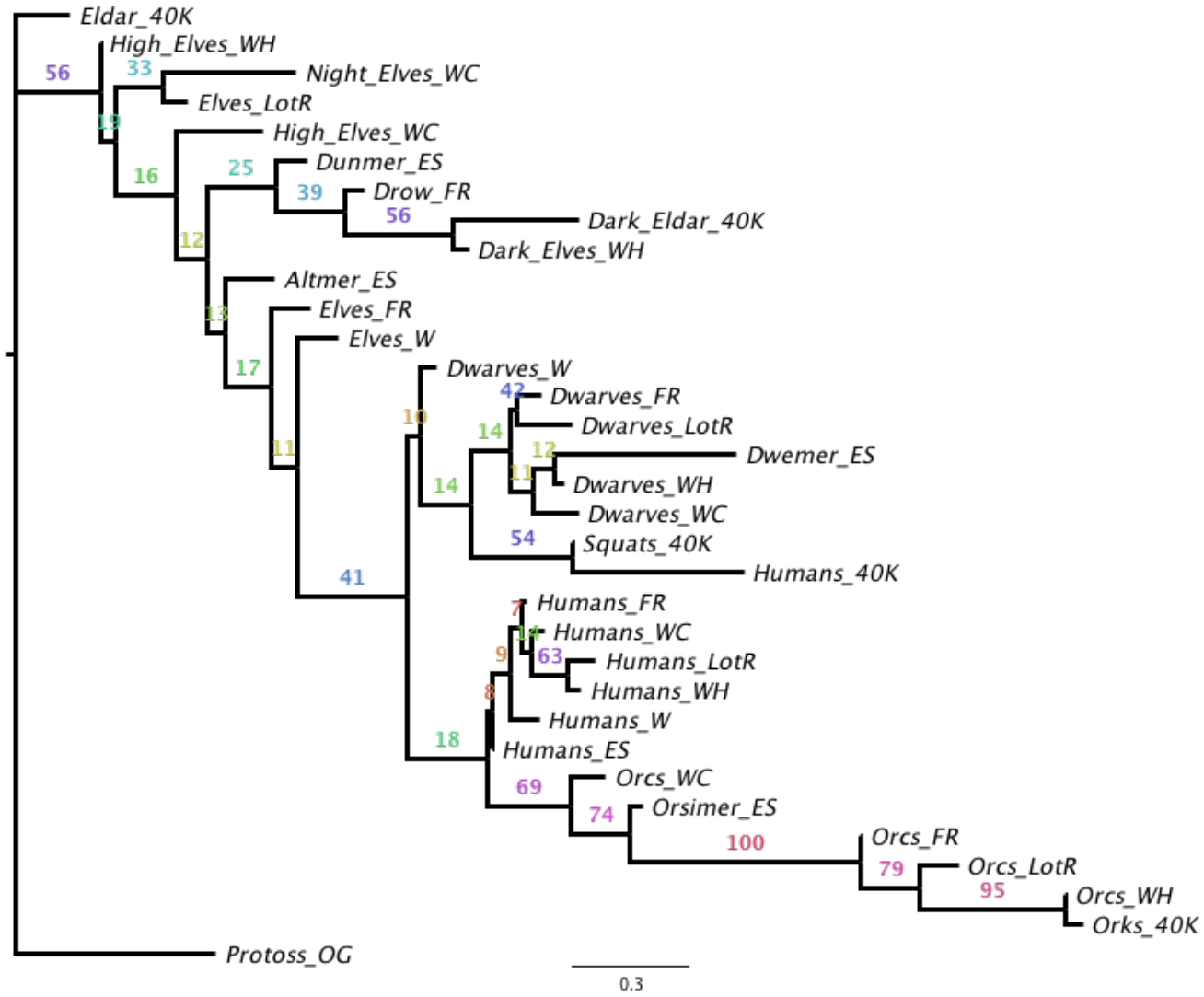

Figure 3. Maximum likelihood phylogeny from concatenated character trait alignment with the Protoss taxon set as the sole outgroup.

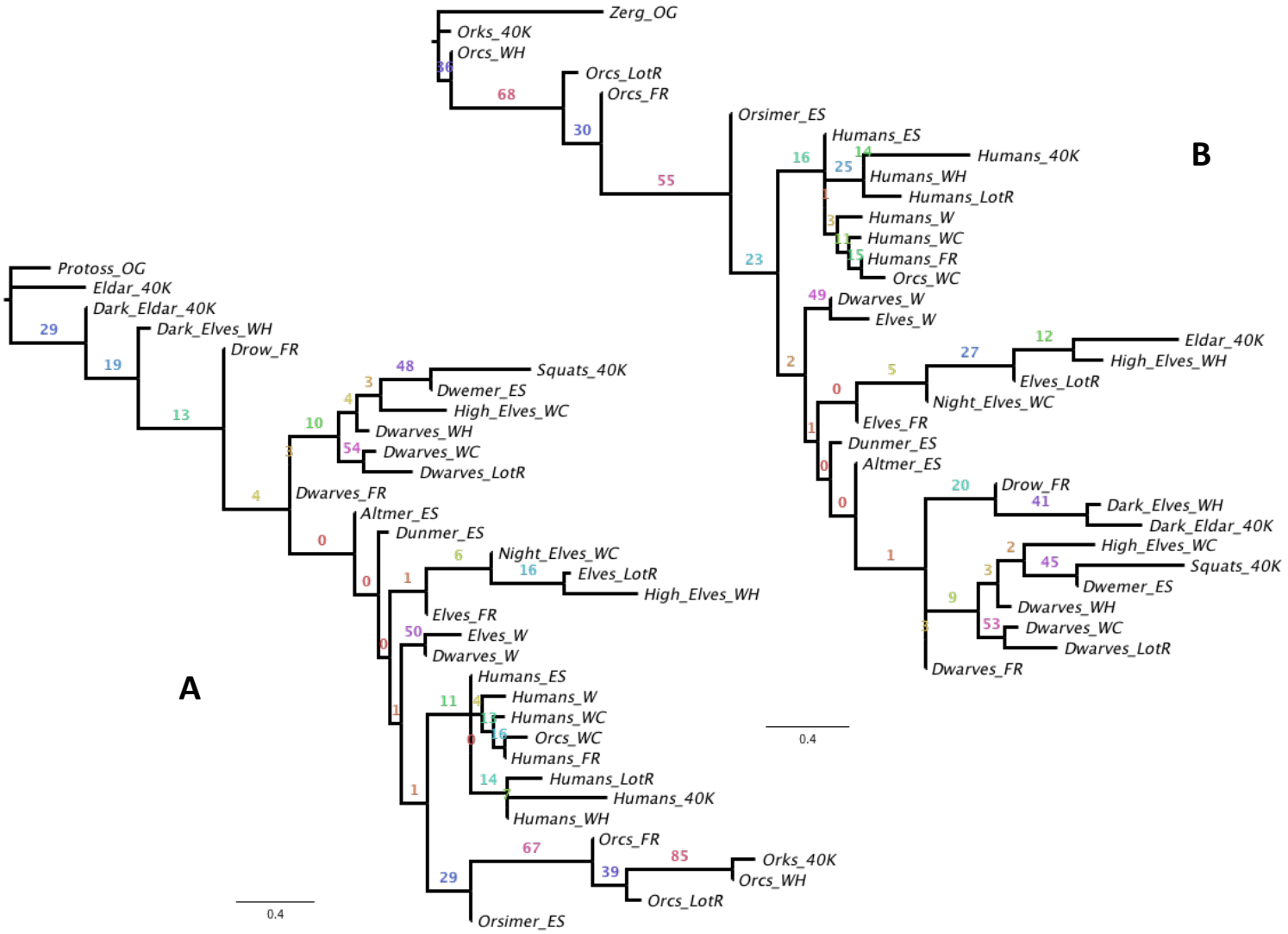

Figure 4. Maximum likelihood phylogenies from archetypal trait alignments; A. Protoss taxon set as outgroup; B. Zerg taxon set as outgroup.

Xenomorphs_OG outgroup taxa were also strongly supported. In the first two trees shown in Figures 1 and 2, the clade containing all the elves, dwarves and humans was moderately supported at 65 percent bootstrap support. All other nodes were poorly of very poorly supported.

Analyses of Categorical Alignments

The categorical character state alignments were problematic to assess. Of the six ML analyses attempted on these datasets (see Table 6), two of them (both behavioral alignment analyses) failed to complete the tree search due to persistent errors encountered in *RaxML* that were unable to be resolved. A third analysis (of the physiological alignment with the Zerg_OG taxon as outgroup) persistently failed to complete the bootstrap assessment due to an error terminating the process as well. No issues running the program were encountered throughout assessment of the concatenated sequence alignment, so the issue may have been related to the small sequence size of the categorical alignments. Nonetheless, analysis of some of the categorical alignments was completed and the results of two of them are described here.

Since both runs of the archetypal sequence alignment were completed successfully, their resulting trees are described and can be seen in Figure 4. In general, the tree topologies were affected similarly by their respective designated outgroups as in the analyses described for the concatenated alignment. In the archetypal alignment where the Zerg_OG taxa was set as the outgroup, the orc taxa were more basal and the elves more derived, with the opposite being the case when the Protoss_OG taxon was made to be the outgroup. Other than that, limiting the dataset to just the archetypal alignment did affect the tree topology somewhat compared to the concatenated dataset. In the tree from the Zerg_OG alignment, the elven and dwarven taxa formed an inter-mixed clade, which placed as sister to a clade that this time contained all the human taxa; although humans all grouped together archetypally, they were still paraphyletic as one of the orc taxa (orcs from the *Warcraft* universe) nested within that clade as well.

Similarly, the humans all grouped together in the topology for the tree built from the archetypal alignment with Protoss_OG set as outgroup. The elf-dwarf clade was broken up however by the outgroup assignment, the two sets of taxa spread out across the basal half of the tree.

Bootstrap support values for nodes of these trees were even lower than in the trees for the concatenated alignment; nearly all nodes had less than 70 percent support.

DISCUSSION

The tree topologies resulting from the analyses of the concatenated sequence alignment suggest that despite alterations creators have made to their versions of Tolkien's iconic fantasy races, most of these fictional peoples have not changed

Table 6. Results of maximum likelihood analyses. Analyses are distinguished by combination of alignment and outgroup(s) used. Summary statistics shown are: the final log likelihood of each analysis; their AIC, AICc and BIC scores; the total number of free parameters in the analysis; the alignment size (total number of character traits); whether or not trees were able to be successfully rooted in the maximum likelihood tree search and bootstrap analyses; and whether or not bootstraps converged in the bootstopping test.

Analysis Dataset	Outgroup(s) Used	Final Log Likelihood	AIC Score	AICc Score	BIC Score	Free Parameters (model + branch length)	Alignment Size	Search Trees Rooted	Bootstrap Trees Rooted	Bootstrap Convergence
Concatenated Alignment	Zerg_OG Tyranids_OG Xenomorphs_OG	-2544.78	5227.56	14887.56	5382.71	69	70	Rooted	Not Rooted	Yes
Concatenated Alignment	Protoss_OG Zerg_OG Tyranids_OG Xenomorphs_OG	-2544.78	5227.56	14887.56	5382.71	69	70	Not Rooted	Not Rooted	Yes
Concatenated Alignment	Protoss_OG	-2145.14	4416.27	5760.27	4557.92	63	70	Rooted	Rooted	Yes
Physiological Alignment	Zerg_OG	-636.42	1398.84	9462.84	1491.18	63	32	Rooted	--	--
Behavioral Alignment	Zerg_OG	--	--	--	--	--	--	--	--	--
Archetypal Alignment	Zerg_OG	-639.63	1405.26	9469.26	1461.36	63	18	Rooted	Rooted	No
Physiological Alignment	Protoss_OG	-627.05	1380.10	9444.10	1472.44	63	32	Rooted	Rooted	No
Behavioral Alignment	Protoss_OG	--	--	--	--	--	--	--	--	--
Archetypal Alignment	Protoss_OG	-632.35	1390.70	9454.70	1446.80	63	18	Rooted	Rooted	Yes

substantially across the IPs investigated to prevent them from forming phylogenetic clades based on race when considering all the character traits utilized in this study. The one exception was the orcish taxa, who did not form distinct clades except when the designated outgroup was the Protoss (discussed further below). The long branches separating subsets of the orc taxa are likely due to substantial creative shifts that have been made by IP producers in terms of the archetypes and behaviors characterizing their brand of orcs. For example, the orcs of the *Elder Scrolls* and *Warcraft* IPs (Orsimer_ES and Orcs_WC, respectively) are much more humanistic than orcs portrayed in other intellectual properties. They have families, form emotionally bound communities, and often fight for causes other than the shear desire to express violence and engage in bloodlust; these departures from the vile and retched orcs of Tolkien and others have resulted in the placement of these taxa as sharing a more recent common ancestor with elves, dwarves and humans than with other orcs. Conversely, the more antagonistic and villainous orc taxa shared more recent common ancestors with the monstrous horrors of the Zerg, Tyranid and Xenomorph outgroups, all three of which are characterized as large, physically powerful and overwhelmingly deadly insectoid creatures who stop at nothing to dispatch their foes.

The exception for orcs not appearing to form distinct clades was when the Protoss where designated as the sole outgroup in an analysis. A bi-pedal, psionically adept alien race from the *Starcraft* science-fiction universe, the Protoss, despite being the product of a different genre in fiction with many physiological traits that differ substantially from the Tolkien-derived elf variants, share much of their design as a fictional people in terms of behavior and story roles with that of the elves. Their behavioral and archetypal trait states are shared so much so that whenever the Protoss are assigned as the outgroup in an alignment it causes all the elven taxa to shift on the tree topology from a derived monophyletic group to a basal paraphyletic group. This shift additionally alters the placement of the orc taxa in the phylogenies. The orcs and the elves were consistently the most distantly related groups out of the Tolkien four, indicating they shared the least amount of character states overall. As a result, a shift of the elven taxa to a more basal position on the trees when the Protoss were used as the outgroup always shifted the orc taxa to more derived positions. In the trees where this was the case, the orcs would appear to be monophyletic. However, given the degree of similarity between the Protoss and the elves it seems reasonable to determine that the Protoss do not serve well as an outgroup for the set of taxa used in this study. The greater dissimilarity between the Zerg, Tyranids and Xenomorphs relative to the rest of the dataset made these taxa more suitable, hence the tree from Figure 1 is here the preferred representation of the present data in terms of maximum likelihood inference.

The inferred tree topology suggests that elves, dwarves and humans from across seven different fictional universes share more in common with their respective cross-universe variants than they do with each other, but the tree nodes representing these relationships were for the most part poorly supported in the bootstrap analyses. Although the preferred tree topology from the concatenated sequence alignment suggests an interesting conclusion about the general fidelity of content creators to the origins established by Tolkien, the present study was unable

to show strong statistical support for the relationships inferred. An extension of the dataset developed and used here may be necessary to estimate phylogenetic relationships among these fictional races with greater statistical support.

The scope of the present dataset was even more limiting in the case of the assessment of the categorical sequence alignments. Although this study was unable to compare sets of tree topologies generated from each categorical sequence alignment based on different outgroups, the topologies successfully generated from the archetypal traits alignment, when compared to those generated with the concatenated alignment, hint at the idea that IP creators may be more inclined to bend the Tolkien mold in some categorical ways than others. Future phylogenetic studies of the fictional races of Tolkien would likely benefit from an expanded suite of categorical character traits to better assess in what ways IP creators are more or less likely to depart from established source materials.