

综述与专论

血浆蛋白质组
——人类蛋白质组计划的“探路者”*

王 英 赵晓航**

(中国医学科学院 肿瘤医院肿瘤研究所, 分子肿瘤学国家重点实验室, 北京 100021)
中国协和医科大学

摘要 概述了血浆蛋白的研究现状、难点和策略。血浆是血液中无形的液体成分, 是一种十分复杂和多样化的基质, 包含数百万种蛋白质和小分子多肽、盐、类脂、氨基酸和糖等。血浆蛋白参与机体免疫、凝血-抗凝血、物质运输、营养和对生长信号调节等多种重要的生理功能。人体器官的病理变化可导致血浆蛋白在结构和数量上的改变, 这种特征性的变化对疾病诊断和疗效监测具有十分重要的意义。然而, 迄今为止人类对血浆蛋白的了解还十分有限, 只有很少一部分血浆蛋白被用于常规的临床诊断。全面而系统地认识健康和疾病状态下血液循环中血浆蛋白的性质, 会极大地加速对具有疾病诊断和治疗监测作用的血浆标志蛋白的研发。国际人类蛋白质组组织于 2002 年首先选择了血浆蛋白质组作为人类蛋白质组首期执行计划之一, 其初期目标是: a. 比较各种蛋白质组分析技术平台的优点和局限性; b. 用这些技术平台分析人类血浆和血清的参考样本; c. 建立人类血浆蛋白质组知识库。

关键词 血浆/血清蛋白, 蛋白质组学, 参考样本, 生物标志, 质谱

学科分类号 Q78

1 血浆蛋白的定义和分类

血浆蛋白的组成十分复杂, 以往把那些在血浆中出现, 并在血浆中发挥作用的蛋白质称为血浆蛋白。随着蛋白质组学研究的发展, 目前把血浆中的所有蛋白质统称为血浆蛋白。根据其来源和功能可分为: a. 组成性蛋白, 主要指肝脏和小肠的分泌蛋白, 它们在血浆中发挥作用, 如白蛋白、纤维蛋白原等; b. 免疫球蛋白, 血浆中大约含有数百万种抗体; c. “长距离”受体和配体, 如胰岛素等生长因子; d. “局部”受体和配体, 如 IL-8 等蛋白质因子; e. 一过性通过蛋白, 如溶酶体蛋白酶; f. 组织渗漏蛋白, 如心肌球蛋白; g. 异常分泌蛋白, 如 PSA 等肿瘤相关蛋白; h. 外来蛋白, 如病毒蛋白等。其中, 前两类是血浆蛋白的主要组成成分, 属于高丰度、大分子质量和易于检测的蛋白质。目前被分离、鉴定并用于临床诊断的血浆蛋白多属于此类。后几类血浆蛋白多为低丰度蛋白, 它们种类繁多、性质各异、作用重要、分离和鉴定相当困难。比如, 通常在细胞内发挥作用, 但当细胞死亡或损伤后被释放到血浆中的组织渗漏蛋白, 肿瘤或其他病变器官释放到血浆中的异常分泌蛋白等, 它们对认识疾病演变过程, 分离疾病特征标志

和药物靶标具有重要意义^[1]。

2 人类已知的水浆蛋白

目前还不知道血浆蛋白的确切种类和数量。保守地估计组成性血浆蛋白约有 500 种, 它们多为糖蛋白, 平均每种蛋白质至少具有 5 种不同的结构, 如前体分子 (precursors)、成熟形式 (mature forms)、降解产物 (degradation products) 以及不同的剪接体 (splice variants) 等, 预计由 5 万种不同的蛋白质分子组成; 另一部分是组织渗漏蛋白, 可能包括全部人类蛋白质组。如果按 5 万种基因产物计算, 平均每种产生 10 种不同的结构, 就会有 50 万种蛋白质分子; 血浆中出现序列不同免疫球蛋白的可能性有近千万种, 但是它们不会同时在血浆中全部出现。Pieper、Tirumalai 和 Adkins 等^[2-4]采用不同的技术策略, 分别从血浆中鉴定出 325、341 和 490 种不同的蛋白质。Anderson 等^[1,5]总结了截至 2004 年初所鉴定的血浆蛋白为 1 175 种,

* 国家自然科学基金资助项目 (30370713, 30225045 和 30171049) 和国家高技术“863”计划资助项目 (2004AA227060)。

** 通讯联系人。

Tel: 010-67709015, E-mail: zhaoxh@pubem.cicams.ac.cn

收稿日期: 2004-02-09, 接受日期: 2004-03-29

其中 117 种已被美国食品药品监督管理局 (FDA) 批准用于临床检查, 它们仅占整个血浆蛋白的很小部分。

3 人类血浆蛋白质组计划简介

随着人类基因组计划的完成, 探究基因组所蕴含的生物学意义, 阐明基因组的生物学功能, 已经成为生命科学的研究热点。

2001 年 10 月, 国际人类蛋白质组组织 (Human Proteome Organization, HUPO) 在美国成立, 并计划启动人类蛋白质组计划 (Human Proteome Project, HPP)。HPP 的研究目的是鉴定人类基因组编码的全部蛋白质及其功能, 揭示: a. 构成各种人类组织不同细胞类型的蛋白质表达谱; b. 蛋白质组翻译后修饰谱; c. 蛋白质组亚细胞定位图; d. 蛋白质-蛋白质相互作用关系图; e. 蛋白质结构与功能联系图等。包括对不同发育阶段、不同生理和病理状态下蛋白质的空间与时间上表达状况的认识。HPP 是继人类基因组计划后又一个国际化的重大科学研究项目, 其首批执行计划包括: 人血浆蛋白质组计划 (Human Plasma Proteome Project, HUPO PPP) 和人肝脏蛋白质组计划 (Human Liver Proteome Project, HLPP)^[6-8]。

由于技术手段和数据处理上的复杂性和多样性, 使得蛋白质组研究比基因组研究困难和复杂得多。作为人类蛋白质组研究的“探路者”, HUPO PPP 的先期 (pilot phase) 目标是: 比较不同样本收集、处理和储存过程对蛋白质分析的影响; 通过分析血浆蛋白质组, 比较不同蛋白质组分析技术平台的敏感性和局限性; 比较不同高丰度蛋白去除 (depletion) 方法, 以及去除与否的差异; 数据提取、整理和储存的标准化等, 全面认识人类正常血浆/血清中的全部表达蛋白及其在不同生理条件下的变异度^[8]。

HUPO PPP 由美国科学家 Gilbert Omenn 牵头, 目前有来自世界十多个国家和地区的 57 个实验室参加。该计划启动阶段要求所有参加实验室均使用相同的“参考样品” (reference specimens), 该参考样本分别由英国、美国和中国按照同一标准提供, 制备成不同人群的混合血浆 (EDTA、肝素和枸橼酸钠抗凝血) 和血清。代表世界上主要人种的参考样本取自部分志愿者, 他们分别属于白种美国人 (Caucasian-American)、非洲裔美国人 (African-American)、亚裔美国人 (Asian-American) 和亚裔

中国人 (Asian-Chinese)。在此基础上, 使用各实验室不同的技术平台、不同的蛋白质组学实验方案, 全面鉴定和分析正常人血浆/血清蛋白质组。

选择血浆作为人类蛋白质组研究的“探路者”主要基于以下考虑: a. 血浆蛋白质组是最复杂的人类蛋白质组, 其中包含着不同组织的亚蛋白质组; b. 血浆蛋白取材方便, 能够获得足够的研究样本, 容易标准化; c. 血浆蛋白的性质具有很大的动态范围; d. 血浆/血清是临床检查最主要的样本来源, 对疾病诊断和疗效监测具有重要的意义; e. 多数血浆蛋白的表达水平与其相应 mRNA 的水平相关性很差, 且普遍存在糖基化、磷酸化等翻译后修饰现象, 只能从蛋白质水平上开展研究。

4 血浆蛋白组研究遇到的问题

4.1 取材方式

血清是指血液凝固后析出的淡黄色透明液体, 与血浆的区别在于血清中没有纤维蛋白原, 但含有一些在凝血过程中生成的分解产物。常用的血浆有三种: 肝素、EDTA 和枸橼酸钠抗凝血浆。血浆蛋白的取材方式不同, 会在一定程度上影响血浆蛋白质组的分析结果。选用血清还是血浆? 选用何种抗凝剂获得血浆? 是血浆蛋白质组研究首先遇到的问题。早在 20 世纪 80 年代, Smith 等^[9]发现使用抗凝剂能引起血浆渗透压改变, 使血细胞内部的液体进入血浆中, 导致血浆蛋白被稀释, 其中枸橼酸盐抗凝剂会引起非常明显的改变, 而肝素则影响不大。Sadagopan 等^[10]发现在使用液相色谱/串联质谱 (LC/MS/MS) 分析血浆样品时, EDTA 抗凝血浆的自动进样错误率小于肝素抗凝血浆, 与此同时进样速度可以提高 3 倍。Glendenning 等^[11]发现甲状旁腺激素在 EDTA 抗凝血浆中的稳定性要好于血清。此外, EDTA 对血浆中的金属离子结合蛋白、胆固醇、甘油三脂等都会产生影响^[12]。

由于血浆蛋白代谢范围很大, 个体差异、遗传因素和某些偶然因素都可能影响血浆蛋白组成。如何使研究结果具有代表性, 避免个体差异带来的偏差? HUPO PPP 采用先选择有代表性的个体取材, 再制备混合参考样本以稀释个体差异的取材方式。

4.2 蛋白质性质的动态范围

血浆蛋白的性质具有很大的动态范围 (dynamic range)。首先, 不同血浆蛋白的表达丰度相差很大, 同一种血浆蛋白在不同生理和病理条件下浓度变化也很大。血浆中的组成性蛋白, 如白蛋

白含量在 g/L 级以上, 而多数组织渗漏蛋白的浓度在 ng/L 级以下. 通常把含量在 g/L 级的血浆蛋白称为高丰度蛋白 (high abundant proteins), 浓度在 $\mu\text{g/L}$ 级的血浆蛋白称为低丰度蛋白 (low abundant

proteins). 研究表明, 血浆蛋白浓度的动态范围至少相差 9 ~ 12 个数量级. 图 1 显示了 70 种常见血浆中蛋白质的浓度范围^[1,2].

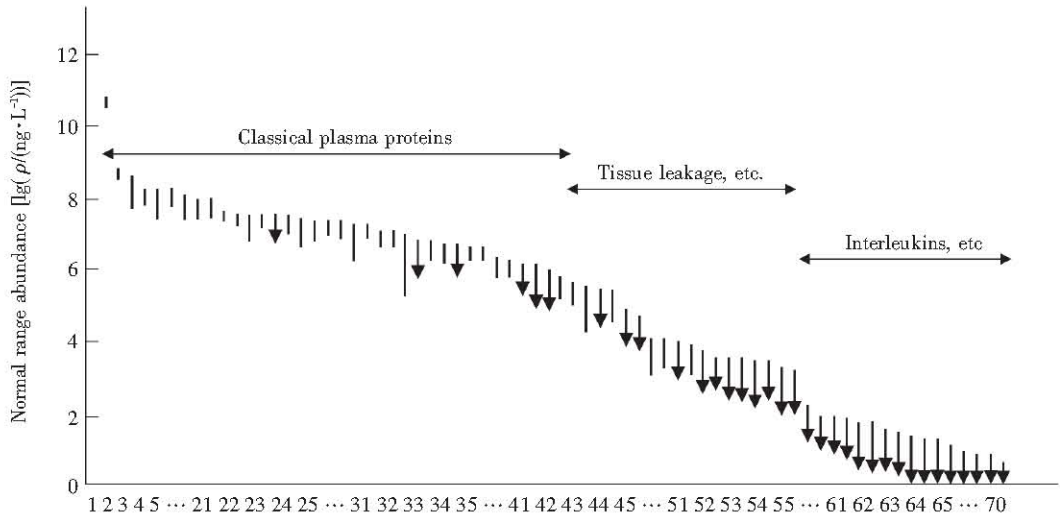


Fig. 1 Reference intervals for 70 proteins in plasma (From Specialty Laboratories)

图 1 70 种血浆蛋白的浓度参考范围 (引自《Specialty Laboratories》)

1: Hemoglobin, 2: Albumin, 3: Factor H, 4: C9 complement, 5: C8 complement, 6: C5 complement, 7: IgG total, 8: C6 complement, 9: C7 complement, 10: Complement, 11: Transferrin, 12: Fibrinogen, 13: IgA total, 14: Thyroxin binding, 15: C3b complemen, 16: α -2-macroglobulin, 17: IgM total, 18: C2 complement, 19: α -1-antitrypsin, 20: C3 complement, 21: Haptoglobin, 22: Apolipoprotein A-1, 23: Apolipoprotein B, 24: α -1-acid glycoprotein, 25: Lipoprotein (a), 26: Thrombus precursor protein, 27: Ceruloplasmin, 28: C4 Complement, 29: CRP, 30: Complement, 31: Prealbumin, 32: C1q complement, 33: Plasminogen, 34: Bb fragment, 35: IgD, 36: C1 inhibitor, 37: Retinol binding protein, 38: C3a complement, 39: Ferritin, 40: Rantes, 41: SC5b-9 complex, 42: Myoglobin, 43: Thyroglobulin, 44: TPA, 45: C5a complement, 46: Neuron-specific enolase, 47: C-peptide, 48: AFP, 49: TNF binding protein, 50: PSA, 51: Prostatic acid phosphatase, 52: Myelin basic protein, 53: Carcinoembryonic, 54: Troponin I, 55: IL-1 ra, 56: MIP-1 β , 57: Troponin T, 58: IL-8, 59: MIP-1 α , 60: Tissue factor, 61: GCSF, 62: Interferon α , 63: IL-2, 64: IL-4, 65: TNF α , 66: Interferon γ , 67: IL-1 β , 68: IL-12, 69: IL-5, 70: IL-10.

此外, 各种血浆蛋白的分子质量、等电点差异也非常大, Anderson 等^[1]总结了 289 种血浆蛋白分子质量的分布范围 (图 2). 虽然大多数血浆蛋白分子质量集中在 10 ~ 100 ku 之间, 但在 <10 ku 和 >100 ku 范围内仍然存在相当数量的血浆蛋白, 尤其是 <10 ku 的小分子蛋白/肽多为疾病时释放到血浆中的特异蛋白, 对分离和鉴定疾病标志蛋白具有重要意义. 同时, 血浆蛋白具有较大异质性 (heterogeneity), 绝大多数血浆蛋白是糖蛋白, 还包括其他多种翻译后修饰和剪切形式.

4.3 高丰度蛋白对分析的干扰

血浆蛋白浓度的动态范围很大, 高丰度蛋白在含量上占有绝对优势. 血浆中前 10 种高丰度蛋白分别为: 白蛋白、免疫球蛋白、转铁蛋白、纤维蛋白原、 α -2 巨球蛋白、 α -1 抗胰蛋白酶、补体 C3 片

段、结合珠蛋白等, 其含量占血浆总蛋白的 90% 以上. 数以万计的低丰度蛋白在血浆中的总含量低于 10%, 而它们与许多疾病相关. 由于蛋白质组分析技术对蛋白质上样量的限制, 高丰度蛋白通常给低丰度蛋白的分离、鉴定带来极大的困难, 成为蛋白质组分析的瓶颈问题. 在进行血浆蛋白质组分析时有必要去除高丰度蛋白的干扰. 去除血浆高丰度蛋白的基本原理是: 用一种对白蛋白具有吸附能力的蓝胶 (blue affinity gel) 非特异性地去除血浆白蛋白; 用 protein A/G 与血浆免疫球蛋白 FC 端结合而去除免疫球蛋白; 用针对主要血浆高丰度蛋白的单克隆抗体亲和层析后, 同时除去几种高丰度蛋白等^[13,14]. 另一方面, 在去除高丰度蛋白的同时有可能带走很多低丰度蛋白, 而丢失部分蛋白质信息. 例如, 占血浆总蛋白一半以上的白蛋白

(55%) 是血浆中重要的转运蛋白, 去除白蛋白的同时很可能会带走与白蛋白结合的细胞因子、肽类

激素以及脂蛋白等低丰度蛋白^[4].

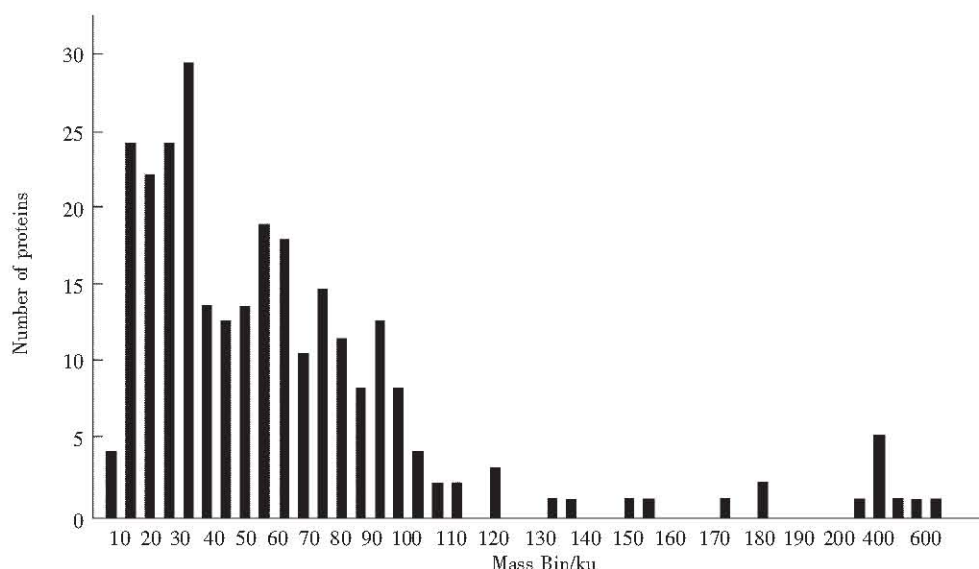


Fig. 2 Histogram of the molecular masses of plasma proteins

图2 血浆蛋白的分子质量分布情况

4.4 血浆蛋白的不稳定性

血浆蛋白在离体后很容易降解, 许多因素会影响血浆蛋白的稳定性。例如, 保存温度和时间、反复冻融、使用蛋白酶和磷酸酶抑制剂, 以及保存容器/试管的质地等等, 目前尚无统一方法。由于样本取材、保存和处理方法不一, 很可能导致来自不同实验室结果间的差异, 降低可比性。为避免这种差异, HUPO PPP 要求所有项目参加实验室使用经同一方法处理的“参考样本”。

4.5 数据间的可比性差

截止到2003年10月的研究发现, 来自不同实验室间的结果可比性较差。由于各个实验室所用的技术平台不同, 相应数据提取、处理的软件不一都会降低数据间可比性。建立一种统一的数据提取、处理和提交标准十分必要。比如, 有人建议在提交质谱数据时将所有数据转换成目前国际上流行的可扩展标记语言 (eXtensible Markup Language, XML) 格式, 以便于各个实验室间远程访问、读取和添加数据等。目前, HUPO PPP 正在建立一个中央操作系统 (central operation) 用统一标准提取和整理数据。

5 血浆蛋白质组研究的主要技术策略

由于血浆蛋白组成复杂, 蛋白质组学研究技术

各有利弊, 靠单一技术平台不能达到最佳分离效果。联合使用不同技术平台 (主要包括: 去除高丰度蛋白、样本预分离系统、分离系统、质谱鉴定系统), 并加以适当改进是目前通常的做法^[15,16]。几种主要的研究策略如下。

a. 高丰度蛋白去除 + 色谱分离 + 2-DE + 联合质谱鉴定: Pieper 等^[2]采集两名健康男性志愿者血清, 首先用免疫亲和吸附法去除血清8种主要的高丰度蛋白; 然后经色谱预分离 (pre-fractionation), 包括离子交换色谱 (anion-exchange chromatography, AEC) 和分子筛色谱 (size-exclusion chromatography, SEC); 将富集的低丰度蛋白组分 (fractions) 经二维电泳 (two-dimensional gel electrophoresis, 2-DE) 分离, 得到近3700个蛋白质点; 这些蛋白质点首先经MALDI-TOF质谱鉴定, 再将未能鉴定的蛋白质点进一步用LC-MS/MS串联质谱分析。最终, 获得了325个完整的蛋白质。

b. 鸟枪测序法 (Shotgun sequencing), 即高丰度蛋白去除 + 全蛋白酶解 + 二维液相色谱 (阳离子/阴离子交换色谱 + 反相色谱) 分离 + 质谱鉴定 (离子阱或Q-TOF串联质谱): Adkins 等^[4]采集女性志愿者血清, 首先用protein A/G吸附去除免疫球蛋白, 胰蛋白酶对血浆全蛋白酶解; 酶解后经强阳离子交换 (strong cation exchange, SCX) 和反相

(reversed phase) 色谱分离后, 进一步经电喷雾液相离子阱质谱 (LCQ Ion Trap) 鉴定, 获得 490 种不同血浆蛋白, 并把蛋白质鉴定效率提高了 3~5 倍。

c. 血浆全蛋白预分离 (色谱聚集 + 反相高压液相色谱) + 组分胰蛋白酶酶解 + 串联质谱鉴定 (LC-ESI-MS-MS 或 2D-micro-LC + MALDI-TOF-TOF-MS): Beckman 公司于 2003 年下旬推出 ProteomeLab PF 2D system 完整血浆蛋白预分离系统。血浆蛋白先经多维色谱分离, 各分离组分经胰蛋白酶酶解后直接进行质谱鉴定。克服了 2-DE 的缺点, 简化了信息处理上的复杂性。

d. 多重窄 pH 范围 2-DE + 联合质谱鉴定 (MALDI-TOF + LC-MS/MS): Starita-Geribaldi 等^[17]通过使用相差 1 个 pH 范围的多重固相 pH 梯度干胶条, 大大提高了传统 2-DE 的分离效率, 经 MALDI-TOF + LC-MS/MS 联合质谱分析后使蛋白质鉴定的效率增长两倍, 并且改善了极酸和极碱性蛋白质的分离和鉴定效果。

e. 非变性等电聚焦预分离 + 变性等电聚焦预分离 + 1D 分离 + 质谱鉴定 (LC/MS/MS): Giometti 等^[18]通过实验证明在非变性条件下 (不破坏蛋白质的三级结构) 进行二维电泳, 分离到的蛋白质仍然保持原始状态, 使蛋白质生物功能和蛋白质之间相互作用的检测成为可能, 克服了传统 2-DE 的缺陷 (采用变性条件)。Manabe 等^[19]通过采用非变性二维电泳与变性二维电泳相结合的方法, 成功地从人血浆中分离、鉴定出多个 IgG 结合蛋白, 这一方法将成为研究蛋白质功能的一个有效手段。

f. SELDI-TOF 分离 + Q-TOF 质谱鉴定: 表面增强激光解吸/离子化-飞行时间质谱 (surface enhanced laser desorption/ionization-time of flight, SELDI-TOF) 是一种以固相表面亲和为基础的质谱分离技术。完整血浆蛋白先经过与不同亲和表面 (化学表面芯片和生物表面芯片) 结合纯化后, 再经过 SELDI 质谱分离。由于结合蛋白是未经剪切的, 很多蛋白质的质荷比非常接近, 因此该技术无法鉴定所分离的蛋白质, 需要进一步结合 Q-TOF 质谱鉴定^[20]。由于该技术能够相对简单、快速地鉴别出不同分析样本的差异蛋白质, 尤其是小分子蛋白, 因此在血浆蛋白的分离和鉴定中与上述其他技术具有很大的互补性, 已被用于疾病相关的血浆蛋白质组研究中。Petricoin 等^[21]用疏水型芯片结合并分离了卵巢癌患者血浆相关蛋白, 在肿瘤和健

康血清之间发现了 5 个差异明显的蛋白质峰, 它们共同构成卵巢癌诊断模型, 用该模型对肿瘤和健康人血浆进行盲筛, 该模型诊断肿瘤的敏感性为 100%, 特异性为 95%。Li 等^[22]选用金属离子螯合芯片结合并分离乳腺癌血浆蛋白, 发现 3 个乳腺癌特异的差异蛋白峰。用这 3 个特异蛋白标志盲筛一组乳腺癌和健康人血浆蛋白, 乳腺癌检出的敏感性为 93%, 特异性为 91%。Adam 等^[23]利用同样的技术策略找到了 9 个区别前列腺癌和良性前列腺增生与健康志愿者血浆的差异蛋白峰, 其敏感性为 83%, 特异性为 97%。此外, SELDI-TOF 技术也在结肠癌、肺癌、食管癌、肝癌等其他恶性肿瘤生物标志物的研究方面取得了较好的结果。

6 展 望

虽然蛋白质组技术有了长足的进步, 新的、高通量和自动化的分离、鉴定技术平台相继出现, 但人类蛋白质组研究尚处在“探路”阶段, 其复杂性也许超乎人们的想象。蛋白质组研究中的许多技术“瓶颈”问题还没有解决。血浆蛋白质组研究将会成为新的研究热点, 完全解析血浆蛋白质组, 阐明每个蛋白质的功能, 描绘出蛋白质间相互作用的网络, 需要全世界科学家继续努力, 任重而道远。对血浆蛋白质组的深入研究, 将为无数疾病患者带来诊治的希望, 成为医学发展史上的又一个里程碑。

参 考 文 献

- 1 Anderson N L, Anderson N G. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics*, 2002, **1** (11): 845~867
- 2 Pieper R, Gatlin C L, Makusky A J, *et al.* The human serum proteome: Display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. *Proteomics*, 2003, **3** (7): 1345~1364
- 3 Tirumalai R S, Chan K C, Prieto D A, *et al.* Characterization of the low molecular weight human serum proteome. *Mol Cell Proteomics*, 2003, **2** (10): 1096~1103
- 4 Adkins J N, Varnum S M, Auberry K J, *et al.* Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Mol Cell Proteomics*, 2002, **1** (12): 947~955
- 5 Anderson N L, Polanski M, Pieper R, *et al.* The human plasma proteome: a non-redundant list developed by combination of four separate sources. *Mol Cell Proteomics*, 2004, **3** (4): 311~326
- 6 Blackstock W P, Weir M P. Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol*, 1999, **17** (3): 121~127
- 7 Hanash S, Celis J E. The Human Proteome Organization: a mission to advance proteome knowledge. *Mol Cell Proteomics*, 2002, **1**

- (6): 413 ~414
- 8 Omenn G S. The human proteome organization plasma proteome project pilot phase: reference specimens, technology platform comparisons, and standardized data submissions and analyses. *Proteomics*, 2004, **4** (5): 1235 ~1240
 - 9 Smith J C Jr, Lewis S, Holbrook J, *et al.* Effect of heparin and citrate on measured concentrations of various analytes in plasma. *Clin Chem*, 1987, **33** (6): 814 ~816
 - 10 Sadagopan N P, Li W, Cook J A, *et al.* Investigation of EDTA anticoagulant in plasma to improve the throughput of liquid chromatography/tandem mass spectrometric assays. *Rapid Commun Mass Spectrom*, 2003, **17** (10): 1065 ~1070
 - 11 Glendenning P, Laffer L L, Weber H K, *et al.* Parathyroid hormone is more stable in EDTA plasma than in serum. *Clin Chem*, 2002, **48** (5): 766 ~767
 - 12 Beheshti I, Wessels L M, Eckfeldt J H. EDTA-plasma vs serum differences in cholesterol, high-density-lipoprotein cholesterol, and triglyceride as measured by several methods. *Clin Chem*, 1994, **40** (11 Pt 1): 2088 ~2092
 - 13 Pieper R, Su Q, Gatlin C L, *et al.* Multi-component immunoaffinity subtraction chromatography: An innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics*, 2003, **3** (4): 422 ~432
 - 14 Zhang K, Zolotarjova N, Nicol G, *et al.* Agilent multiple affinity removal system for the development of high-abundant proteins from human serum—a new technology from agilent (Application) © Aligent Technologies, Inc. <http://www.chem.agilent.com/Scripts/PDS.asp?IPage=10784>, 2003
 - 15 Harry J L, Wilkins M R, Herbert B R, *et al.* Proteomics: capacity versus utility. *Electrophoresis*, 2000, **21** (6): 1071 ~1081
 - 16 Gygi S P, Corthals G L, Zhang Y, *et al.* Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci USA*, 2000, **97** (17): 9390 ~9395
 - 17 Starita-Gerbaldi M, Roux F, Garin J, *et al.* Development of narrow immobilized pH gradients covering one pH unit for human seminal plasma proteomic analysis. *Proteomics*, 2003, **3** (8): 1611 ~1619
 - 18 Giometti C S, Khare T, Tollaksen S L, *et al.* Analysis of the *Shewanella oneidensis* proteome by two-dimensional gel electrophoresis under nondenaturing conditions. *Proteomics*, 2003, **3** (5): 777 ~785
 - 19 Manabe T, Yamaguchi N, Mukai J, *et al.* Detection of protein-protein interactions and a group of immunoglobulin G-associated minor proteins in human plasma by nondenaturing and denaturing two-dimensional gel electrophoresis. *Proteomics*, 2003, **3** (6): 832 ~846
 - 20 Watkins B, Szaro R, Ball S, *et al.* Detection of early-stage cancer by serum protein analysis. *Am Lab*, 2001, **33** (1): 32 ~36
 - 21 Petricoin E F, Ardekani A M, Hitt B A, *et al.* Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*, 2002, **359** (9306): 572 ~577
 - 22 Li J, Zhang Z, Rosenzweig J, *et al.* Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin Chem*, 2002, **48** (8): 1296 ~1304
 - 23 Adam B L, Qu Y, Davis J W, *et al.* Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res*, 2002, **62** (13): 3609 ~3614

Plasma Proteome: A “Pathfinder” of Human Proteome Project*

WANG Ying, ZHAO Xiao-Hang**

(National Laboratory of Molecular Oncology, Cancer Institute, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100021, China)

Abstract Blood plasma is a complex body fluid that contains various proteins and small molecules including peptides, salts, lipids, amino acids and sugars. Plasma proteins perform different kinds of housekeeping functions, such as immunoresponse, coagulation, anticoagulation, transportation, nutrition and regulating signaling cascades, which normally are altered both in structure and amount during pathogenesis. These characteristics of plasma proteome are critical for disease diagnosis and therapeutic monitoring. Nevertheless, only a handful of proteins are currently well known and used in routine clinical diagnosis. A comprehensive, systematic characterization of circulating proteins and peptides in health and disease will greatly facilitate development of biomarkers for many human diseases. The human proteome organization has launched the Human Plasma Proteome Project as one of the major initiatives since April 2002. In the pilot phase of the project, as a “pathfinder” of human proteomics study, the major goals are (1) to compare the advantages and limitations of many technology platforms; (2) to compare reference specimens of human plasma (EDTA, heparin, citrate-anti-coagulated) and serum prepared specifically for this project with the technology platforms; and (3) to create a knowledge base. The brief introduction about the plasma proteome project including research status, major problems faced and technical strategies been used are reviewed.

Key words plasma/serum proteins, proteomics, reference specimen, biomarkers, mass spectrometry

* This work was supported by grants from The National Natural Sciences Foundation of China (30370713, 30225045 and 30171049) and State 863 High Technology R&D Project of China (2004AA227060).

** Corresponding author. Tel: 86-10-67709015, E-mail: zhaoxh@pubem.cicams.ac.cn

Received: February 9, 2004 Accepted: March 29, 2004