

びRNAL OF FISHERIES OF CHINA





DOI: 10.11964/jfc.20210712937

中华绒螯蟹代谢蛋白互作网络构建及分子功能、 亚细胞定位和途径分析

杨佳睿,郝 彤*,李倩一,孙金生* (天津师范大学生命科学学院,天津市动植物抗性重点实验室,天津 300387)

摘要:为了构建中华绒螯蟹代谢过程研究的系统工具,实验在已经构建的中华绒螯蟹蛋白 互作网络的基础上,首先采用邻接节点注释法对未知蛋白的分子功能进行预测。随后采用 GO 回溯法,构建了代谢蛋白网络并对网络中蛋白分子功能、亚细胞定位和途径分布进行 了分析。分子功能注释中,确定了932个蛋白的分子功能,占所有未知分子功能蛋白的 97%。最终构建的代谢蛋白互作网络包含2045个蛋白及这些蛋白之间的15927条互作关 系。网络中94.2% (1926/2045)的蛋白具有亚细胞定位信息,大多分布于有膜细胞器中; 96.1% 的蛋白 (1966/2045)具有分子功能信息,大多具有催化活性和结合活性。进一步对 确定了分子功能和亚细胞定位的蛋白在40个 KEGG 子系统中的分布进行分析,发现参与 翻译和氨基酸代谢过程的蛋白较多,也有一部分参与免疫和运输过程。本实验结果可为中 华绒螯蟹代谢相关的蛋白功能、定位的研究提供重要的数据参考,对系统研究中华绒螯蟹 代谢过程及代谢相关疾病具有重要价值。

关键词:中华绒螯蟹;蛋白互作网络;分子功能;亚细胞定位;代谢 中图分类号:Q 591.2; S 917.4 文献标志码:A

蛋白互作是指在同一生物体内,2个或多个 蛋白质相互作用,共同完成某个生理功能的过程。 遵循互作的蛋白质可能处于同一网络中或处于相 同的分子途径这一原理,实验者可依靠蛋白互作 的信息从而重构途径。随着技术的逐步提高,生 物网络现已广泛运用到生物研究的各个领域。生 物网络的研究主要集中于蛋白互作网络、代谢网 络、信号转导网络、转录调控网络等。蛋白功能 的预测始终是非常重要的研究目标与难题,随着 后基因组时代的到来,基于越来越多的转录组测 序数据,蛋白互作网络已经成为研究蛋白功能的 重要工具。所以越来越多的学者着手于研究蛋白 互作网络,并通过分析该网络进行蛋白功能预测 和分析。至今为止,蛋白互作网络的构建涉及到 很多物种,包括真核生物、原核生物、病毒、动 物和植物。其中,模式生物的蛋白互作网络的研 究尤为深入。目前获取大规模蛋白互作关系的技 术包括酵母双杂交实验^[1]、双分子荧光互补技术^[2]、 噬菌体展示技术^[3-4]、免疫共沉淀法^[5-7]、蛋白芯片 技术^[3-9]、计算机预测^[10]等。其中,酵母双杂交实 验是蛋白互作网络构建时最常使用的互作关系技 术手段,酿酒酵母(*Saccharomyces cerevisiae*)^[11-12]、

- 收稿日期: 2021-07-03 修回日期: 2021-11-01
- 资助项目:国家重点研发计划 (2018YFD0901301);国家自然科学基金 (31770904);天津市人才发展特殊支持计划 高层次创新创业团队项目 (ITTFRS2017007);天津市高等学校创新团队建设规划 (TD13-5076);天津市 自然科学基金 (19JCYBJC29700)
- 第一作者:杨佳睿(照片),从事水生生物学生物信息学研究, E-mail: yanng0524@163.com
- 通信作者:郝彤,从事生物信息学、代谢网络、蛋白互作网络构建和分析研究,E-mail: skyht@tjnu.edu.cn;孙金生,从事水生生物生理、免疫研究,E-mail: skysjs@tjnu.edu.cn

版权所有 ©《水产学报》编辑部(CC BY-NC-ND 4.0) 中国水产学会主办 sponsored by China Society of Fisheries Copyright © Editorial Office of Journal of Fisheries of China (CC BY-NC-ND 4.0) https://www.china-fishery.cn

黑腹果蝇 (Drosophila melanogaster)^[13]、秀丽隐杆 线虫 (Caenorhabditis elegans)^[14]、大肠杆菌 (Escherichia coli)^[15]、人 (Homo sapiens)^[16] 等许多物 种的蛋白质互作关系均在酵母双杂交实验的基础 上构建完成。对于没有高通量实验检测蛋白互作 关系的物种, 计算机预测则成为其互作网络构建 时蛋白互作关系的主要技术手段。2009年, De Bodt 等^[17] 借助计算机预测方法,通过整合各数据 库,分析了拟南芥 (Arabidopsis thaliana) 与黑腹果 蝇、秀丽隐杆线虫、酿酒酵母和人的同源蛋白数 据,构建了拟南芥的蛋白互作网络。此外,幽门 螺旋杆菌 (Helicobacter pylori)^[18],枯草芽孢杆菌 (Bacillus subtilis)^[19]等生物蛋白互作网络也通过计 算机预测方法构建。蛋白互作网络应用十分广泛, 主要包括预测未知蛋白功能^[20]、功能子网络^[21-22]、 进化分析^[23]和 Hub 蛋白分析^[24]等。

中华绒螯蟹 (Eriocheir sinensis), 俗称大闸蟹、 河蟹,隶属于十足目 (Decapoda) 绒螯蟹属 (Eriocheir),是我国重要的经济水产动物。然而,在中 华绒螯蟹的养殖过程中,由代谢引起的各种疾病 造成了中华绒螯蟹的营养需求异常,从而造成其 产量大幅下降。因此了解中华绒螯蟹代谢机制对 于其健康养殖尤为重要。有关中华绒螯蟹代谢活 动方面的研究较少,Zou等^[25-26]关于中华绒螯蟹 在温度和低溶解氧含量等对其代谢进行了相关研 究。动物的代谢过程所需要的能源物质有蛋白质、 脂质和碳水化合物。根据这3种物质在代谢活动 中所占的比例,可以将甲壳动物的代谢活动分为 以下3种不同的类型:蛋白质为主要能源型、蛋 白质与脂质混合型、脂质和碳水化合物混合型。 由于水生动物的代谢活动在碳水化合物的利用上 较为微量,因此蛋白质和脂质成为提供中华绒螯 蟹代谢活动能量的主要来源[27],对蛋白质的功能、 位置等信息进行分析也成为对中华绒螯蟹代谢过 程进行深入分析的重要途径。

本实验室前期利用中华绒螯蟹的肝胰腺、眼 柄和Y器官的转录组测序数据,利用计算机方法 构建了中华绒螯蟹蛋白互作网络,该网络是第一 个水生甲壳类生物的蛋白互作网络。网络中含有 3223个蛋白质和35787条互作关系^[28]。并在此基 础上,运用邻接节点注释的方法,对未知功能的 蛋白质进行生物过程(biological process)^[28]和细胞 组分(cellular component)^[29]2个方面的GO注释。 为了探索中华绒螯蟹的代谢过程,本实验基于实 验室前期构建的中华绒螯蟹蛋白互作网络,对其 进行了分子功能 (molecular function) 的注释,并通 过 GO 注释回溯提取了其中的代谢子网络,从而 构建了中华绒螯蟹代谢蛋白互作网络,并分析了 网络中蛋白的分子功能、亚细胞定位和途径,为 中华绒螯蟹代谢蛋白的研究提供了数据参考。构 建中华绒螯蟹代谢蛋白互作网络有助于了解甲壳 动物代谢过程中相关蛋白互作过程,为进一步研 究甲壳动物代谢活动提供了数据基础与理论依据。

1 材料与方法

1.1 中华绒螯蟹蛋白互作网络蛋白分子功能注释

在前期研究中,构建了中华绒螯蟹蛋白互作 网络,该网络包含3223个蛋白质和35787条互 作关系^[28]。在此基础上,基于邻接节点注释的方 法对未知蛋白进行了GO注释,分别在生物过程^[28] 和细胞组分^[29]2个方面进行了注释。通过这种方 法,分别为549个蛋白添加了生物过程注释,为 830个蛋白添加了细胞组分注释。注释完成后, 在生物过程和细胞组分2个方面含有注释的蛋白 分别占总网络的94.5%和97.7%。

中华绒螯蟹蛋白互作网络中包含了很多缺乏 分析功能注释的蛋白,本实验首先采用相同的方 法对这些分子功能未知的蛋白进行了注释。基于 蛋白网络中已经具备分析功能注释信息的蛋白, 依据相邻蛋白具有相似功能的原则,利用相邻蛋 白质的分子功能注释信息对未知功能的蛋白质进 行注释。具体方法:

①从构建的蛋白互作网络中获得蛋白间相互 作用的信息,并通过网络中 GO 分子功能的注释 获得未知分子功能的蛋白和含有分子功能注释信 息的蛋白,及其相互间的关系。

②计算每个未知分子功能蛋白质的邻近蛋白 数量 (*m*) 和每个 GO 编号注释的蛋白质数量 (*n*)。

③确定具有 GO 注释的相邻节点数是否为1, 如果为1则将所有功能注释给未知分子功能的蛋 白;若不为1,则从任意一个 GO 编号开始计算 *m/n*,若*m/n*≥0.75,则对未知分子功能的蛋白质 进行 GO 注释,若*m/n*<0.75,则进行下一个 GO 编号的计算。

④重复上述步骤,直到网络中 GO 注释蛋白 质的数量保持不变。

1.2 中华绒螯蟹代谢蛋白互作网络的构建

中华绒螯蟹代谢蛋白互作网络数据是以其全 中国水产学会主办 sponsored by China Society of Fisheries 局蛋白互作网络(即总网络)为基础进行构建的, 从总网络中根据 GO"生物过程"注释提取与代谢 功能相关的蛋白及其互作关系构成代谢蛋白互作 网络。

GO 术语之间具有层级结构关系,处于越低 层次的 GO 术语注释的信息越详细,而越高层次 的术语注释则越宽泛,这种结构使研究者能够根 据需求在不同层面上对基因或蛋白功能进行分析。 蛋白互作网络中包含的蛋白注释来自于转录组数 据,均为其能注释到的最详细的注释,即最低层 次的注释,但这样详细的注释并不便于进行大类 功能的研究,例如代谢功能。因此,为了构建代 谢子网络,需要利用层级关系将这些详细的注释 回溯到高层次的代谢功能注释上,并提取相关蛋 白。首选需要确定 GO 术语的层级结构中与代谢 相关的最高层的术语,即为 GO: 0008152 (代谢过 程, metabolic process), 该注释位于层级结构的第 二层。随后,采用 VBA 语言编写程序,通过 GO 回溯的方式,将中华绒螯蟹蛋白互作网络中低层 次 GO 沿层级结构回溯到高层次,并将所有能够 回溯到 GO: 0008152 的 GO 条目提取出来,这些 条目即为与代谢功能相关的 GO 条目。将这些 GO条目与中华绒螯蟹代谢蛋白互作网络中的蛋 白 GO 注释相匹配,能够匹配上的即为代谢相关 蛋白。最后从全局网络中将这些蛋白质及其相互 之间的作用关系提取出来,即形成中华绒螯蟹代 谢蛋白互作网络。

1.3 中华绒螯蟹代谢相关蛋白的分子功能及亚 细胞定位

在中华绒螯蟹蛋白互作网络分子功能注释中, 对网络中的蛋白质所对应的分子功能进行了注释, 将这些注释匹配到代谢蛋白互作网络包含的蛋白 质上,得到这些蛋白的分子功能信息。同理,在 前期研究中的中华绒螯蟹蛋白互作网络细胞组 分注释中,为网络中的蛋白在细胞中的位置进行 了注释,将这些注释匹配到代谢蛋白互作网络包 含的蛋白上,即得到这些蛋白的亚细胞定位信息。

1.4 中华绒螯蟹代谢相关蛋白的途径分析

中华绒螯蟹蛋白互作网络的构建基于其转录 组测序数据,转录组测序结果中包含各单基因簇 (unigene) 对应的 KEGG 途径注释,蛋白互作网络 中包含 unigene 与蛋白的对应关系,根据这些数据, 将转录组中 unigene 的 KEGG 途径注释匹配到蛋 白互作网络中对应的蛋白上,即能够得到每个蛋 白的途径信息。因此,对于代谢蛋白互作网络, 实验首先从总网络中确定代谢蛋白互作网络中每 个蛋白对应的 unigene,并将 unigene 的途径信息 匹配到其编码的蛋白上,即得到了代谢蛋白互作 网络中蛋白的途径信息。

2 结果

2.1 蛋白互作网络分子功能注释

在中华绒螯蟹蛋白互作网络中共有 3 223 个 蛋白,构成了 35 787 条互作关系。该网络中分子 功能未知的蛋白共有 961 个,本实验采用了邻接 节点注释法,按照方法中描述的步骤,为其中 932 个蛋白添加了分子功能注释,占所有未知蛋 白的 97%。注释完成后,网络中在分子功能方面 没有注释的蛋白仅占网络中蛋白总数的 0.9% (29/ 3 223) (图 1),即网络中绝大部分的蛋白是具有分 子功能注释的,中华绒螯蟹蛋白互作网络中仅有 7 个蛋白在细胞过程、分子功能和细胞位置 3 方 面都没有功能注释。

2.2 代谢蛋白互作网络构建

从 GO 数据库下载的数据中共包含 47 347 个 GO 条目,其中 30 811 个条目属于生物过程注释。 在生物过程注释中,10 582 个 GO 条目最终能够 回溯到 GO: 0008152 (代谢过程)的注释上,这些 条目即构成了代谢相关 GO 条目。利用 GO 回溯 的方法,将这些代谢相关 GO 条目。利用 GO 回溯 的方法,将这些代谢相关 GO 条目与在中华绒螯 蟹蛋白互作网络中的蛋白 GO 注释相匹配,在所有 3 223 个蛋白中,确定了与代谢相关的蛋白 2 045 个,其中 1 614 个为原网络中已知分子功能的蛋 白,431 个为新注释分子功能的蛋白。从总网络 中提取这些蛋白参与的互作关系,共得到 15 927 条互作关系,因此最终得到的中华绒螯蟹代谢蛋 白互作网络包含 2 045 个蛋白和 15 927 条互作关 系(图 2)。

2.3 中华绒螯蟹代谢相关蛋白的分子功能及亚 细胞定位

根据本实验添加的分子功能注释和中华绒螯 蟹蛋白互作网络中原有的细胞组分注释,能够得 到网络中蛋白的分子功能和亚细胞定位,将这些 信息匹配到代谢蛋白互作网络中,即能够得到代



图1 中华绒螯蟹蛋白互作网络

蓝色表示原有注释蛋白,粉色表示新注释蛋白,紫色表示未注释到的蛋白。

Fig. 1 The protein-protein interaction network of *E. sinensis*

Blue represents the original annotated protein, pink represents the newly annotated protein, purple represents the protein with no annotation.

谢相关蛋白的分子功能及亚细胞定位。最终,在 中华绒螯蟹代谢蛋白互作网络的所有 2045 个蛋 白中,1966 个蛋白能够确定其分子功能,79 个 蛋白的分子功能无法确定; 共有 1 929 个蛋白能 够确定其亚细胞定位, 116 个蛋白在细胞内的位 置无法确定(图 3)。



图 2 中华绒螯蟹代谢蛋白互作网络

紫色表示新注释分子功能的蛋白,绿色表示原有分子功能注释的蛋白。

Fig. 2 The metabolic protein-protein interaction network of *E. sinensis*

Purple represents the newly annotated protein, blue represents the original annotated protein.

https://www.china-fishery.cn



图 3 中华绒螯蟹代谢蛋白互作网络中蛋白的分子功能和亚细胞定位

橘色表示能够确定亚细胞定位的蛋白,蓝色表示不能确定亚细胞定位的蛋白;三角形表示能够确定分子功能的蛋白,圆形表示未能确定 分子功能的蛋白。



Orange represents protein with subcellular location, blue represents the protein with no subcellular location; triangle represents protein with molecular function, circular represents protein with no molecular function.

代谢蛋白互作网络中各蛋白在不同分子功能 中的分布如图 4 所示,其中具有催化活性 (catalytic activity) 功能和结合活性 (binding) 功能的蛋白 居多,具有催化活性功能的蛋白占网络总蛋白数 的 59.7%,也有少部分代谢蛋白具有酶调节活性 (enzyme regulator activity)。

代谢蛋白互作网络中,各蛋白在不同亚细胞 定位中的分布显示,其中有 94% 的代谢蛋白定位 到细胞部分上 (cell part),并且有大部分代谢蛋白 定位到细胞器 (organelle) 上 (图 5)。

由于亚细胞定位中细胞器的种类繁多,因此, 实验又对网络中的蛋白在各细胞器中的分布进行 了分析,结果显示,大多蛋白分布于有膜细胞器 中,其中包括线粒体、高尔基体等,在无膜细胞 器如核糖体中也有分布(图 6)。

2.4 中华绒螯蟹代谢相关蛋白的途径分析

通过与中华绒螯蟹总蛋白互作网络中 KEGG 途径信息的匹配,得到了代谢相关蛋白的途径信 息。代谢相关蛋白在代谢途径中的分布情况如 图 7 所示。中华绒螯蟹代谢蛋白分布于 40 个子系 统中,其中参与翻译和氨基酸代谢途径的蛋白较 多,也有一部分参与免疫、运输等生物过程。参 与氨基酸代谢的蛋白有 122 个,例如 CAT、MIF、 GOT 和 GPT 等。网络中参与翻译的蛋白有 158 个, 例如 Sec13 等。参与免疫应答的蛋白有 78 个,例 如 HSC70 等。除此之外,在代谢子网络中参与运 输的蛋白有 80 个,如 Crc、Src 和 Sod 等。

3 讨论

本实验以中华绒螯蟹蛋白互作网络为基础, 首先对其中分子功能未知的蛋白进行了分子功能 注释,随后利用 GO 注释信息提取与代谢相关的 蛋白,构建了包含 2 045 个蛋白和 19 527 条相互 作用关系的代谢蛋白互作网络。对该网络中的蛋 白进行分子功能、亚细胞定位和途径分析,在分 子功能分布中,在代谢蛋白互作网络中具有催化 活性功能和结合活性的蛋白居多,也有少部分代 谢蛋白具有酶调节活性。在亚细胞定位分布中, 有 94% 的代谢蛋白定位到了细胞内部,并且大部 分代谢蛋白定位在细胞器上,其中在有膜细胞器

中国水产学会主办 sponsored by China Society of Fisheries



图 4 代谢蛋白在不同分子功能中的分布

1. 结合, 2. 抗氧化, 3. 转运, 4. 核酸结合转录因子, 5. 受体, 6. 电 子载体, 7. 分子传感器, 8. 蛋白质结合转录因子, 9. 结构分子, 10. 营养储备, 11. 翻译调节, 12. 酶调节, 13. 催化, 14. 蛋白质标记物。

Fig. 4 Distribution of metabolic protein in different molecular functions

1. binding, 2. antioxidant activity, 3. transporter activity, 4. nucleic acid binding transcription factor activity, 5. receptor activity, 6. electron carrier activity, 7. molecular transducer activity, 8. protein binding transcription factor activity, 9. structural molecule activity, 10. nutrient reservoir activity, 11. translation regulator activity, 12. enzyme regulator activity, 13. catalytic activity, 14. protein tag.





1. 突触组分, 2. 细胞外基质, 3. 细胞器组分, 4. 细胞组分, 5. 细胞, 6. 细胞膜组分, 7. 细胞核, 8. 膜腔, 9. 突触, 10. 细胞器, 11. 细胞外基质组分, 12. 胞外区组分, 13. 细胞膜, 14. 高分子复合物, 15. 胞外区。

Fig. 5 Distribution of metabolic proteins in cellular components

1. synapse part, 2. extracellular matrix, 3. organelle part, 4. cell part, 5. cell, 6. membrane part, 7. nucleoid, 8. membrane-enclosed lumen, 9. synapse, 10. organelle, 11. extracellular matrix part, 12. extracellular part, 13. membrane, 14. macromolecular complex, 15. extracellular region.



图 6 代谢蛋白在不同细胞器中的分布

1. 收缩纤维组分, 2. 核糖体, 3. 剂量补偿复合体, 4. 细胞器包膜腔, 5. 无膜细胞器, 6. 细胞内细胞器, 7. 内质网膜, 8. 突触小泡, 9. 细胞内细胞器组分, 10. 囊泡, 11. 细胞器腔, 12. 细胞器膜, 13. 线粒体膜组分, 14. 膜结合细胞器。

Fig. 6 Distribution of metabolic proteins in different cellular organelle components

 contractile fiber part, 2. organellar ribosome, 3. dosage compensation complex, 4. organelle envelope lumen, 5. non-membrane-bounded organelle, 6. intracellular organelle, 7. endoplasmic reticulum membrane, 8. synaptic vesicle, 9.intracellular organelle part, 10. vesicle, 11. organelle lumen, 12. organelle membrane, 13. mitochondrial membrane part, 14. membrane-bounded organelle.

中分布最多。对代谢蛋白的途径信息分析,发现 中华绒螯蟹的代谢蛋白分布在40个子系统中,其 中较多分布于翻译和氨基酸代谢,也有一部分具 有免疫、运输等功能。本实验首次构建了中华绒 螯蟹代谢蛋白互作网络,并分析了其中蛋白的分 子功能、亚细胞定位和途径分布,为进一步研究 中华绒螯蟹的代谢活动及代谢有关的蛋白提供了 数据基础。

3.1 中华绒螯蟹蛋白互作网络的分子功能注释

从 GO 数据库下载的 47 347 个 GO 条目中, 30 811 个条目属于生物过程注释,可见生物过程 注释在 3 个注释体系中所占的比重最大。在生物 过程注释中,代谢相关的条目共得到了 10 582 个, 而与其同层次的 GO 条目共有 29 个,可见在整个 生物过程注释体系中,代谢功能具有非常重要的 地位,包含的注释分类约占所有生物过程注释的 三分之一。

在分子功能层面新注释的 932 个蛋白中,有 448 个蛋白具有亚细胞定位信息,其他 484 个蛋 白没有亚细胞定位注释。具有亚细胞定位的蛋白 中分布于细胞部分中的最多,共有 414 个蛋白,

https://www.china-fishery.cn



图 7 代谢蛋白在不同 KEGG 子系统中的分布

1. 氨基酸代谢, 2. 其他次级代谢产物生物合成, 3. 癌症, 4. 碳水化合物代谢, 5. 心血管疾病, 6. 细胞生成与死亡, 7. 细胞运动, 8. 细胞群落, 9. 循环系统, 10. 发育, 11. 消化系统, 12. 内分泌和代谢疾病, 13. 内分泌系统, 14. 能量代谢, 15. 环境适应, 16. 排泄系统, 17. 折叠、分类和降解, 18. 全局系统, 19. 甘氨酸生物合成和代谢, 20. 免疫疾病, 21. 免疫系统, 22. 传染病, 23. 脂质代谢, 24. 膜转运, 25. 辅因子和维生素的代谢, 26. 其他氨基酸代谢, 27. 萜类和聚酮类化合物的代谢, 28. 神经系统, 29. 神经变性疾病, 30. 核苷酸代谢, 31. 纵览系统, 32. 复制和修复, 33. 感觉系统, 34. 信号转导, 35. 信号分子与相互作用, 36. TNF 家族, 37. 转录, 38. 翻译, 39. 运输和分解代谢, 40. 异种生物降解和代谢。

Fig. 7 Distribution of metabolic proteins in KEGG subsystems

1.amino acid metabolism, 2. biosynthesis of other secondary metabolites, 3. cancers, 4. carbohydrate metabolism, 5. cardiovascular diseases, 6. cell growth and death, 7. cell motility, 8. cellular community, 9. circulatory system, 10. development. 11. digestive system, 12. endocrine and metabolic diseases, 13. endocrine system, 14. energy metabolism, 15. environmental adaptation, 16. excretory system, 17. folding, sorting and degradation, 18. global and overview maps, 19. glycan biosynthesis and metabolism. 20. immune diseases, 21. immune system, 22. infectious diseases, 23. lipid metabolism, 24. membrane transport, 25. metabolism of cofactors and vitamins, 26. metabolism of other amino acids, 27. metabolism of terpenoids and polyketides, 28. nervous system, 29. neurodegenerative diseases, 30. nucleotide metabolism, 31. overview, 32. replication and repair, 33. sensory system, 34. signal transduction, 35. signaling molecules and interaction, 36. TNF family, 37. transcription, 38. translation, 39. transport and catabolism, 40. xenobiotics biode-gradation and metabolism.

其中定位于细胞器部分的蛋白共有 339 个。将亚 细胞定位结果更加细化,发现新注释的 932 个功 能蛋白中,大部分蛋白分布于胞内细胞器组分, 其中分布于有膜细胞器的蛋白比分布于无膜细胞 器的蛋白多。将分子功能层面新注释的 932 个蛋 白进行途径分布分析,仅有 155 个蛋白可回溯到 KEGG 途径分布中,其中大部分蛋白分布于折叠、 定位和降解 (folding, sorting and degradation)途径, 也有少部分蛋白分布于转录 (transcription)、翻译 (translation)、信号转导 (signal transduction)以及免 疫系统 (immune system)等途径。

3.2 代谢蛋白互作网络分析

代谢蛋白互作网络中一些蛋白的分子功能、 中国水产学会主办 sponsored by China Society of Fisheries 亚细胞定位和途径分析是与文献报道相符的。例 如在分子功能方面,于丰军^[30]在研究铅和镉2种 重金属对中华绒螯蟹的毒性效应时提到乙酞胆碱 酯酶 (AChE)、谷光甘肽转硫酶 (GST) 和乳酸脱氢 酶 (LDH) 这3种生物生命活动中的重要代谢酶类。 酶作为构成整个生命活动中不可或缺的部分,在 物质进入机体之后,进行催化作用与代谢转化。 这说明中华绒螯蟹中许多代谢蛋白作为催化酶参 与催化调节等代谢作用,与代谢网络蛋白分子功 能分析的结果相符。

经文献查证,中华绒螯蟹中存在 Cullin 家族 蛋白,该家族在进化上十分保守^[31]。Cullin-Ring 类 E3 连接酶 (CRLs) 是以 Cullin 为支架的多亚基 E3 泛素连接酶复合体,主要参与细胞周期、基因 转录、信号转导等生物学进程^[32]。Cullin 作为催化 核心,通过与不同的衔接蛋白结合形成 CRLs,促 进大量的靶蛋白进行泛素化降解。在代谢-蛋白互 作网络中分子功能结果中找到 Cullin 蛋白主要具 有催化活性功能,这与文献中的结果相符合。

根据 Yang 等^[33] 研究,周期蛋白依赖激酶 (cyclin-dependent protein kinases, CDKs) 作为调控 细胞周期的核心分子,主要依赖周期蛋白发挥功 能。CDKs 与细胞周期蛋白亚基结合发挥催化作 用,激活处于静息状态下的细胞,并开启细胞周 期进程。因此,CDKs 具有结合活性功能,与中 华绒螯蟹代谢-蛋白互作网络中分子功能结果一致。

小 G 蛋白 (GTP-binding protein) 是质膜受体 下游信号转导的关键元件,根据其形态以及生理 功能可分为 5 个家族, Ras、Rho、Rab、Arf 和 Ran,其中 Rab 蛋白是最大的一个分支。Rab 蛋白 主要有两种形式,一种是与 GTP 结合的活性形式, 另一种为 GDP 结合的未活化形式。在中华绒螯蟹 代谢-蛋白互作网络分子功能结果中 Rab 蛋白具有 调节活性功能,且根据 Meng 等^[34]的研究,Rab 蛋白参与许多重要生物学过程的调节,包括生长 和分化、细胞分裂、囊泡运输等,Rab6 蛋白有效 调动肌动蛋白,调节宿主细胞的吞噬作用。Rab7 作为发挥内吞作用的关键调节因子,主要调节溶 酶体区室的生物合成以及平衡。

在亚细胞定位方面,在杨洁^[33]研究的中华绒 螯蟹体内几种同工酶中涉及到了ACP,即磷酸单 酯酶。这是一种广泛存在于各类动物体内的,能 够进行磷酸单酯水解,并且在部分营养物质的消 化、吸收及转运的过程中起着重要作用的代谢相 关蛋白。ACP 酶主要定位在动物器官组织上的线 粒体、细胞核周围或溶酶体等细胞器上。而分布 在线粒体内的 ACP 很可能与氧化磷酸化过程有着 密切关联。这与已知的特异性代谢,即代谢蛋白 大多定位在细胞器上,且大部分细胞代谢蛋白存 在于有膜细胞器上是吻合的。

热休克蛋白 60 (heat shock protein 60, HSP60) 是热休克蛋白家族中的一员,主要作为分子伴侣介导多肽的正确折叠和装配,协助蛋白质 的合成和跨膜运输。在 Hong 等^[36]的研究中发现, HSP60作为合适的热应激生物标志物,参与了线 粒体蛋白质的正确折叠,该蛋白定位于线粒体内, 这与实验构建的代谢-蛋白互作网络中的亚细胞定 位结果相吻合。 钙连蛋白 (Cnx)^[37] 是目前研究地最为深入的 特异性糖蛋白的伴侣蛋白,在蛋白质合成中能稳 定多肽链折叠构象,促进新生肽链的折叠以及蛋 白的跨膜运输。Cnx 主要定位于内质网,且大约 有 100 种新生糖蛋白会和 Cnx 在内质网结合。当 蛋白与 Cnx 分离后,会被立即转运到高尔基体。 这与中华绒螯蟹代谢-蛋白互作网络的亚细胞定位 结果一致。

在中华绒螯蟹代谢-蛋白互作网络的亚细胞定 位结果中发现,存在于核糖体中的蛋白大多为核 糖体蛋白。经文献查证,核糖体蛋白 L40 参与核 糖体 60S 大亚基的组装,核糖体蛋白 S27 参与核 糖体 40S 小亚基的形成^[38]。根据 Wang 等^[39]研究 发现,泛素核糖体融合蛋白首先通过剪切释放泛 素单体,使其进入泛素-蛋白酶体途径 (ubiquitinproteasome pathway, UPP),同时核糖体蛋白参与 核糖体的组装,且泛素分子可充当分子伴侣,大 大提高了核糖体蛋白与核糖体的结合效率,对细 胞中蛋白质的合成起着重要作用。

在途径分布方面,参与氨基酸代谢的蛋白有 122个,例如CAT、MIF、GOT和GPT等。目前 关于MIF蛋白的文献仅报道过该蛋白参与先天性 免疫^[40]。朱筛成等^[41]证实GOT和GPT为水生动 物体内2种重要的转氨酶,参与水生生物氨基酸 的分解和合成,GOT和GPT的活性也反映着水生 生物体内氨基酸代谢的强度。在该网络中参与翻 译的蛋白有158个,例如Sec13等。参与免疫应 答的蛋白有78个,例如HSC70等。HSC70为热 休克蛋白,其具有与变性蛋白和热休克因子结合 的能力。Li等^[42]在2016年报道的文献中证实 HSC70蛋白参与中华绒螯蟹的免疫应答。除此之 外,在该代谢子网络中参与运输的蛋白有80个。 例如Crc、Src、Sod等蛋白都参与运输过程。

(作者声明本文无实际或潜在的利益冲突)

参考文献 (References):

- [1] 林存刚, 徐慧, 陈玉成. 酵母双杂交系统研究及其应用 进展[J]. 微生物学杂志, 2005, 25(6): 85-89.
 Lin C G, Xu H, Chen Y C. Advance in research and application of yeast double-hybrid system[J]. Journal of Microbiology, 2005, 25(6): 85-89 (in Chinese).
- [2] 樊晋宇, 崔宗强, 张先恩. 双分子荧光互补技术[J]. 中 国生物化学与分子生物学报, 2008, 24(8): 767-774.
 中国水产学会主办 sponsored by China Society of Fisheries

https://www.china-fishery.cn

Fan J Y, Cui Z Q, Zhang X E. Bimolecular fluorescence complementation technique[J]. Chinese Journal of Biochemistry and Molecular Biology, 2008, 24(8): 767-774 (in Chinese).

- [3] 李珣,任珍珍,钟国华. 噬菌体展示技术在蛋白质研究中的应用[J]. 生命的化学, 2009, 29(4): 588-594.
 Li X, Ren Z Z, Zhong G H. Application of phage display technology in protein researches[J]. Chemistry of Life, 2009, 29(4): 588-594 (in Chinese).
- [4] Zani M L, Moreau T. Phage display as a powerful tool to engineer protease inhibitors[J]. Biochimie, 2010, 92(11): 1689-1704.
- [5] 冯晓琴,徐峰,王嵬,等.通过对免疫共沉淀技术的优 化验证蛋白质弱相互作用[J].中国科学院研究生院学 报,2013,30(1):18-23.

Feng X Q, Xu F, Wang W, *et al.* Confirmation of weak protein-protein interactions by optimizing co-immuno-precipitation[J]. Journal of Graduate University of Chinese Academy of Sciences, 2013, 30(1): 18-23 (in Chinese).

- [6] Kuo M H, Allis C D. *In vivo* cross-linking and immunoprecipitation for studying dynamic protein: DNA associations in a chromatin environment[J]. Methods, 1999, 19(3): 425-433.
- [7] 李玲,杨鹏跃,朱本忠,等.染色质免疫共沉淀技术的应用和研究进展[J].中国食品学报,2012,12(6):124-132.

Li L, Yang P Y, Zhu B Z, *et al.* The application and research progress of chromatin immunoprecipitation[J]. Journal of Chinese Institute of Food Science and Technology, 2012, 12(6): 124-132 (in Chinese).

[8] 孙平,张逢春,张影.蛋白质芯片技术的研究及应用现 状[J].北华大学学报(自然科学版),2009,10(2):115-120.

> Sun P, Zhang F C, Zhang Y. Protein microarray technology and application status[J]. Journal of Beihua University (Natural Science), 2009, 10(2): 115-120 (in Chinese).

[9] 刘康栋 赵建龙. 蛋白质芯片技术进展[J]. 中国生物工
 程杂志, 2004, 24(12): 48-52,58.
 Liu K D, Zhao J L. Progress of protein chip

technology[J]. China Biotechnology, 2004, 24(12): 48-52,58 (in Chinese).

- [10] Raman K. Construction and analysis of protein-protein interaction networks[J]. Automated Experimentation, 2010, 2(1): 2.
- [11] Ho Y, Gruhler A, Heilbut A, *et al.* Systematic identification of protein complexes in *Saccharomyces cerevisiae*

by mass spectrometry[J]. Nature, 2002, 415(6868): 180-183.

- [12] Ito T, Chiba T, Ozawa R, et al. A comprehensive twohybrid analysis to explore the yeast protein interactome[J]. Proceedings of the National Academy of Sciences of the United States of America, 2001, 98(8): 4569-4574.
- [13] Giot L, Bader J S, Brouwer C, et al. A protein interaction map of *Drosophila melanogaster*[J]. Science, 2003, 302(5651): 1727-1736.
- [14] Li S M, Armstrong C M, Bertin N, et al. A map of the interactome network of the metazoan C. elegans[J]. Science, 2004, 303(5657): 540-543.
- [15] Arifuzzaman M, Maeda M, Itoh A, et al. Large-scale identification of protein-protein interaction of Escherichia coli K-12[J]. Genome Research, 2006, 16(5): 686-691.
- [16] Rual J F, Venkatesan K, Hao T, *et al.* Towards a proteome-scale map of the human protein-protein interaction network[J]. Nature, 2005, 437(7062): 1173-1178.
- [17] De Bodt S, Proost S, Vandepoele K, et al. Predicting protein-protein interactions in Arabidopsis thaliana through integration of orthology, gene ontology and coexpression[J]. BMC Genomics, 2009, 10(1): 288.
- [18] Schuette S, Piatkowski B, Corley A, et al. Predicted protein-protein interactions in the moss *Physcomitrella patens*: a new bioinformatic resource[J]. BMC Bioinformatics, 2015, 16(1): 89.
- [19] Kim H B, Kim K K. Protein interaction network related to *Helicobacter pylori* infection response[J]. World Journal of Gastroenterology, 2009, 15(36): 4518-4528.
- [20] Lin M Z, Zhou X, Shen X L, *et al.* The predicted Arabidopsis interactome resource and network topologybased systems biology analyses[J]. The Plant cell, 2011, 23(3): 911-922.
- [21] Stelzl U, Worm U, Lalowski M, et al. A human proteinprotein interaction network: a resource for annotating the proteome[J]. Cell, 2005, 122(6): 957-968.
- [22] Hu Y, Yang Y C, Fang Z H, *et al.* Detecting pathway relationship in the context of human protein-protein interaction network and its application to Parkinson's disease[J]. Methods, 2017, 131: 93-103.
- [23] Li L, Tibiche C, Fu C, et al. The human phosphotyrosine signaling network: evolution and hotspots of hijacking in cancer[J]. Genome Research, 2012, 22(7): 1222-1230.
- [24] Blasche S, Wuchty S, Rajagopala S V, *et al.* The protein interaction network of bacteriophage lambda with its

host, *Escherichia coli*[J]. Journal of Virology, 2013, 87(23): 12745-12755.

- [25] Zou E M, Du N S, Lai W. The effects of mass temperature and thermal acclimation on the respiration rate of the Chinese freshwater crab *Eriocheir sinensis* (Crustacea: Decapod)[J]. Zoological Research, 1995, 16(1): 49-58.
- [26] Zou E M, Du N S, Lai W. The effects of acute progressive hypoxia on the respiration rate of the chinese crab *Eriocheir sinensis*[J]. Zoological Research, 1993, 14(4): 327-334.
- [27] 陈立侨,李二超. 中华绒螯蟹营养需求的研究现状和 进展[J]. 饲料工业, 2009, 30(10): 1-6.
 Chen L Q, Li E C. Current situation and development on nutrition requirement of *Eriocheir sinensis*[J]. Feed Industry, 2009, 30(10): 1-6 (in Chinese).
- [28] Hao T, Zeng Z, Wang B, et al. The protein-protein interaction network of eyestalk, Y-organ and hepatopancreas in Chinese mitten crab Eriocheir sinensis[J]. BMC Systems Biology, 2014, 8(1): 39.
- [29] Wang B, Ning Q J, Wang Q, et al. Reconstruction and subcellular localization analysis of *Eriocheir sinensis* molting protein-protein interaction network[J]. Pakistan Journal of Zoology, 2018, 50(5): 1777-1784.
- [30] 于丰军. 铅和镉两种重金属对中华绒螯蟹的毒性效应研究 [D]. 上海: 华东师范大学, 2005.
 Yu F J. The study of Cadmium and lead toxicity on the Chinese mitten-handed crab, *Eriocheir sinensis*[D].
 Shanghai: East China Normal University, 2005 (in Chinese).
- [31] Wang Y L, Li Q, Xie J, *et al.* Involvement of the single *Cul*4 gene of Chinese mitten crab *Eriocheir sinensis* in spermatogenesis[J]. Gene, 2014, 536(1): 9-17.
- [32] Petroski M D, Deshaies R J. Function and regulation of cullin-RING ubiquitin ligases[J]. Nature Reviews Molecular Cell Biology, 2005, 6(1): 9-20.
- [33] Yang B, Jia Y K, Jia Z H, *et al.* The cyclin-dependent kinase 2 (CDK2) mediates hematopoiesis through G1-to-S transition in Chinese mitten crab *Eriocheir sinensis*[J]. Developmental & Comparative Immunology, 2018, 81: 156-166.
- [34] Meng Q G, Hou L B, Zhao Y, et al. Itraq-based proteomic study of the effects of *Spiroplasma eriocheiris* on Chinese mitten crab *Eriocheir sinensis* hemocytes[J]. Fish & Shellfish Immunology, 2014, 40(1): 182-189.

- [35] 杨洁. 中华绒螯蟹体内几种同工酶及在不同发育阶段的研究 [D]. 上海: 上海师范大学, 2010.
 Yang J. Studies on enzymes physicochemical property and the comparative in different phases in *Eriocheir sinensis*[D]. Shanghai: Shanghai Normal University, 2010 (in Chinese).
- [36] Hong Y H, Huang Y, Yan G W, et al. Effects of deltamethrin on the antioxidant defense and heat shock protein expression in Chinese mitten crab, *Eriocheir sinen*sis[J]. Environmental Toxicology and Pharmacology, 2019, 66: 1-6.
- [37] Huang Y, Hui K M, Jin M, *et al.* Two endoplasmic reticulum proteins (calnexin and caleticulin) are involved in innate immunity in Chinese mitten crab (*Eriocheir sinensis*)[J]. Scientific reports, 2016, 6(27578): 1-11.
- [38] Finley D, Bartel B, Varshavsky A. The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis[J]. Nature, 1989, 338(6214): 394-401.
- [39] Wang Q, Chen L L, Wang Y, et al. Expression characteristics of two ubiquitin/ribosomal fusion protein genes in the developing testis, accessory gonad and ovary of Chinese mitten crab, *Eriocheir sinensis*[J]. Molecular Biology Reports, 2012, 39(6): 6683-6692.
- [40] Li W W, Jin X K, He L, et al. Molecular cloning, characterization and expression analysis of macrophage migration inhibitory protein (MIF) in Chinese mitten crab, *Eriocheir sinensis*[J]. Fish & Shellfish Immunology, 2011, 30(1): 324-329.
- [41] 朱筛成, 龙晓文, 向朝林, 等. 复合蛋白源替代鱼粉对 中华绒螯蟹幼蟹生长性能、生理代谢和生化组成的 影响[J]. 南方水产科学, 2019, 15(2): 83-92.
 Zhu S C, Long X W, Xiang C L, *et al.* Effects of dietary fishmeal replacement with protein mixtures on growth performance, physiological metabolism and biochemical composition of juvenile Chinese mitten crab (*Eriocheir sinensis*)[J]. South China Fisheries Science, 2019, 15(2): 83-92 (in Chinese).
- [42] Li P, Jiang X F, Guo W B, et al. Expression patterns of two heat-shock cognate 70 genes during immune responses and larval development of the Chinese mitten crab Eriocheir sinensis[J]. Genetics & Molecular Research, 2016, 15(3): gmr.15036319.

Construction of metabolic protein-protein interaction network of *Eriocheir sinensis* and molecular function, subcellular location and pathway analysis

YANG Jiarui, HAO Tong^{*}, LI Qianyi, SUN Jinsheng^{*}

(Tianjin Key Laboratory of Animal and Plant Resistance, College of Life Sciences, Tianjin Normal University, Tianjin 300387, China)

Abstract: Various diseases caused by metabolic abnormalities have caused great losses to the breeding industry of Eriocheir sinensis. Therefore, the study of the metabolic system of E. sinensis is very important for its healthy growth and development. However, there is still a lack of systematic analysis and research on the metabolic process of E. sinensis., and this study aims to construct a systematic tool for the study on metabolic process of E. sinensis. Firstly, the proteins with unknown molecular functions were annotated based on the protein-protein interaction network (PIN) by annotating adjacent nodes. Secondly, based on the global network, the metabolic PIN of E. sinensis was constructed by extracting metabolism related proteins and their interaction relationships from the global network according to the GO annotation. The GO terms were organized in a hierarchical structure, among which, the highest level of metabolic system is annotated as GO: 0008152 (metabolic process). All the proteins that can be back traced to GO: 0008152 through the hierarchical structure were considered as metabolism related proteins. Then, the metabolism related proteins and their interaction relationships were extracted from the global network, which formed the metabolic PIN. Subsequently, the proteins in the network were annotated according to their molecular function and cellular component. Furthermore, the pathway distribution of the proteins in the metabolic PIN was analyzed according to the KEGG annotation of unigenes. The molecular functions of 932 proteins were determined, accounting for 97% of all unknown functional proteins in the network. The extracted metabolic PIN contains 2 045 proteins and 15 927 interactions. According to the molecular function and cellular component annotations, the functions and subcellular locations of the proteins in the network were further analyzed. 94.2% (1 926/2 045) of the proteins in the network had subcellular location information, and most of them were distributed in membranous organelles. 96.1% of the proteins (1 966/2 045) have molecular function information, and most of them have catalytic and binding activities. Analysis on the distribution of proteins in 40 subsystems showed that more proteins were involved in translation and amino acid metabolism, and some of them were involved in immune and transport processes. In this study, the metabolic PIN of E. sinensis was constructed for the first time, and the molecular function, subcellular location and pathway of the proteins were analyzed, which provided data basis for further study of metabolic activities and metabolism related proteins in E. sinensis. The results of this study can provide important data for the study of the function and location of metabolic proteins in E. sinensis.

Key words: *Eriocheir sinensis*; metabolic protein-protein interaction network; molecular function; subcellular localization; metabolism

Corresponding authors: HAO Tong. E-mail: skyht@tjnu.edu.cn;

SUN Jinsheng. E-mail: skysjs@tjnu.edu.cn

Funding projects: National Key Research and Development Program (2018YFD0901301); National Natural Science Foundation of China (31770904); Tianjin Development Program for Innovation and Entrepreneurship Team (ITTFRS2017007); Program for Innovative Research Team in University of Tianjin (TD13-5076); Natural Science Foundation of Tianjin (19JCYBJC29700)

中国水产学会主办 sponsored by China Society of Fisheries