

参赛队员姓名: 成果、徐游新

学校: 华南师范大学附属中学

指导教师: 栾云霞、罗宇立

省份: 广东省

国家/地区: 中国

论文题目: 昆虫呼吸蛋白的分子演化

六足动物（昆虫）呼吸蛋白的分子演化

论文摘要:

六足动物（昆虫）是地球上物种最丰富、数量最大的动物类群。前人研究表明，有些昆虫类群使用血蓝蛋白，而有些昆虫类群使用血红蛋白。迄今，呼吸蛋白在昆虫各类群中的分布，及其演化规律与意义仍不明朗。本研究基于GenBank数据和现有的昆虫基因组和转录组数据，全面查找了六足动物所有现生类群中血蓝蛋白和血红蛋白的同源分子序列，阐明了呼吸蛋白在各类群中的分布，并通过构建分子演化树，探讨了昆虫呼吸蛋白的演化规律。另外，我们还发现昆虫的起源、翅的发生、血蓝蛋白的丢失和全变态类群的发生等重要的演化节点都与地球氧气浓度变化密切相关，从而为昆虫的生物多样性的演化机制提供了全新的思路和证据。

关键词:

昆虫 血蓝蛋白 血红蛋白 分子演化 氧气浓度变化

本参赛团队声明所提交的论文是在指导老师指导下进行的研究工作和取得的研究成果。尽本团队所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果。若有不实之处，本人愿意承担一切相关责任。

参赛队员：成果，徐游新

指导老师：栾云霞，罗宇立

2019年 9月 15日

目录

1. 研究背景	1
1.1 血蓝蛋白	1
1.2 血红蛋白	1
1.3 六足动物（昆虫）的演化	2
1.4 地球氧气浓度的变化	2
2. 材料和方法	2
2.1 研究类群和数据来源	2
2.2 昆虫呼吸蛋白候选分子查找	5
2.3 呼吸蛋白分子的序列比对和系统发生分析	5
2.4 昆虫呼吸蛋白的演化与地球氧气浓度变化的整合分析	6
3. 结果	6
3.1 呼吸蛋白在昆虫各目中的分布	6
3.2 昆虫血蓝蛋白和血红蛋白的分子演化	8
3.3 昆虫呼吸蛋白的演化与地球氧气浓度变化的关系	8
4. 讨论	11
5. 参考文献	13
6. 附件	15
6.1 项目分工及致谢	15
6.2 参赛选手简历	15
6.3 辅导教师简历	15

1. 研究背景

血蓝蛋白和血红蛋白为动物界最主要的两类呼吸蛋白。血红蛋白广泛分布于各个动物门类, 而血蓝蛋白迄今只发现于软体动物和节肢动物中 (Burmester, 2002)。六足动物 (昆虫) 隶属于节肢动物门, 是地球上物种最为丰富、数量最大的动物类群。前人的研究 (Burmester 2015) 表明, 有些昆虫类群使用血蓝蛋白, 而有些昆虫类群使用血红蛋白。迄今, 呼吸蛋白在昆虫各类群中的分布, 及其演化规律与意义仍不明朗, 这很可能与昆虫的生物多样性的演化机制密切相关。

1.1 血蓝蛋白

血蓝蛋白 (hemocyanin, Hc) 是仅在节肢动物和软体动物中发现的一种螯合铜的呼吸蛋白, 但节肢动物和软体动物的血蓝蛋白被认为起源于不同的铜离子蛋白 (Burmester 2001)。节肢动物血蓝蛋白起源于酚氧化酶, 具有保守的6个组氨酸 (His) 位点, 协同结合铜离子, 进而结合氧气, 能自由的穿梭在细胞间和血淋巴内, 为生物体运输氧气分子 (Burmester 2002)。在昆虫演化的历程中, 血蓝蛋白发生复制, 其中一支血蓝蛋白的6个保守的组氨酸位点 (His) 发生了不同程度的突变, 失去了结合铜离子的能力, 不能携带氧气, 而转变成为储存蛋白

(hexamerin, Hx)。储存蛋白通常与蜕皮周期有关, 在昆虫非进食时期提供能量和营养, 支持昆虫的发育 (Burmester 1999)。

1.2 血红蛋白

血红蛋白 (Hemoglobin, Hb) 广泛分布于古细菌、细菌、植物、真菌和动物各类生物类群中 (Hardison 1996, Freitas et al. 2004)。血红蛋白隶属于珠蛋白家族 (globin superfamily), 有一个以铁元素为中心的卟啉环结构 (Schmidt-Rhaesa 2007), 能够与氧气可逆性结合, 目前已知的珠蛋白包括血红蛋白、肌红蛋白、神经珠蛋白和细胞珠蛋白。在两侧对称动物中, 血红蛋白成为最重要的氧气输送、储存分子 (Schmidt-Rhaesa 2007), 其大小和组织分布均高度多样化。血红蛋白单体长约为120到170氨基酸, 包含两个 α 球蛋白多肽和两个 β 球蛋白多肽 (Hardison 1996)。

1.3 六足动物（昆虫）的演化

六足动物（昆虫）是地球上演化非常成功、生物多样性最丰富的生物种群，数量几乎占了所有物种的一半（Gullan & Cranston 2005）。现生六足动物类群包括原尾纲、弹尾纲、双尾纲和狭义昆虫纲。狭义昆虫纲有28个目，包括不变态的无翅类群（石蛎目和衣鱼目），不完全变态的有翅类群（蜉蝣目、蜻蜓目、缺翅目、革翅目、襀翅目、直翅目、螳螂目、蚤蟻目、纺足目、蝻目、螳螂目、蜚蠊目、缨翅目、半翅目、啮虫目）和完全变态的有翅类群（膜翅目、蛇蛉目、广翅目、脉翅目、捻翅目、鞘翅目、毛翅目、鳞翅目、蚤目、长翅目、双翅目）。昆虫起源于水生的甲壳动物，从水生到陆生，从陆地到空中，昆虫翅的产生和变态发育（幼体与成虫具有迥异的食物来源、生活环境和形态）的形成对其生物多样性的爆发起着决定性的作用，但这背后的演化机制一直悬而未决，呼吸蛋白的演化可能起着决定性的作用。

1.4 地球氧气浓度变化

氧气对于生命的重要性不言而喻。氧气参与大多数生物体内的新陈代谢，为维持正常生命活动提供能量。在3亿年前，地球大气含氧量曾经上升到35%，当时的昆虫都变得异常巨大，后来在1.2亿年的时间内氧气浓度降至15%，导致很多大型昆虫灭绝（Berner 2009）。因此，地球氧气浓度变化与昆虫利用氧气的方式一定密切相关，但迄今鲜有相关研究和报道。

随着组学测序技术的快速发展，目前昆虫的基因组和转录组数据已有大量报道。基于这些高通量数据信息，本研究在昆虫所有现生类群中查找血蓝蛋白和血红蛋白的同源分子序列，初步阐明呼吸蛋白在昆虫各类群中的分布，并对其序列进行比对分析，构建分子演化树，探讨昆虫呼吸蛋白的演化规律，结合地球氧气浓度变化数据，发掘昆虫的生物多样性的演化机制。

2. 材料和方法

2.1 研究类群和分析数据来源

本研究选取六足动物31个类群的代表物种，具体物种信息及数据来源见表1。

表1 本研究涉及的六足动物（昆虫）的物种及其序列信息¹

类群 Taxonomy	物种 Species	序列名称（序列号） Sequence name (GenBank accession No.)			数据来源 Sources
		I 型血蓝蛋白 (Hc1)	II 型血蓝蛋白 (Hc2)	血红蛋白 (Hb/Gb)	
原尾纲 Protura	<i>Acerentomon</i> sp.	—	—	AceGb(GAXE02041270.1)	Blast 转录组(Misof et al. 2014)
弹尾纲 Collembola	<i>Folsomia candida</i>	FcaHc1(CAR85703.1)	FcaHc2 (KF670722)	FcaGb(GAXI02015810.1)	Genbank 公布 (Xie & Luan 2014) ; Blast 基因组 (Faddeeva-Vakhru sheva et al. 2017)
	<i>Pogonognathellus</i> sp.	PogHc1(QCC26464.1)	PogHc2(QCC26465.1)	?	GenBank 公布
双尾纲 Diplura	<i>Occasjapyx japonicus</i>	—	—	OjaGb(GAXJ02020824.1)	Blast 转录组(Misof et al. 2014)
	<i>Camptodea augens</i>	—	—	?	
石蛎目 Archaeognatha	<i>Machilis germanica</i>	MgeHc1(CAR85692.1)	?	?	GenBank 公布
	<i>Machilis hrabei</i>	?	?	MhrGb(GAUM02028425.1)	Blast 转录组(Misof et al. 2014)
衣鱼目 Zygentoma	<i>Thermobia domestica</i>	TdoHc1 (FM165288)	TdoHc2(CAQ63322.1)	?	GenBank 公布
	<i>Atelura formicaria</i>	?	?	AfoGb(GAYJ02037127.1)	Blast 转录组(Misof et al. 2014)
蜻蜓目 Odonata	<i>Mortonagrion selenion</i>	—	—	?	
	<i>Cordulegaster boltonii</i>	—	—	CboGb(GAYO02026537.1)	Blast 转录组(Misof et al. 2014)
	<i>Epiophlebia superstes</i>	—	—	?	
蜉蝣目 Ephemeroptera	<i>Ecdyonurus insignis</i>	—	—	EinGb(GCCL01023965.1)	Blast 转录组(Misof et al. 2014)
	<i>Eurylophella</i>	—	—	?	
	<i>Isonychia bicolor</i>	—	—	?	
缺翅目 Zoraptera	<i>Zorotypus caudelli</i>	—	—	ZcaGb(GAYA02043503.1)	Blast 转录组(Misof et al. 2014)
革翅目 Dermaptera	<i>Chelidura acanthopygia</i>	CacHc1(CAR85694.1)	CacHc2(CAR85707.1)	?	GenBank 公布
	<i>Hemimerus</i> sp.	?	?	HemGb(GDCU01043786.1)	Blast 转录组(Misof et al. 2014)
襖翅目 Plecoptera	<i>Protonemura ausonia</i>	?	?	PauGb(GDBP01066472.1)	Blast 转录组(Misof et al. 2014)
	<i>Perla marginata</i>	PmaHc1(CAD87762.1)	PmaHc2(CAD87763.1)	?	GenBank 公布
直翅目 Orthoptera	<i>Locusta migratoria</i>	LmiHc1(AIE39570.1)	LmiHc2(AIE39569.1)	?	
	<i>Schistocerca americana</i>	SamHc1(AAC16760.1)	?	?	GenBank 公布
	<i>Calliptamus italicus</i>	CitHc1(AGG40957.1)	?	?	
	<i>Nippancistroger testaceus</i>	?	?	NteGb(GCPE01053008.1)	Blast 转录组(Misof et al. 2014)
螳螂目 Mantophasmatodea	<i>Mantophasma</i> sp.	—	—	ManGb(GDBV01005567.1)	Blast 转录组(Misof et al. 2014)
	<i>Tanzaniophasma</i> sp.	—	—	?	
蜚蠊目 Grylloblattodea	<i>Grylloblatta bifratrilecta</i>	—	—	GbiGb(GAWP02046080.1)	Blast 转录组(Misof et al. 2014)
	<i>Galloisiana yuasai</i>	—	—	?	
纺足目 Embioptera	<i>Antipaluria urichi</i>	—	—	AurGb(GDCK01049722.1)	Blast 转录组(Misof et al. 2014)
	<i>Haploembia palaui</i>	—	—	?	
蟪目 Phasmatodea	<i>Carausius morosus</i>	CmoHc1 (CAR85693.1)	?	?	Genbank 公布
	<i>Antongilia madagassa</i>	?	?	AmaGb(GDYR01088628.1)	Blast 转录组(Misof et al. 2014)
螳螂目 Mantodea	<i>Hierodula membranacea</i>	HmeHc1(CAR85695.1)	HmeHc2(CAR85696.1)	?	GenBank 公布
	<i>Stagmatoptera biocellata</i>	?	?	SbiGb(GDBT01049140.1)	Blast 转录组(Misof et al. 2014)
蜚蠊目 Blattodea	<i>Cryptotermes secundus</i>	CseHc1(CAR85697.1)	CseHc2(CAR85698.1)	?	
	<i>Zootermopsis nevadensis</i>	ZneHc1(KDR21642.1)	ZneHc2(KDR21641.1)	?	GenBank 公布
	<i>Blattella germanica</i>	BgeHc1(PSN36662.1)	BgeHc2(PSN36663.1)	?	
	<i>Blaptica dubia</i>	BduHc1(CAR85699.1)	BduHc2(CAR85700.1)	?	

¹ 其中“—”表示在该物种中未找到该基因,“?”表示未对该物种进行查找

	<i>Mastotermes darwiniensis</i>	?	?	MdaGb(GAZE02155359.1)	Blast 转录组(Misof et al. 2014)
缨翅目 Thysanoptera	<i>Frankliniella occidentalis</i>	—	—	?	Blast 基因组 (Stafford-Banks CA, et al. 2014)
	<i>Franklinothrips vespiformis</i>	?	?	FveGb(GDFM01025047.1)	Blast 转录组(Misof et al. 2014)
半翅目 Hemiptera	<i>Homalodisca vitripennis</i>	—	—	?	GenBank 公布
	<i>Glycaspis brimblecombei</i>	—	—	?	Blast 转录组(Misof et al. 2014)
	<i>Anisops deanei</i>	?	?	AdeHbA(CCG97778.1) AdeHbB(CCG97779.1) AdeHbC(CCG97780.1)	GenBank 公布
啮虫目 Psocodea	<i>Menopon gallinae</i>	—	—	MgaGb(GAWR02010788.1)	Blast 转录组(Misof et al. 2014)
膜翅目 Hymenoptera	<i>Neodiprion pinetum</i>	NpiHc1(SSWZ01000161.1)	—	?	Blast 基因组 (未发表)
	<i>Eupelmus urozonus</i>	—	—	?	Blast 转录组(Misof et al. 2014)
	<i>Cephus cinctus</i>	CciHc1(XP_015605372.1)	CciHc2(XP_015605378.1)	?	GenBank 公布
	<i>Neodiprion lecontei</i>	NleHc1(XP_015519891.1)	NleHc2(XP_015519877.1)	?	
	<i>Apis mellifera</i>	?	?	AmeHb(CAJ43389.1)	
	<i>Linepithema humile</i>	?	?	LhuGb(XP_012231019.1)	
蛇蛉目 Raphidioptera	<i>Inocellia crassicornis</i>	—	—	IcrGb(GAZH02009168.1)	Blast 转录组(Misof et al. 2014)
	<i>Subilla</i>	—	—	?	
广翅目 Megaloptera	<i>Corydalus cornutus</i>	—	—	CcoGb(GATG02010693.1)	Blast 转录组(Misof et al. 2014)
脉翅目 Neuroptera	<i>Osmylus fulvicephalus</i>	—	—	OfuGb(GAYC02012919.1)	Blast 转录组(Misof et al. 2014)
	<i>Euroleon nostras</i>	—	—	?	
捻翅目 Strepsiptera	<i>Mengenilla moldrzyki</i>	—	—	?	Blast 基因组 (Niehuis O et al. 2012)
	<i>Stylops melittae</i>	?	?	SmeGb(GAZM02010098.1)	Blast 转录组(Misof et al. 2014)
鞘翅目 Coleoptera	<i>Tribolium castaneum</i>	—	—	?	Blast 基因组 (Kim HS et al. 2010)
	<i>Aethina tumida</i>	?	?	AtuGb(XP_019881478.1)	GenBank 公布
毛翅目 Trichoptera	<i>Annulipalpia sp.</i>	—	—	AnuGb(GATX01081972.1)	Blast 转录组(Misof et al. 2014)
	<i>Rhyacophila fasciata</i>	—	—	?	
鳞翅目 Lepidoptera	<i>Amata phegea</i>	—	—	?	Blast 转录组(Misof et al. 2014)
	<i>Triodia sylvina</i>	—	—	?	
	<i>Vanessa tameamea</i>	?	?	VtaGb(XP_026490893.1)	GenBank 公布
蚤目 Siphonaptera	<i>Ctenocephalides felis</i>	—	—	?	Blast 基因组(Misof B, et al.2014)
	<i>Ceratophyllus gallinae</i>	—	—	CgaGb(GAWK02008448.1)	Blast 转录组(Misof et al. 2014)
长翅目 Mecoptera	<i>Boreus hyemalis</i>	—	—	Bhy(GAYK02021577.1)	Blast 转录组(Misof et al. 2014)
	<i>Panorpa vulgaris</i>	—	—	VtaHb(XP_026490893.1)	
双翅目 Diptera					
	<i>Gasterophilus intestinalis</i>	?	?	GinHb (AAC80435.1)	GenBank 公布
	<i>Drosophila melanogaster</i>	—	—	DmeHb(CAB52584.1)	GenBank 公布 ; Blast 基因组(Berlin K, et al. 2015)
	<i>Chironomus thummi</i>	?	?	CthHbVIIIB5-9 (CAA39712-	GenBank 公布

				AA39716) CthHbVIIIBZ, V, W, Y (CAA39717-CAA 39720)	
--	--	--	--	----------------------------------------------------------------	--

2.2 昆虫呼吸蛋白候选分子查找

我们采取了三种方法查找昆虫的呼吸蛋白分子。首先, 在GenBank的 Protein 库中直接查找每个代表类群的血蓝蛋白 (hemocyanin) 和血红蛋白 (hemoglobin) 的氨基酸序列, 找到后保存为fasta格式的本地文件。

未能直接找到相关蛋白的代表类群采取第二种方法继续查找, 在GenBank中的Genome库里搜索该目已测定的基因组数据, 挑选数据较完整、拼接程度较高的物种, 并从先前查找到的序列中筛选几条较为完整的确定的序列作为query (血蓝蛋白query包括衣鱼的TdoHc1和TdoHc2、飞蝗的LmiHc1和LmiHc2、松叶蜂的NleHc, 血红蛋白的query包括半翅目的AdeHbA、膜翅目的AmeHb1、双翅目的CthHbVIIIB-9和DmaHb1、鳞翅目的VtaHb1、鞘翅目的AtuHb1) 对基因组组装数据进行在线tblastn比对。将血蓝蛋白比对结果中Identities>40%的序列和血红蛋白比对结果中Identities>30%的序列分别下载保存 (基于前期的数据分析发现, 血蓝蛋白的序列相似度>40%, 血红蛋白的序列相似度>30%)。

对于其余尚未找到血蓝蛋白和血红蛋白序列的类群, 我们采用本地 Blast 进一步挖掘转录组中潜在的血蓝蛋白家族分子。安装 Blast 程序至本地后, 我们下载使用 Misof et al. (2014) 报道的昆虫转录组数据, 将上述血蓝蛋白和血红蛋白的 query 序列作为检索序列, 分别在各昆虫代表类群 cDNA 测序组装的 TSA 数据中, 用 tblastn 进行比对分析, 阈值为 $E \times 10^{-5}$ (Camacho et al. 2009)。将血蓝蛋白比对结果中 Identities>40%的序列和血红蛋白比对结果中 Identities>30%的序列分别下载保存。

2.3 呼吸蛋白分子的序列比对和系统发生分析

将获得的候选血蓝蛋白和血红蛋白序列分别使用BioEdit (Hall, T.A. 1999.) 中ClustalW方法, 进行序列对齐。

血蓝蛋白序列较长, 比较保守, 其中有六个参与结合氧气的组氨酸 (His) 高度保守, 将这些His中有突变位点的序列 (储存蛋白Hexamerin, Hx) 从后续分

析中去除。

血红蛋白序列较短（150aa左右），序列变异较大，仅根据序列比对很难断定是否真正是血红蛋白的直系同源分子。我们筛选保留了具有保守氨基酸（结合氧气的两个His，与heme结合的Phe和Leu等）的序列。

使用MEGA（Kumar et al. 2018）的邻接法（Neighbor-joining）对获得的血蓝蛋白和血红蛋白候选序列分别进行系统发生分析，重抽样（Bootstrap）检验1000次，甲壳动物血蓝蛋白和血红蛋白分别设为外群。根据聚类关系，将未与已知血蓝蛋白或血红蛋白聚类的序列删除。

最终保留的候选血蓝蛋白序列命名为血蓝蛋白Hc，而得到的候选血红蛋白的序列变异度较大，仅根据序列比对和构树聚类，只能确认是珠蛋白家族的成员，不能完全确定是否是血红蛋白的直系同源分子，所以我们暂时命名为珠蛋白Gb，与已经通过实验确认过的血红蛋白Hb做以区分。

将最终确认的血蓝蛋白和血红蛋白（珠蛋白）序列分别使用MEGA（Kumar et al. 2018）的邻接法（Neighbor-joining）构建分子系统树，氨基酸替代模型设为Jones-Taylor-Thornton model (JTT)，Gap/Missing data treatment设为Pairwise deletion。

2.4 昆虫呼吸蛋白的演化与地球氧气浓度变化的整合分析

将目前较为公认的昆虫系统演化树（Misof et al. 2014）与地质演化年代以及地球氧气浓度变化数据（Berner 2009）整合分析，加入血蓝蛋白和血红蛋白（珠蛋白）的演化信息，展示昆虫呼吸蛋白的演化与地球氧气浓度变化的关系。

3. 结果

3.1 呼吸蛋白在昆虫各目中的分布

呼吸蛋白在昆虫各目中的分布见表 1。

我们在 10 个类群中找到血蓝蛋白，包括不变态的无翅类（弹尾纲、石蛎目和衣鱼目）、不完全变态的复新翅类（襁翅目、蜚蠊目、螳螂目、螽蟴目、直翅目和革翅目），以及全变态类的内翅类（膜翅目）。值得注意的是，我们首次在膜翅目三个物种（*Neodiprion pinetum*, *Cephus cinctus*, *Neodiprion lecontei*）中找到

了血蓝蛋白(表 1)。这是目前发现的唯一使用血蓝蛋白的全变态类昆虫类群，从序列比对(图 1)可以看出，膜翅目血蓝蛋白序列与其它血蓝蛋白高度保守，6 个保守的 His 位点显示它们具有结合氧气的的能力。

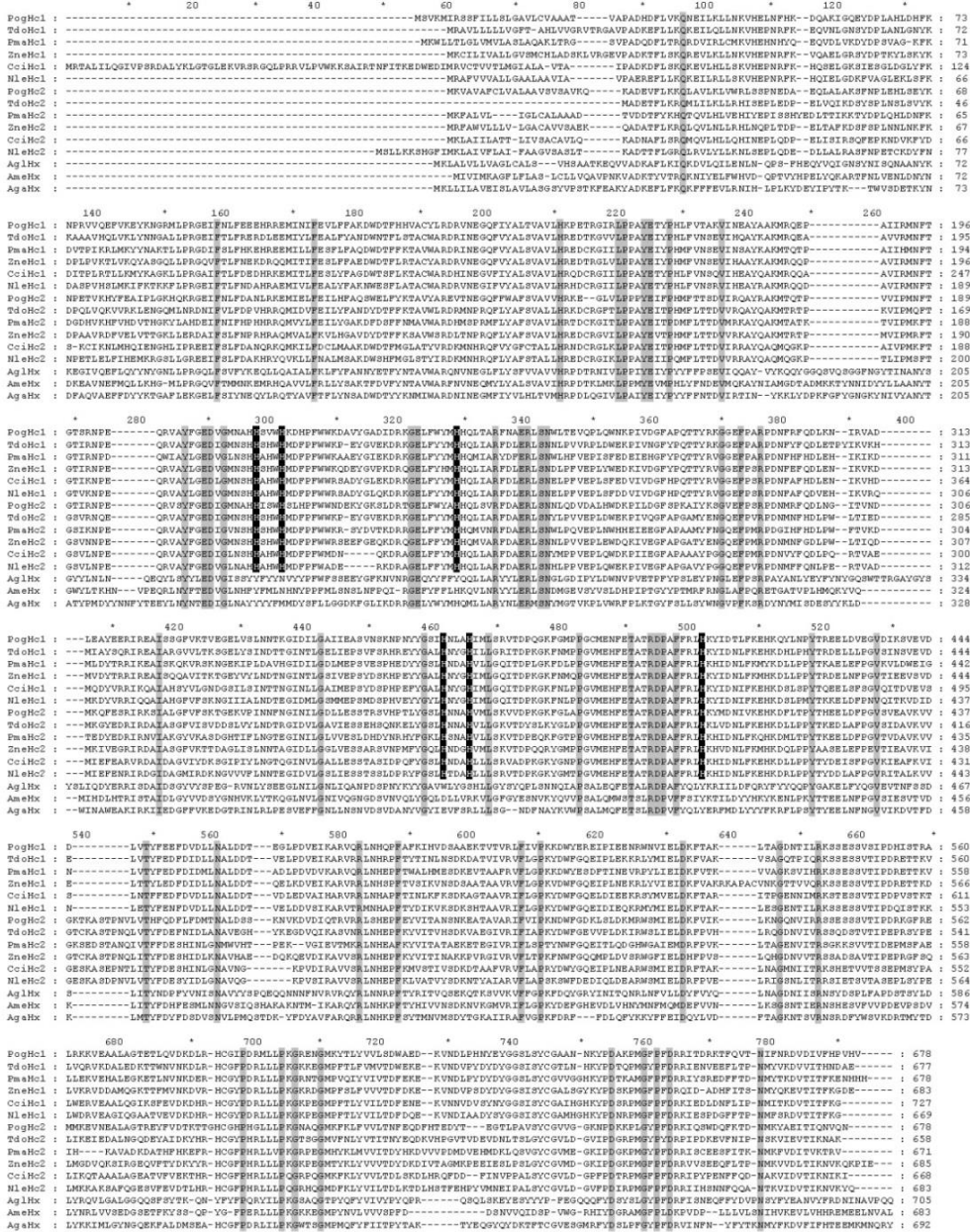


图 1 部分昆虫血蓝蛋白 Hc1、Hc2 和储存蛋白 Hx 的多重序列比对。黑色阴影部分显示 Hc 中高度保守、结合铜离子运输氧气的 6 个组氨酸 (Hx 中均有不同程度的丢失)，浅灰色阴影显示其它保守位点。血蓝蛋白序列选自：弹尾目 (*Pogonognathellus* sp. AD2013, PogHc1 和 PogHc2)、衣鱼目 (*Thermobia domestica*, TdoHc1 和 TdoHc2)、積翅目 (*Perla marginata*, PmaHc1 和 PmaHc2)、蜚蠊目 (*Zootermopsis nevadensis*, ZneHc1 和 ZneHc2) 和膜翅目 (*Cephus cinctus*, CciHc1 和 CciHc2; *Neodiprion lecontei*, NleHc1 和 NleHc2)，昆虫储存蛋白序列选自：鞘翅目 (*Anoplophora glabripennis*, AglHx)、膜翅目 (*Apis mellifera*, AmeHx)、双翅目 (*Anopheles gambiae*, AgaHx)。

有趣的是, 我们在几乎所有六足动物(昆虫)类群的转录组数据中都找到了血红蛋白的同源序列(表 1)。如前所述, 血红蛋白序列保守性较低, 仅凭氨基酸序列比对难以判断其是否是血红蛋白的直系同源分子。因此, 前人已通过结构和功能确认的序列(包括半翅目、双翅目、膜翅目、鳞翅目和鞘翅目)我们用 Hb 表示, 本研究新查找的序列, 我们暂用珠蛋白(Gb)表示, 留待后续实验验证是否是真正的血红蛋白。无论如何, 我们获得的序列都具有保守的氧气结合位点(His), 表明相关蛋白均具有携氧能力。

3.2 昆虫血蓝蛋白和血红蛋白的分子演化

六足动物(昆虫)血蓝蛋白在祖先已发生复制, 分化为 Hc1 和 Hc2 两支, 随后各自演化(图 2)。[[(蜚蠊目+螳螂目)+蝽目]+直翅目]的 Hc1 和 Hc2 均与其物种聚类关系一致, 说明昆虫血蓝蛋白 Hc1 和 Hc2 的分子演化与昆虫早期分化出来的类群的演化关系较好地吻合, 暗示昆虫血蓝蛋白运输氧气的功能高度保守。

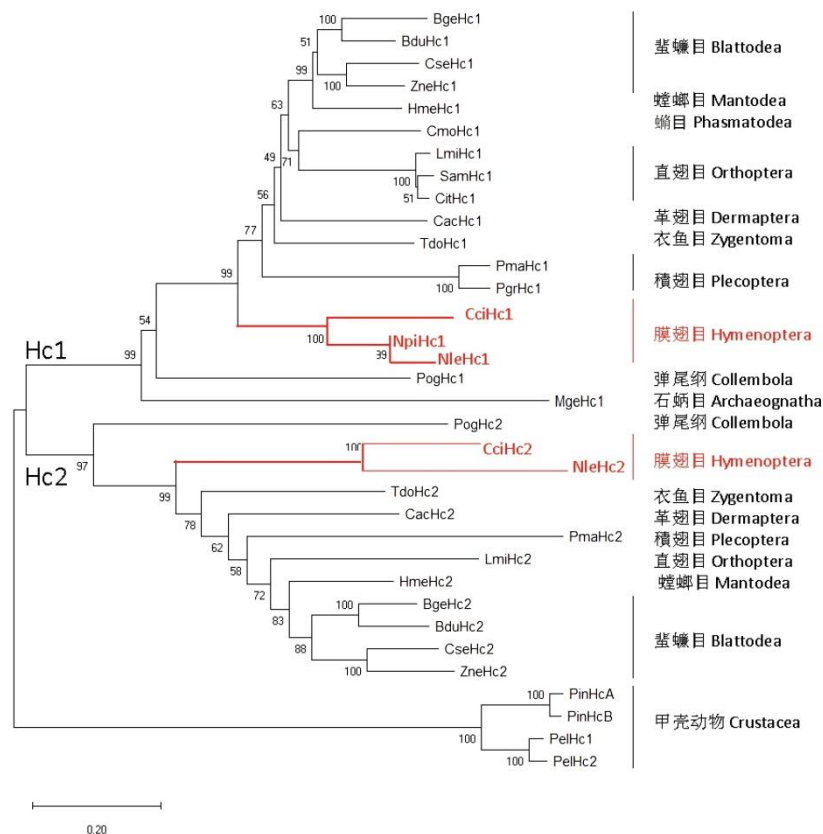


图 2 昆虫血蓝蛋白分子系统发生树(NJ tree)。以甲壳动物(*Panulirus Interruptus*, *Palinurus Elephas*)血蓝蛋白氨基酸序列为外群。六足动物缩写对应的物种和序列信息见表 1。膜翅目血蓝蛋白分子用红色突出显示。每个节点的数字为其对应的重抽样检验值(Bootstrap 1000 次)。

需要特别关注的是, 我们首次在膜翅目类群中发现 Hc1 和 Hc2, 而在其它不完全变态的副新翅类 (Paraneoptera) 和全变态类群中均未发现血蓝蛋白的踪迹。从血蓝蛋白的分子系统树上看, 膜翅目血蓝蛋白 Hc1 和 Hc2 分别聚类在昆虫血蓝蛋白较外源的位置, 这与膜翅目在昆虫演化过程中较晚出现的时间并不一致, 结合其它较高等类群都没有血蓝蛋白的事实, 说明血蓝蛋白在昆虫演化过程中已发生丢失, 膜翅目血蓝蛋白是后期单独通过某种未知的方式和来源重新获得的。

六足动物 (昆虫) 血红蛋白 (珠蛋白) 的分子系统树 (图 3) 表明该蛋白序列变异较大, 在某些类群中高度复制 (例如双翅目), 分子演化模式与六足动物 (昆虫) 物种的演化不一致, 暗示该蛋白功能趋于多元化。

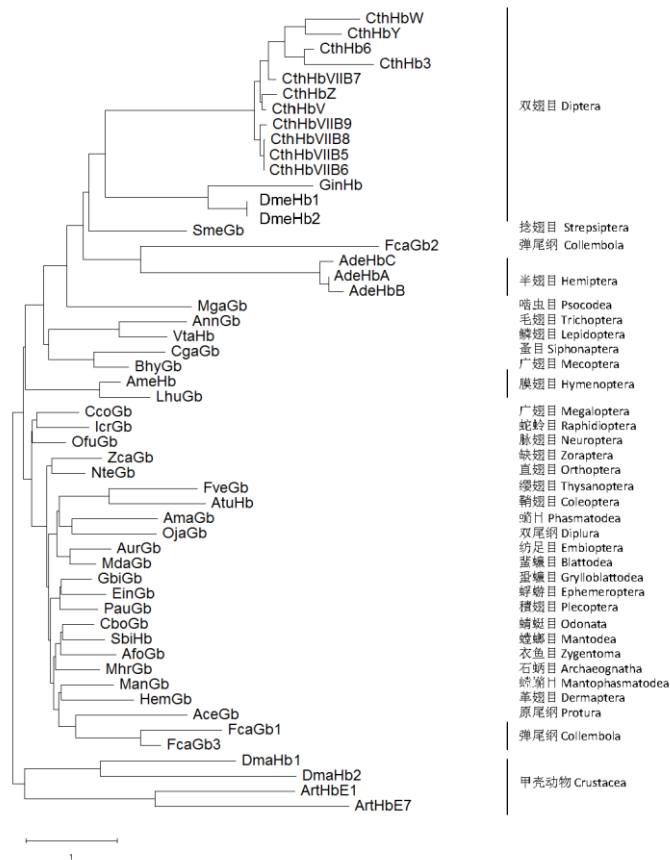


图 3 昆虫血红蛋白 (珠蛋白) 的分子系统发生树 (NJ tree)。以甲壳动物 (*Daphnia magna*, *Artemia* sp.) 血红蛋白氨基酸序列为外群。六足动物缩写对应的物种和序列信息见表 1。

3.3 昆虫呼吸蛋白的演化与氧气浓度的关系

通过整合六足动物 (昆虫) 各类群的系统发生关系、分化时间和呼吸蛋白的分布,

以及地球氧气浓度变化等信息, 我们发现其中存在着密切的关联(图4)。六足动物(昆虫)起源于约4.8亿年前, 当时地球氧气浓度处于历史上中等水平(~18%), 为源自水生甲壳动物的六足动物登陆到陆地生活创造了有利条件。六足动物(昆虫)在登陆后, 沿用来自甲壳动物的血蓝蛋白运输氧气。随着地球氧气浓度逐渐升高, 昆虫在地球氧气浓度达到一个小高峰(24%左右)时演化出了翅的结构, 充足的氧气使得昆虫得以增加耗能, 在空中飞行, 大大加强生活能力。随后氧气浓度开始迅速下降, 在3.8亿年前, 地球氧气降至昆虫存在以来的最低点(15%左右), 新演化出来的昆虫类群开始弃用血蓝蛋白, 使用效率更高的血红蛋白输送氧气, 进行新陈代谢活动。随后, 氧气浓度又快速上升, 在3.4亿年前, 氧气浓度上升至20%以上, 昆虫演化出了更为高级的全变态发育方式, 克服恶劣环境的能力再次大大加强。随着氧气浓度的继续上升, 在3亿年前, 当地球氧气浓度处于最高峰(>30%)时, 昆虫类群出现了爆发式的增长。在随后漫长的昆虫演化过程中, 氧气浓度又迅速下降, 很多昆虫类群出现了不同的应对策略, 例如膜翅目的某些类群再次获得了血蓝蛋白, 而双翅目昆虫中血红蛋白大量复制。

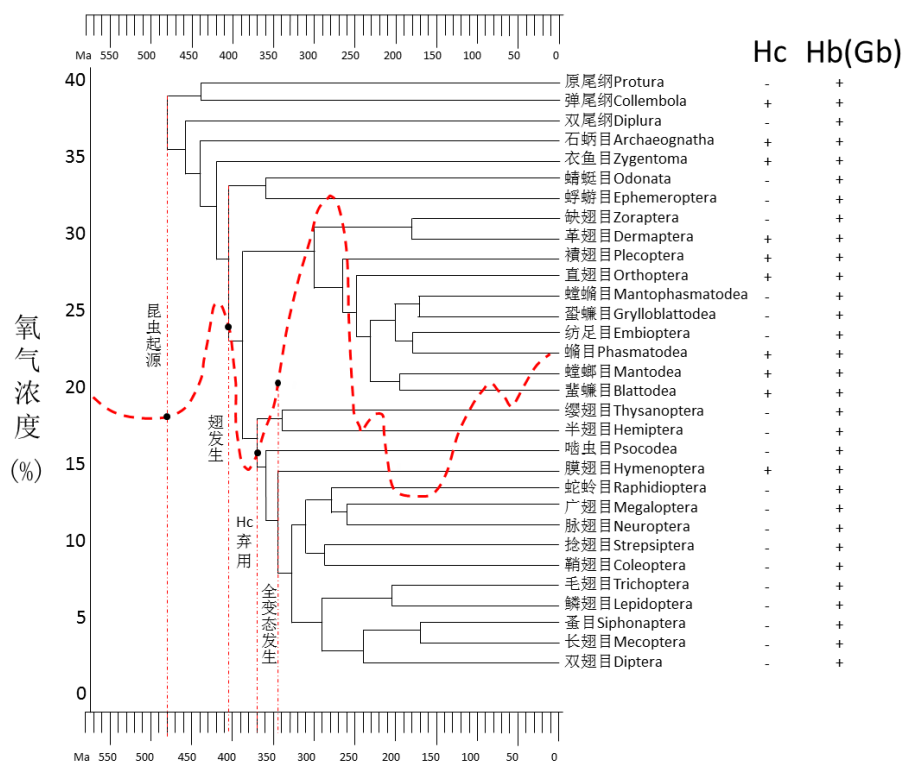


图4 六足动物(昆虫)各类群的系统发生关系、分化时间和呼吸蛋白的分布, 及其与地球氧气浓度变化的潜在关系。昆虫系统演化树引自 Misof et al. (2014), 地球氧气浓度变化数据引自 Burner (2009)。

4. 讨论

本研究全面揭示了呼吸蛋白在六足动物（昆虫）各类群中的分布情况。如前人（XIE Wei, LUAN Yun-Xia, 2011）所述，血蓝蛋白主要分布在较低等的昆虫类群中，包括不变态无翅类和不完全变态的复新翅类。而与前人研究（XIE Wei, LUAN Yun-Xia, 2011）所不同的是，我们在所有类群中都找到了血红蛋白的同源序列，尽管只根据序列和分子系统树，我们不能断定它们一定是血红蛋白的直系同源分子，但我们获得的序列都具有保守的氧气结合位点（His），表明它们应该具有一定的携氧能力。其实，血红蛋白广泛存在于古细菌、细菌、植物、真菌和动物各类生物类群中（Hardison 1996, Freitas et al. 2004），在六足动物（昆虫）中广泛存在也不足为奇。

本研究首次确认高等昆虫全变态类群的膜翅目中存在血蓝蛋白，更新了我们对于昆虫血蓝蛋白的认知。不过，基于血蓝蛋白在昆虫中的分布，及其分子系统树，我们同意前人的观点：昆虫血蓝蛋白在不变态类群的副新翅类和全变态类群的共同祖先中已经丢失（Burmester 2015 ; Pick et al. 2009），后期演化过程中，膜翅目的某些类群可能是通过某种水平转移的方式重新获得了血蓝蛋白分子，并保留下来。使用血蓝蛋白的膜翅目类群是否有某些特征或共性，尚有待于进一步深入的研究。

在揭示了六足动物（昆虫）呼吸蛋白演化规律的基础上，我们发现地球氧气浓度变化也与之密切相关。六足动物（昆虫）起源于水生的甲壳动物，由于水的溶氧量很低，在使用血红蛋白的基础上，甲壳动物需要使用血蓝蛋白加强呼吸作用（Burmester 2015）。4.8亿年前，在六足动物登陆后，部分类群继续沿用血蓝蛋白，但有些类群各自丢失了血蓝蛋白，这其中可能与空气含氧量较高、某些类群活动能力弱、以及昆虫气管系统的形成有关，较低含量的血红蛋白足以满足它们的生活要求。

六足动物（昆虫）出现后，地球氧气浓度逐渐升高，在达到 24% 左右时，昆虫演化出了翅的结构，昆虫的代谢速率加快、运动能力增强，而当时地球氧气浓度可以支撑其进行消耗大量能量和氧气的飞行运动。

随后氧气浓度开始迅速下降，3.8 亿年前，地球氧气降至昆虫存在以来的最低点（15% 左右），新演化出来的昆虫类群全面弃用血蓝蛋白，而使用大量血红

蛋白输送氧气和进行新陈代谢活动, 因为血红蛋白运输氧气和进行机体代谢的效率远高于血蓝蛋白, 具有更强的携氧能力 (Mangum 1985; Jameson & Ibers 1994), 可支持机体进行更为剧烈、复杂的运动。

随后, 氧气浓度又快速上升, 血红蛋白的高效利用促使昆虫演化出了更为高级的全变态发育方式, 进一步增强了克服恶劣环境的能力。3 亿年前左右的晚石炭期, 地球氧气浓度 >30%, 昆虫类群出现了爆发式的增长。随后氧气浓度又迅速下降, 很多昆虫类群出现了不同的应对策略, 例如双翅目昆虫中血红蛋白大量复制, 而膜翅目的某些类群再次获得了血蓝蛋白。本研究六足动物 (昆虫) 呼吸蛋白的分子演化分析, 为昆虫的生物多样性的演化机制提供了全新的思路和证据。

5. 参考文献

- Berner RA.2009.Phanerozoic atmospheric oxygen: new results using the GEOCARBSULF model. *American Journal of Science*, 309: 603–606.
- Burmester.1999.Evolution and function of the insect hexamerins. *Eur J Entomol* 96:213-25.
- Burmester.2001.Molecular evolution of the arthropod hemocyanin superfamily. *Mol Biol Evol*.Feb;18(2):184-95.
- Burmester.2002.Origin and evolution of arthropod hemocyanins and related proteins. *J Comp Physiol B*. Feb;172(2):95-107.
- Burmester T.2015.Evolution of respiratory proteins across the Pancrustacea. *Integrative and Comparative Biology*, 55: 792–801.
- Camacho et al.2009.BLAST+: architecture and applications. *BMC Bioinformatics*. volume 10, Article number: 421 (2009).
- Faddeeva-Vakhrusheva A, Kraaijeveld K, Derks MFL, et al. 2017.Coping with living in the soil: the genome of the parthenogenetic springtail *Folsomia candida*. *BMC Genomics*, 18: 493.
- Gullan PJ, Cranston PS. 2005 *The Insects: An Outline of Entomology*. 3rd Edition, Blackwell Publishing Ltd., Hoboken, 505 p. (彩万志, 花保祯, 宋敦伦, 梁广文, 沈佐锐 译. 2009 昆虫学概论. 第3版. 北京: 中国农业大学出版社).
- Hardison 1996, Freitas et al. 2004 A brief history of hemoglobins: Plant, animal, protist, and bacteria. *Proc Natl Acad Sci USA* 93:5675-9.
- Kumar S, Stecher G, Li M, et al. 2018.MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*, 35:1547-1549.
- Misof et al. 2014 The 1KITE project: evolution of insects. *Science*, 2014 Nov 7;346(6210):763-7.
- Schmidt-Rhaesa A. 2007 *The evolution of organ systems*. Oxford: Oxford University Press.
- Burmester T. , Hankeln,T.1999.: A globin gene of *Drosophila melanogaster*. *Mol. Biol. Evol.* 16 (12), 1809-1811 (1999).
- Mangum CP, Scott JL, Black REL, et al.1985.Centipedal hemocyanin: its structure and implication for arthropod phylogeny. *Proc Natl Acad Sci USA*, 82(11):3721-5.

Jameson, Ibers, et al.1994.Biological and Synthetic Dioxygen Carriers. In: Ibertini et al., Bioinorganic Chemistry. University Science Books, Mill Valley, CA.

Berlin K, et al. 2015.Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. Nat Biotechnol. Jun;33(6):623-30.

Stafford-Banks CA, et al. 2014. Analysis of the salivary gland transcriptome of *Frankliniella occidentalis*. PLoS One. 2014 Apr 15;9(4):e94447.

Niehuis O et al.2012.Genomic and morphological evidence converge to resolve the enigma of Strepsiptera. Curr Biol. 2012 Jul 24;22(14):1309-13.

Kim HS et al. 2010. The genome of the model beetle and pest *Tribolium castaneum*. Nature. 2008 Apr 24;452(7190):949-55.

XIE Wei, LUAN Yun-Xia.2011.Jan.Members of the arthropod hemocyanin superfamily and their evolution. Chinese Bulletin of Life Sciences. 1004-0374(2011)01-0106-09.

6. 附件

6.1 项目分工及致谢

在研究过程中, 老师协助我们理清研究思路, 对我们进行了操作方法上的指导, 为我们搭建了论文的基本框架并且指导我们不断完善论文的写作, 所给予的帮助均为无偿指导。成果和徐游新同学共同参与并且完成了数据查找、序列搜索比对、数据分析收集以及后期系统发育树等图表填写与制作的全过程, 论文撰写的大部分工作由成果同学承担。

感谢两位老师对本研项目研究过程及论文撰写和修改的指导!

6.2 参赛选手简历

成果, 女, 2003年5月13日出生于广州, 初中就读于华南师范大学附属中学初中部奥班, 现为华南师范大学附属中学奥林匹克竞赛班生物组高二级的学生。她德智体美劳全面发展, 特长突出, 为学校高二生物竞赛组的一员。她于2019年6月获得全国中学生生物联赛二等奖, 于2019年4月参加英国生物竞赛获入围奖, 并于2019年参加中山大学生命科学学院第一届全国优秀高中生“生命色彩”夏令营。她还于2018年11月参加美国数学建模比赛获得二等奖, 曾于2016年5月参加美国学术五项全能获社会科学金奖、数学金奖、艺术铜奖、总分铜奖, 于2019年5月被评为广东省“最美南粤少年”之“智慧好少年”。

徐游新, 女, 2002年11月28日出生于广州, 初中就读于华南师范大学附属中学初中部奥班, 现为华南师范大学附属中学奥林匹克竞赛班生物组高二级的学生。她德智体美劳全面发展, 生物特长突出, 于2019年6月获全国中学生生物联赛二等奖, 参加英国生物竞赛获入围奖, 并于2019年暑假参加中山大学生命科学学院第一届全国优秀高中生“生命色彩”夏令营。

6.3 辅导教师简历

栾云霞, 博士、研究员, 主要研究方向为综合分子生物学、遗传发育生物学和形态学等方法研究低等昆虫的系统发生和演化, 探讨昆虫的起源、翅的发生及其变态发育的演化等问题, 曾任职于中国科学院上海植物生理生态研究所, 现任职于广东省昆虫发育生物学与应用技术重点实验室, 华南师范大学生命科学学院。

先后主持 7 项国家自然科学基金面上项目, 2 项国家自然科学基金国际合作交流项目和 1 项国家转基因生物新品种培育科技重大专项子课题。

罗宇立, 华南师范大学附属中学, 生物一级教师, 奥林匹克竞赛教练, 有多年教学及辅导学生参赛经验, 指导学生多次在生物方面获国际、国内大赛奖项。