



(12)发明专利申请

(10)申请公布号 CN 105758928 A

(43)申请公布日 2016.07.13

(21)申请号 201610109049.X

(22)申请日 2016.02.26

(71)申请人 中国科学院计算技术研究所
地址 100190 北京市海淀区中关村科学院南路6号

申请人 中国科学院生物物理研究所

(72)发明人 孙世伟 卜东波 杨飞 王耀军
李岩 黄纯翠 陈润生 高枫
刘亚名

(74)专利代理机构 北京律诚同业知识产权代理有限公司 11006

代理人 祁建国 鲍俊萍

(51)Int.Cl.
G01N 27/62(2006.01)

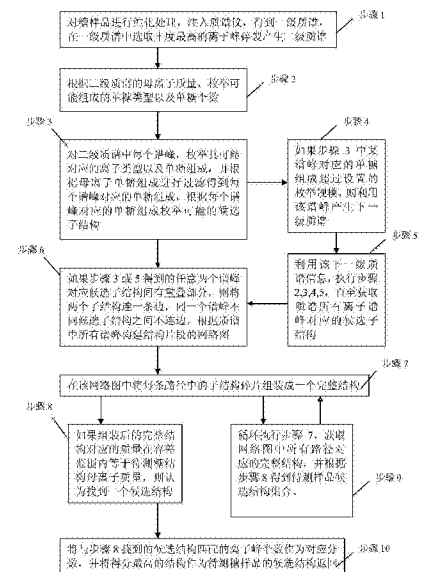
权利要求书2页 说明书8页 附图7页

(54)发明名称

一种糖结构鉴定方法及糖结构鉴定装置

(57)摘要

本发明公开了一种糖结构鉴定方法及糖结构鉴定装置,所述糖结构鉴定方法包括:步骤一:利用多级质谱信息进行待测糖结构谱峰对应的子结构预测,所述子结构作为对应谱峰的候选子结构;以及步骤二:利用De Novo糖结构鉴定技术,根据质谱谱峰对应的候选子结构组装出完整的糖结构。本发明基于糖结构多级质谱数据,采用De Novo糖结构鉴定技术,实现糖结构的鉴定,实现了多级质谱De Novo糖结构鉴定算法,通过有效的利用多级质谱信息,不仅显著提高了De Novo过程中枚举糖结构片段的效率,而且明显缩小了糖候选结构集合。



1. 一种糖结构鉴定方法,其特征在于,包括步骤:

步骤一:利用多级质谱信息进行待测糖结构谱峰对应的子结构预测,所述子结构作为候选子结构;以及

步骤二:利用De Novo糖结构鉴定技术,根据质谱谱峰对应的候选子结构组装出完整的糖结构。

2. 根据权利要求1所述的糖结构鉴定方法,其特征在于,所述步骤一包括:

步骤1:对糖样品进行纯化处理,注入质谱仪,得到一级质谱,在一级质谱中选取丰度最高的离子峰碎裂产生二级质谱;

步骤2:根据二级质谱的母离子质量,枚举可能组成的单糖类型以及每种类型的单糖个数,得到母离子单糖组成;

步骤3:对二级质谱中每个谱峰,枚举可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤得到每个谱峰对应的单糖组成,根据每个谱峰对应的单糖组成枚举可能的候选子结构。

3. 根据权利要求2所述的糖结构鉴定方法,其特征在于,所述步骤一还包括:

步骤4:如果步骤3中某谱峰对应的单糖组成超过设置的枚举规模,则利用该谱峰产生下一级质谱,谱峰 P_3 对应的单糖组成超过枚举规模,则采用下一级质谱组装出该谱峰对应的候选子结构;以及

步骤5:利用该下一级质谱信息,执行步骤2,3,4,5,直至获取质谱所有离子谱峰对应的候选子结构。

4. 根据权利要求3所述的糖结构鉴定方法,其特征在于,所述步骤二包括:

步骤6:如果步骤3或5得到的任意两个谱峰对应候选子结构间存在重叠部分,则将两个子结构连一条边,同一个谱峰不同候选子结构之间不连边,根据质谱中所有谱峰构建候选子结构的网络图;

步骤7:在该网络图中将每条路径中的结构碎片组装成一个完整结构;以及

步骤8:如果组装后的完整结构对应的质量在容差范围内等于待测糖结构母离子质量,则认为找到一个待测糖样品候选结构。

5. 根据权利要求4所述的糖结构鉴定方法,其特征在于,所述步骤二还包括:步骤9:循环执行步骤7,获取网络图中所有路径对应的完整结构,并根据步骤8得到待测样品候选结构集合。

6. 根据权利要求4所述的糖结构鉴定方法,其特征在于,还包括:

步骤10:将与步骤8找到的待测糖样品候选结构匹配的离子峰个数作为对应分数,并将得分最高的结构作为糖的候选结构返回。

7. 一种糖鉴定装置,其特征在于,包括:

用户web交互界面;

谱峰候选子结构鉴定模块,与所述用户web交互界面连接,所述谱峰候选子结构鉴定模块包括:

质谱数据模块:利用质谱仪对糖样品进行处理得到一级质谱,在一级质谱中选取丰度最高的离子峰碎裂产生所述二级质谱数据;

二级质谱枚举模块:根据所述二级质谱的母离子质量,枚举可能组成的单糖类型以及

单糖个数,得到母离子单糖组成;对二级质谱中每个谱峰,枚举其可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤得到每个谱峰对应的单糖组成,根据每个谱峰对应的单糖组成枚举可能的候选子结构;

De Novo糖结构鉴定模块,包括:

候选子结构网络图构建模块:接收二级质谱枚举模块传送的候选结构,如果任意两个谱峰对应的候选子结构间存在重叠部分,则将两个子结构连一条边,同一个谱峰不同候选子结构之间不连边,根据质谱中所有谱峰构建候选子结构网络图;

组装模块:在所述网络图中将每条路径中的结构碎片组装成一个完整结构;以及

候选结构质量判断模块:如果组装后的完整结构对应的质量在容差范围内等于待测糖结构母离子质量,则认为找到一个待测糖样品候选结构,循环执行步骤7和步骤8,将网络图中所有的路径对应的可能待测糖样品候选结构,构成待测糖样品候选结构集合,将此待测糖样品候选结构集合返回用户web交互界面。

8. 根据权利要求7的糖鉴定装置,其特征在于,所述谱峰候选子结构鉴定模块还包括:

判断与生成模块:如果二级质谱枚举模块中某谱峰对应的单糖组成超过枚举模块设置的二级质谱枚举规模,则利用该谱峰产生下一级质谱;以及

多级质谱枚举模块:根据所述下一级质谱的母离子质量,枚举可能组成的单糖类型以及单糖个数,得到母离子单糖组成;对下一级质谱中每个谱峰,枚举其可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤得到每个谱峰对应的单糖组成,根据每个谱峰对应的单糖组成枚举可能的候选子结构,将所述可能的候选子结构传送至候选子结构网络图构建模块。

9. 根据权利要求7的糖鉴定装置,其特征在于,所述De Novo糖结构鉴定模块还包括:

筛选模块:将与组装模块中找到的待测糖样品候选结构匹配的离子峰个数作为对应分数,并将得分最高的结构作为待测糖样品的候选结构返回用户web交互界面。

一种糖结构鉴定方法及糖结构鉴定装置

技术领域

[0001] 本发明涉及生物信息领域,尤其涉及一种糖结构鉴定方法及糖结构鉴定装置。

背景技术

[0002] 糖类化合物是生命活动中起重要功能的生物大分子,其以多种结构构型存在于细胞中,在细胞周期调控、凋亡衰老、细胞表面的相互作用等生命过程起到了不可替代的作用。

[0003] 糖类化合物通过由多个单糖通过糖苷键连接而成,并呈现树形分支结构,因此,糖类化合物的鉴定主要包含糖类组成分子、单糖连接顺序以及分支位点等信息的鉴定。图1a及图1b为单糖对应的环形结构以及糖结构中糖苷键的表示方式。图2a至图2d表示N-linked糖G1cNAc2Man9结构的多种表示方式,其中根节点位于结构的最右端,即N端,子节点向左逐步延伸,形成C端,其中每个节点代表一个单糖,每条边代表两个单糖连接的糖苷键。由于一个单糖可以通过糖苷键与其他一个或者多个单糖连接形成分支结构,导致了糖类化合物结构构型的多样性。

[0004] 目前基于质谱的糖类化合物鉴定策略大部分基于一级质谱和二级质谱数据。一级质谱只能得到糖化合物母离子的质量,而无法得到糖化合物结构的单糖组成、单糖间的连接位点以及结构信息。二级质谱是将糖在质谱仪中进一步打碎进行子结构信息分析,目前主要包含以下几种策略:

[0005] 1)结构库搜索策略

[0006] 在糖结构库中找出母离子质量与待鉴定质谱母离子质量相近的糖结构,然后预测这些糖结构的理论谱,将预测的理论谱与待鉴定质谱一一比较,返回相似度最高的理论谱对应的糖结构作为待鉴定质谱对应的糖结构。

[0007] 优势:根据结构库限定了可能的糖结构的候选范围,从而降低了鉴定难度。

[0008] 不足:a).结构库搜索依赖于库中结构的理论质谱的预测,而目前对于糖结构质谱形成机制认识仍然有局限,导致理论质谱预测精度不高,在一定程度上影响了鉴定结果的准确性。b).结构库搜索局限于库中的糖结构,如果待鉴定质谱对应的糖结构不在库中,则无法得到正确的鉴定结果。

[0009] 2)De Novo结构鉴定策略

[0010] 与结构库搜索策略不同,De Novo策略不依赖于已知的糖结构库,而是直接通过对质谱数据的分析,利用谱图中谱峰间的m/z差值,推断可能的糖结构。

[0011] 优势:相对于结构库搜索而言,De Novo策略能够发现结构库中不存在的糖结构。

[0012] 缺陷:a).De Novo策略鉴定的准确度严重依赖于质谱数据的质量。在高质量的质谱数据中,每个糖苷键都会有相应的离子峰出现;而在低质量的质谱中,部分离子峰的缺失,导致De Novo策略无法得出准确的鉴定结果。b).De Novo需要枚举所有可能的糖结构,因此效率比较低,鉴定规模比较有限。

[0013] 3)谱库搜索策略

[0014] 谱库搜索策略将已知糖结构的二级质谱以“结构-质谱”对的形式保存在数据库中,然后用待鉴定质谱与数据库中的真实谱进行比较,返回相似度最高的质谱对应的糖结构作为鉴定结果。

[0015] 优势:相比较糖结构库搜索策略,谱库搜索策略使用已知的真实谱,而不是预测出的理论谱进行比较,从而使得鉴定结果的可信度比较高。

[0016] 缺陷:谱库搜索策略仅适用于已鉴定过的糖结构,对于未鉴定过的质谱则无法确定其对应的结构。

[0017] 二级质谱提供的信息相对于糖类化合物复杂的结构来说还是具有很大的局限性,造成基于二级质谱的糖类化合物鉴定策略准确度不高。因此,有些研究者提出基于多级质谱鉴定策略,即利用多级质谱数据中包含更多的糖结构碎裂信息,实现提升糖结构鉴定的准确率。现有一种基于多级质谱鉴定方法是使用已知糖结构的二级以及三级质谱构建“质谱-结构”对数据库,此方法在一定程度上提升了鉴定的准确度,但是由于其谱库中谱数量比较少,所以对于有些糖结构仍无法很好地鉴定。

发明内容

[0018] 本发明提供一种糖结构鉴定方法及糖结构鉴定装置,以解决现有技术DeNovo策略鉴定的准确度严重依赖于质谱数据的质量及谱库搜索策略对于未鉴定过的质谱无法确定其对应的结构的缺陷。

[0019] 为实现上述目的,本发明提供一种糖结构鉴定方法,包括步骤:

[0020] 步骤一:利用多级质谱信息进行待测糖结构谱峰对应的子结构预测,所述子结构作为候选子结构;以及

[0021] 步骤二:利用De Novo糖结构鉴定技术,根据质谱谱峰对应的候选子结构组装出完整的糖结构。

[0022] 较佳地,所述步骤一包括:

[0023] 步骤1:对糖样品进行纯化处理,注入质谱仪,得到一级质谱,在一级质谱中选取丰度最高的离子峰碎裂产生二级质谱;

[0024] 步骤2:根据二级质谱的母离子质量,枚举可能组成的单糖类型以及每种类型的单糖个数,得到母离子单糖组成;以及

[0025] 步骤3:对二级质谱中每个谱峰,枚举可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤得到每个谱峰对应的单糖组成,根据每个谱峰对应的单糖组成枚举可能的候选子结构。

[0026] 较佳地,所述步骤一还包括:

[0027] 步骤4:如果步骤3中某谱峰对应的单糖组成超过设置的枚举规模,则利用该谱峰产生下一级质谱,谱峰 P_3 对应的单糖组成超过枚举规模,则采用下一级质谱组装出该谱峰对应的候选子结构;以及

[0028] 步骤5:利用该下一级质谱信息,执行步骤2,3,4,5,直至获取质谱所有离子谱峰对应的候选子结构。

[0029] 较佳地,所述步骤二包括:

[0030] 步骤6:如果步骤3或5得到的任意两个谱峰对应候选子结构间存在重叠部分,则将

两个子结构连一条边,同一个谱峰不同候选子结构之间不连边,根据质谱中所有谱峰构建候选子结构的网络图;

[0031] 步骤7:在该网络图中将每条路径中的结构碎片组装成一个完整结构;以及

[0032] 步骤8:如果组装后的完整结构对应的质量在容差范围内等于待测糖结构母离子质量,则认为找到一个待测糖样品候选结构。

[0033] 较佳地,所述步骤二还包括:步骤9:循环执行步骤7,获取网络图中所有路径对应的完整结构,并根据步骤8得到待测样品候选结构集合。

[0034] 较佳地,所述的糖结构鉴定方法,还包括:

[0035] 步骤10:将与步骤8找到的待测糖样品候选结构匹配的离子峰个数作为对应分数,并将得分最高的结构作为糖的候选结构返回。

[0036] 而且,为实现上述目的,本发明还提供一种糖鉴定装置,包括:

[0037] 用户web交互界面;

[0038] 谱峰候选子结构鉴定模块,与所述用户web交互界面连接,所述谱峰候选子结构鉴定模块包括:

[0039] 质谱数据模块:利用质谱仪对糖样品进行处理得到一级质谱,在一级质谱中选取丰度最高的离子峰碎裂产生所述二级质谱数据;及

[0040] 二级质谱枚举模块:根据所述二级质谱的母离子质量,枚举可能组成的单糖类型以及单糖个数,得到母离子单糖组成;对二级质谱中每个谱峰,枚举其可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤得到每个谱峰对应的单糖组成,根据每个谱峰对应的单糖组成枚举可能的候选子结构;以及

[0041] De Novo糖结构鉴定模块,包括:

[0042] 候选子结构网络图构建模块:接收二级质谱枚举模块传送的候选结构,如果任意两个谱峰对应的候选子结构间存在重叠部分,则将两个子结构连一条边,同一个谱峰不同候选子结构之间不连边,根据质谱中所有谱峰构建候选子结构网络图;

[0043] 组装模块:在所述网络图中将每条路径中的结构碎片组装成一个完整结构;及

[0044] 候选结构质量判断模块:如果组装后的完整结构对应的质量在容差范围内等于待测糖结构母离子质量,则认为找到一个待测糖样品候选结构,循环执行步骤7和步骤8,将网络图中所有的路径对应的可能待测糖样品候选结构,构成待测糖样品候选结构集合,将此待测糖样品候选结构集合返回用户web交互界面。

[0045] 较佳地,所述谱峰候选子结构鉴定模块还包括:

[0046] 判断与生成模块:如果二级质谱枚举模块中某谱峰对应的单糖组成超过枚举模块设置的二级质谱枚举规模,则利用该谱峰产生下一级质谱;以及

[0047] 多级质谱枚举模块:根据所述下一级质谱的母离子质量,枚举可能组成的单糖类型以及单糖个数,得到母离子单糖组成;对下一级质谱中每个谱峰,枚举其可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤得到每个谱峰对应的单糖组成,根据每个谱峰对应的单糖组成枚举可能的候选子结构,将所述可能的候选子结构传送至候选子结构网络图构建模块。

[0048] 较佳地,所述De Novo糖结构鉴定模块还包括:

[0049] 筛选模块:将与组装模块中找到的待测糖样品候选结构匹配的离子峰个数作为对

应分数,并将得分最高的结构作为待测糖样品的候选结构返回用户web交互界面。

[0050] 本发明基于糖结构多级质谱数据,采用De Novo糖结构鉴定技术,实现糖结构的鉴定。解决了现有技术De Novo策略鉴定的准确度严重依赖于质谱数据的质量及谱库搜索策略对于未鉴定过的质谱无法确定其对应的结构的缺陷。实现了多级质谱De Novo糖结构鉴定算法,通过有效的利用多级质谱信息,不仅显著提高了De Novo过程中枚举糖结构片段的效率,而且明显缩小了糖候选结构集合。

[0051] 以下结合附图和具体实施例对本发明进行详细描述,但不作为对本发明的限定。

附图说明

[0052] 图1a为现有技术单糖环形结构示意图,其中1,2,3,4和6对应的C原子为单糖的连接位点;图1b为现有技术单糖糖苷键结构示意图,其中,两个单糖连接形成1-4糖苷键;

[0053] 图2a至图2d为现有技术N-linked糖GlcNAc2Man9结构的多种表示方式示意图;

[0054] 图3为本发明糖结构鉴定方法步骤图;

[0055] 图4为本发明糖结构鉴定装置示意图;

[0056] 图5为6种N-linked糖使用本发明鉴定方法得到的候选结果;

[0057] 图6利用谱峰P₃的下一级质谱产生其对应的候选结构;

[0058] 图7为本发明糖结构组装过程示意图。

[0059] 其中,附图标记:

[0060] 1用户web交互界面

[0061] 2谱峰候选子结构鉴定模块

[0062] 21质谱数据模块

[0063] 22二级质谱枚举模块

[0064] 23判断与生成模块

[0065] 24多级质谱枚举模块

[0066] 3 De Novo糖结构鉴定模块

[0067] 31候选子结构网络图构建模块

[0068] 32组装模块

[0069] 33候选结构质量判断模块

[0070] 34筛选模块

具体实施方式

[0071] 本发明构建有效的De Novo策略,结合多级质谱数据实现糖结构的鉴定,本发明的糖结构鉴定方法主要包括2方面:

[0072] De Novo糖结构组装算法:根据质谱谱峰对应的候选子结构组装出完整的糖结构。本发明通过各个候选子结构之间的重叠关系构建网络图,在图中找出能够组装得到的完整糖结构的路径。以及

[0073] 多级质谱实现糖结构的子结构的预测算法:由于下一级质谱包含上一级质谱谱峰对应子结构碎裂后产生的所有片段信息,所以本发明综合多级质谱信息进行谱峰的子结构预测,显著提高了子结构候选集的鉴定准确率和效率。

[0074] 其中,本发明的De Novo糖结构组装算法根据二级质谱中所有离子谱峰对应的候选子结构,利用不同离子候选子结构间的重叠关系构建网络图,然后在图中找出组装后质量等于糖样品母离子质量的路径,具体为:

[0075] 1、如果任意两个谱峰对应候选子结构间存在重叠部分,则将两个子结构连一条边,同一个谱峰不同候选子结构之间不连边,根据质谱中所有谱峰构建子结构片段的网络图;

[0076] 2、在网络图中将每条路径中的结构碎片组装成一个完整的糖结构,如果组装后的糖结构对应的质量等于所求糖结构母离子质量(容差范围内),则认为找到一个候选糖结构;

[0077] 3、将候选糖结构匹配的离子峰个数作为对应分数,并将得分最高的糖结构作为糖的候选结构返回。

[0078] 并且,糖结构的复杂树形结构导致了明显的同分异构现象,即同一个分子质量对应的糖结构可能有很多,从而给De Novo进行结构枚举带来了很大的难度,严重制约了De Novo的鉴定规模。本发明针对现有De Novo鉴定规模不足的问题,设计了利用多级质谱信息实现对于复杂谱峰的候选结构鉴定,从而提升De Novo鉴定的效率。本发明的多级质谱实现糖结构的子结构预测算法为:

[0079] 1、对糖样品进行纯化处理,注入质谱仪,得到一级质谱,在一级质谱中选取丰度最高的离子峰碎裂产生二级质谱;

[0080] 2、根据二级质谱的母离子质量,枚举可能组成的单糖类型以及每种类型的单糖个数;

[0081] 3、对二级质谱中每个谱峰,枚举其可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤;

[0082] 4、根据每个谱峰对应的单糖组成枚举可能的候选子结构,如果谱峰对应的单糖组成超过设置的枚举规模,则利用该谱峰产生下一级质谱,并转到5,见图6,谱峰P₃对应的单糖组成超过枚举规模,则采用下一级质谱组装出该谱峰对应的候选子结构;

[0083] 5、利用下一级质谱信息,执行2,3,4,5步,获取质谱所有离子谱峰对应的候选子结构,并利用糖结构组装算法,组装出该谱对应的候选子结构作为上一级谱相应谱峰的候选子结构。

[0084] 具体而言,图3为本发明的糖结构鉴定方法的方法步骤图,如图3所示,本发明的方法包括:

[0085] 步骤一:利用多级质谱信息进行待测糖结构谱峰对应的子结构预测,所述子结构作为候选子结构;以及

[0086] 步骤二:利用De Novo糖结构鉴定技术,根据质谱谱峰对应的候选子结构组装出完整的糖结构。

[0087] 其中,所述步骤一包括:

[0088] 步骤1:对糖样品进行纯化处理,注入质谱仪,得到一级质谱,在一级质谱中选取丰度最高的离子峰碎裂产生二级质谱;

[0089] 步骤2:根据二级质谱的母离子质量,枚举可能组成的单糖类型以及每种类型的单糖个数,得到母离子单糖组成;以及

[0090] 步骤3:对二级质谱中每个谱峰,枚举其可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤得到每个谱峰对应的单糖组成,根据每个谱峰对应的单糖组成枚举可能的候选子结构。

[0091] 较佳地,所述步骤一还包括:

[0092] 步骤4:如果步骤3中某谱峰对应的单糖组成超过设置的枚举规模,则利用该谱峰产生下一级质谱,见图6,谱峰P₃对应的单糖组成超过枚举规模,则采用该谱峰对应的下一级质谱进行该谱峰候选子结构的预测,组装出该谱峰对应的候选子结构;以及

[0093] 步骤5:利用该下一级质谱信息,执行步骤2,3,4,5,直至获取质谱所有离子谱峰对应的候选子结构。

[0094] 其中,所述步骤二包括:

[0095] 步骤6:如果步骤3或5得到的任意两个谱峰对应候选子结构间存在重叠部分,则将两个子结构连一条边,同一个谱峰不同候选子结构之间不连边,根据质谱中所有谱峰构建候选子结构的网络图;

[0096] 步骤7:在该网络图中将每条路径中的结构碎片组装成一个完整结构,见图7,图7是本发明鉴定方法的组装过程,通过多级质谱得到质谱图中谱峰对应的可能候选子结构,并根据候选子结构是否重叠构建候选子结构网络图,然后利用图中的路径得到待测糖样品对应的候选结构集合,以及每个候选结构对应的打分(匹配的离子峰个数);以及

[0097] 步骤8:如果组装后的完整结构对应的质量在容差范围内等于待测糖结构母离子质量,则认为找到一个待测糖样品候选结构;

[0098] 较佳地,所述步骤二还包括:

[0099] 步骤:10:将与步骤8找到的待测糖样品候选结构匹配的离子峰个数作为对应分数,并将得分最高的结构作为待测糖样品候选结构返回。

[0100] 并且,较佳地,所述步骤二还包括:步骤9:循环执行步骤7,获取网络图中所有路径对应的完整结构,并根据步骤8得到待测样品候选结构集合。

[0101] 另外,本发明还提出了一种糖鉴定装置,如图4所示,包括:

[0102] 用户web交互界面1;

[0103] 谱峰候选子结构鉴定模块2,与所述用户web交互界面连接,所述谱峰候选子结构鉴定模块包括质谱数据模块21及二级质谱枚举模块22,其中,

[0104] 利用串联质谱仪对糖样品进行处理得到一级质谱,在一级质谱中选取丰度最高的离子峰碎裂产生所述二级质谱数据;及

[0105] 二级质谱枚举模块22根据所述二级质谱的母离子质量,枚举可能组成的单糖类型以及单糖个数,得到母离子单糖组成;对二级质谱中每个谱峰,枚举其可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤得到每个谱峰对应的单糖组成,根据每个谱峰对应的单糖组成枚举可能的候选子结构;以及

[0106] De Novo糖结构鉴定模块3,包括候选子结构网络图构建模块31、组装模块32及候选结构质量判断模块33,其中,

[0107] 候选子结构网络图构建模块31:接收二级质谱枚举模块传送的候选子结构,如果任意两个谱峰对应的候选子结构间存在重叠部分,则将两个子结构连一条边,同一个谱峰不同候选子结构之间不连边,根据质谱中所有谱峰构建候选子结构网络图;

[0108] 组装模块32:在所述网络图中将每条路径中的结构碎片组装成一个完整结构;及
[0109] 候选结构质量判断模块33:如果组装后的完整结构对应的质量在容差范围内等于待测糖结构母离子质量,则认为找到一个待测糖样品候选结构,循环执行步骤7和步骤8,将网络图中所有的路径对应的可能待测糖样品候选结构,构成待测糖样品候选结构集合。将此待测糖样品候选结构集合返回用户web交互界面;

[0110] 其中,所述谱峰候选子结构鉴定模块还包括:

[0111] 判断与生成模块23:如果二级质谱枚举模块中某谱峰对应的单糖组成超过枚举模块设置的二级质谱枚举规模,则利用该谱峰产生下一级质谱,如图6中谱峰P₃对应的单糖组成超过枚举规模,则采用下一级质谱组装出该谱峰对应的候选子结构;以及

[0112] 多级质谱枚举模块24:根据所述下一级质谱的母离子质量,枚举可能组成的单糖类型以及单糖个数,得到母离子单糖组成;对下一级质谱中每个谱峰,枚举其可能对应的离子类型以及单糖组成,并根据母离子单糖组成进行过滤得到每个谱峰对应的单糖组成,根据每个谱峰对应的单糖组成枚举可能的候选子结构,将所述可能的候选子结构传送至候选子结构网络图构建模块。

[0113] 较佳地,所述De Novo糖结构鉴定模块还包括筛选模块34:将与组装模块中找到的待测糖样品候选结构匹配的离子峰个数作为对应分数,并将得分最高的候选结构作为糖的候选结构返回用户web交互界面。

[0114] 实施例:多级质谱De Novo N-linked糖样品结构鉴定

[0115] 为了验证本发明糖结构鉴定策略的有效性,利用本发明上述方法对已知N-linked纯糖样品(来自Ludger公司)进行多级质谱De Novo鉴定实验,实验仪器为岛津公司的MALDI-IT-TOF质谱仪AXIMA Resonance。

[0116] 具体步骤:

[0117] 1).将纯糖样品经过处理后注入质谱仪,产生一级质谱,在一级质谱中按分子质量选择一级质谱中对应谱峰产生二级质谱。

[0118] 2).根据二级质谱的母离子质量,枚举可能的母离子单糖组成。

[0119] 3).枚举二级质谱中所有谱峰可能离子类型以及单糖组成,并根据母离子单糖组成过滤得到每个谱峰对应的单糖组成,并枚举产生相应的候选子结构。

[0120] 4).如果谱峰对应的单糖组成包含的单糖类型以及单糖个数比较多,则利用该谱峰产生下一级质谱,重复步骤2,3,4和5,得到该谱峰对应的候选子结构,见图6所示。

[0121] 5).根据二级质谱相应谱峰对应的候选子结构构建候选子结构网络图,任意两个谱峰对应候选子结构间存在重叠部分,则将两个子结构连一条边,同一个谱峰不同候选子结构之间不连边。

[0122] 6).根据候选子结构网络图得到待测糖样品候选结构集合,并根据候选结构匹配离子峰个数对候选结构集合中的候选结构进行打分,将分值对应的候选结构作为待测糖样品的候选结构返回。

[0123] 图5显示了6种N-linked糖的实验结果。第一列表示已知的N-linked糖样品名称,第二列表示糖对应的真实结构,第三列表示通过本发明策略得到的待测糖样品候选结构的个数,第四列为本发明策略得到的对应待测糖样品候选结构。

[0124] 结果分析:

[0125] 1)使用本发明的多级质谱De Novo鉴定策略能够得到待测糖样品的候选结构。

[0126] 2)本发明鉴定策略能将待测糖样品候选结构集合限制在一个很小的规模。

[0127] 技术效果

[0128] 本发明以寡糖的多级质谱数据作为输入,通过使用枚举和组装获取待测糖结构候选结构集合作为输出,可以应用于待测糖结构候选集合的预测,实现了多级质谱De Novo糖结构鉴定算法,通过有效的利用多级质谱信息,不仅显著提高了De Novo过程中枚举糖结构片段的效率,而且明显缩小了糖候选结构集合。

[0129] 当然,本发明还可有其它多种实施例,在不背离本发明精神及其实质的情况下,熟悉本领域的技术人员可根据本发明作出各种相应的改变和变形,但这些相应的改变和变形都应属于本发明权利要求的保护范围。

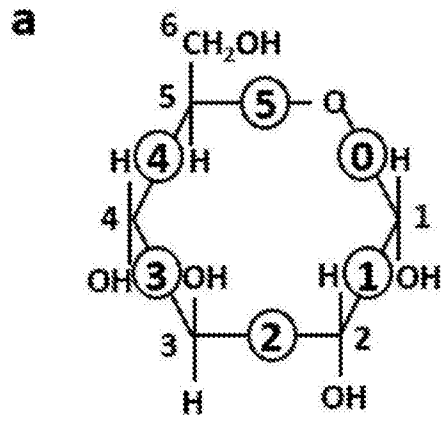


图1a

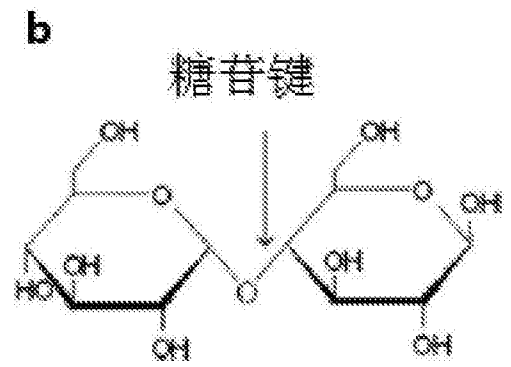


图1b

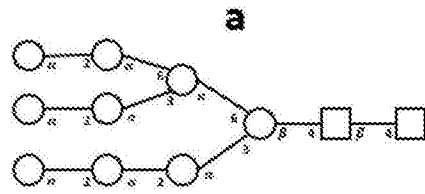


图2a

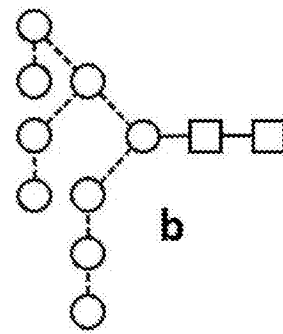


图2b

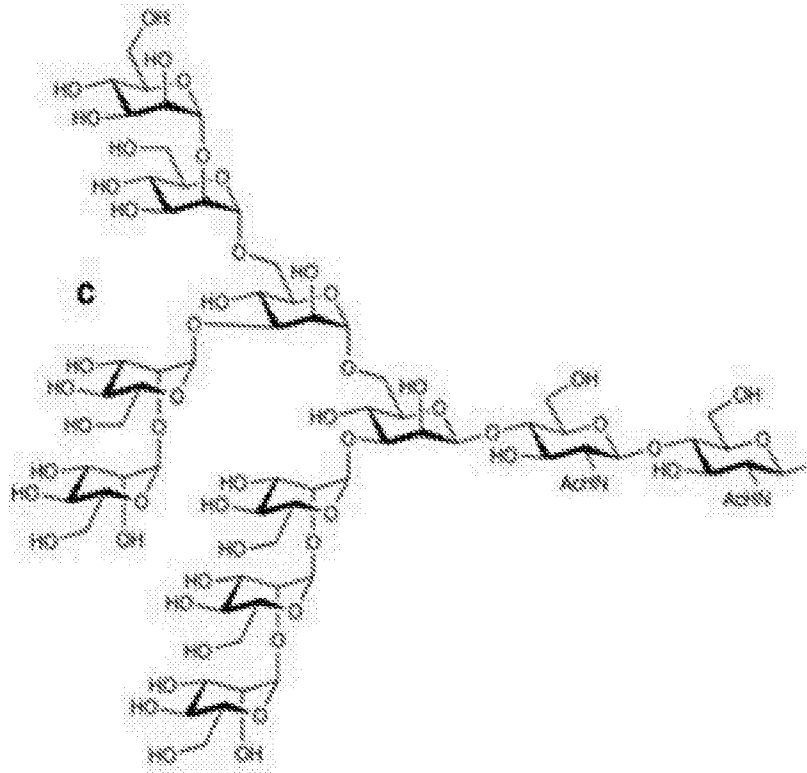


图2c

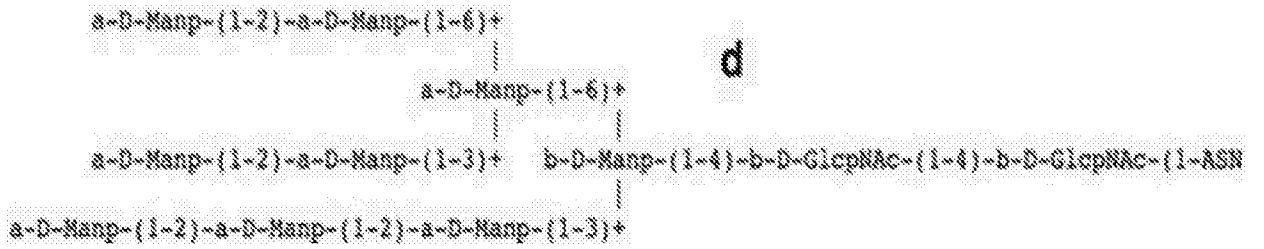


图2d

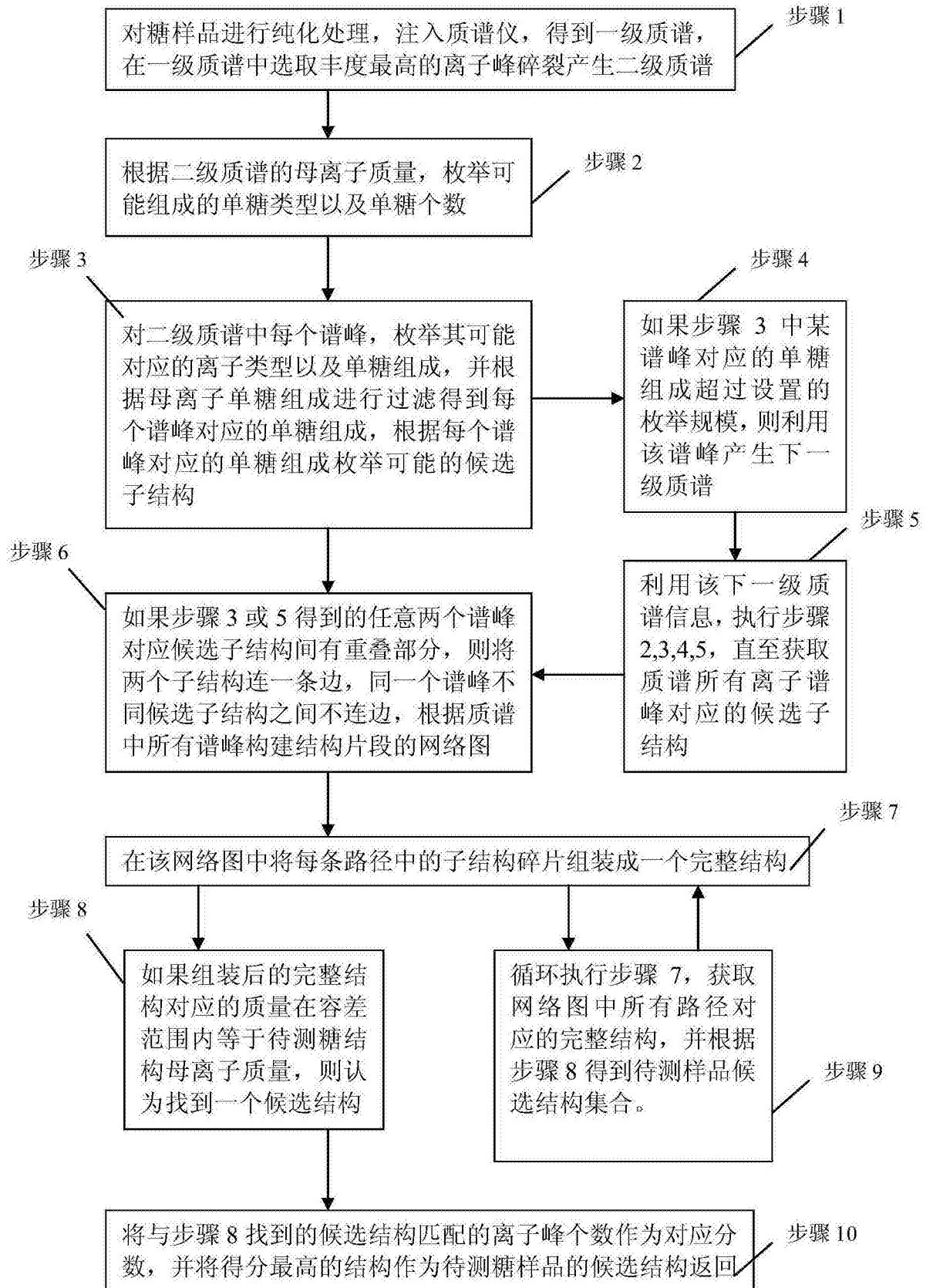


图3

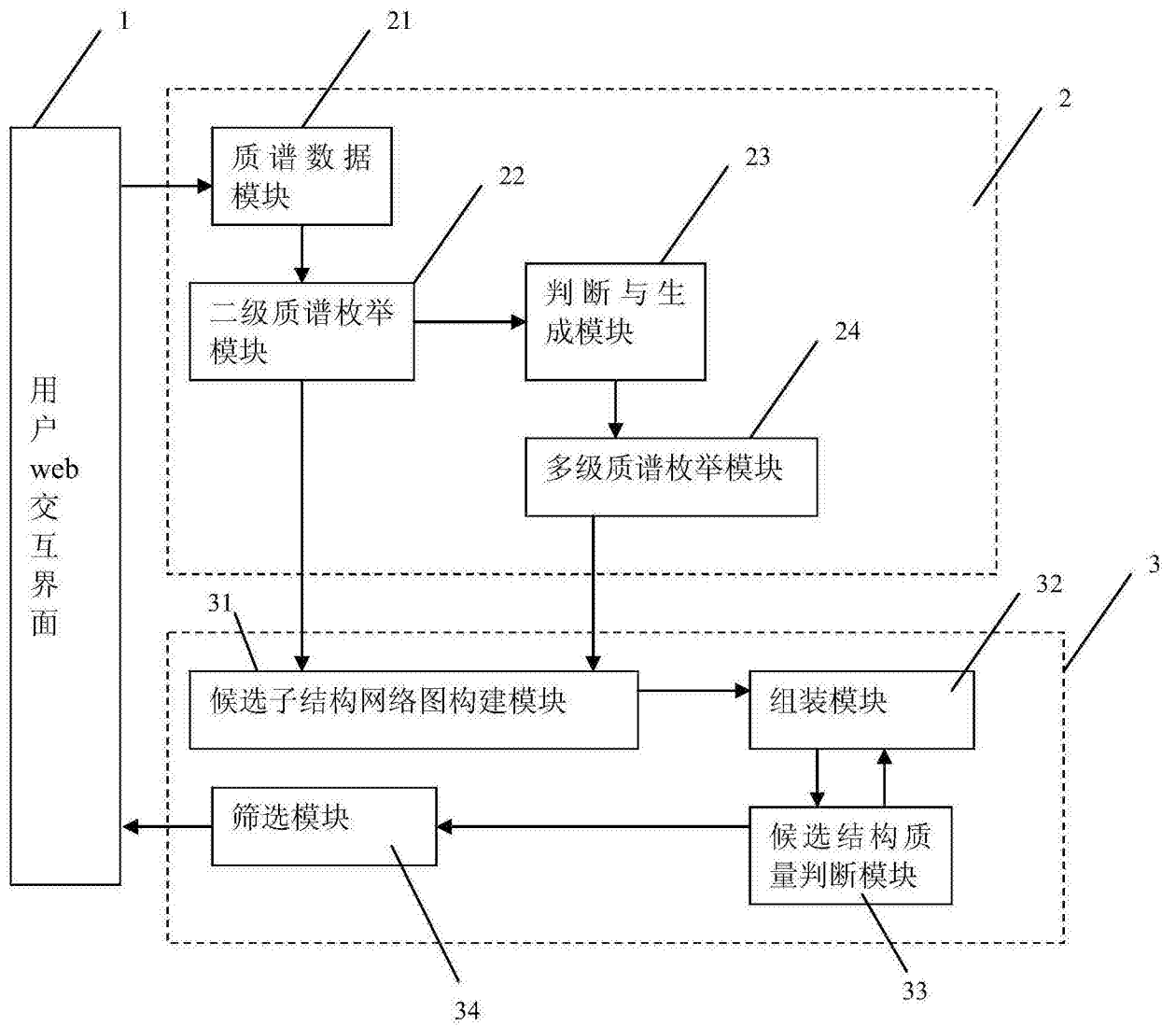


图4

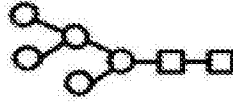
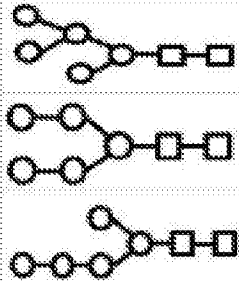
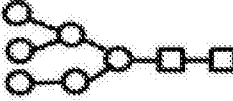
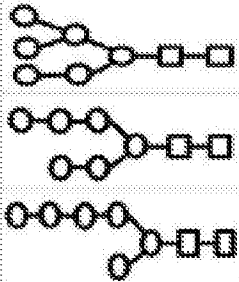
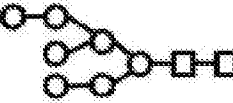
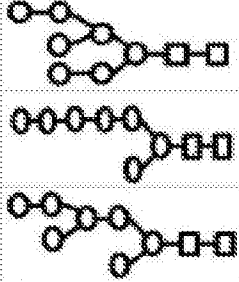
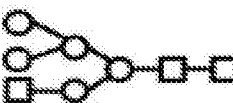
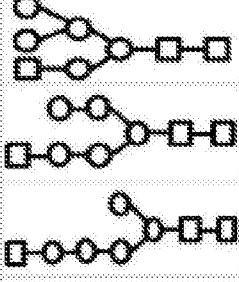
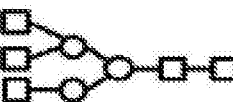
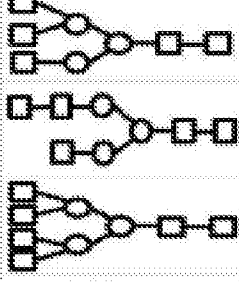
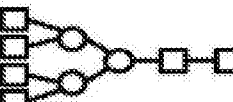
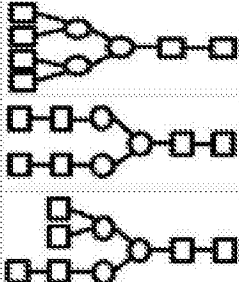
糖样品	真实结构	待测糖样品候选结构个数	待测糖样品候选结构
nan5		3	
nan6		3	
nan7		3	
hybrid		3	
NCA3		3	
NCA4		3	

图5

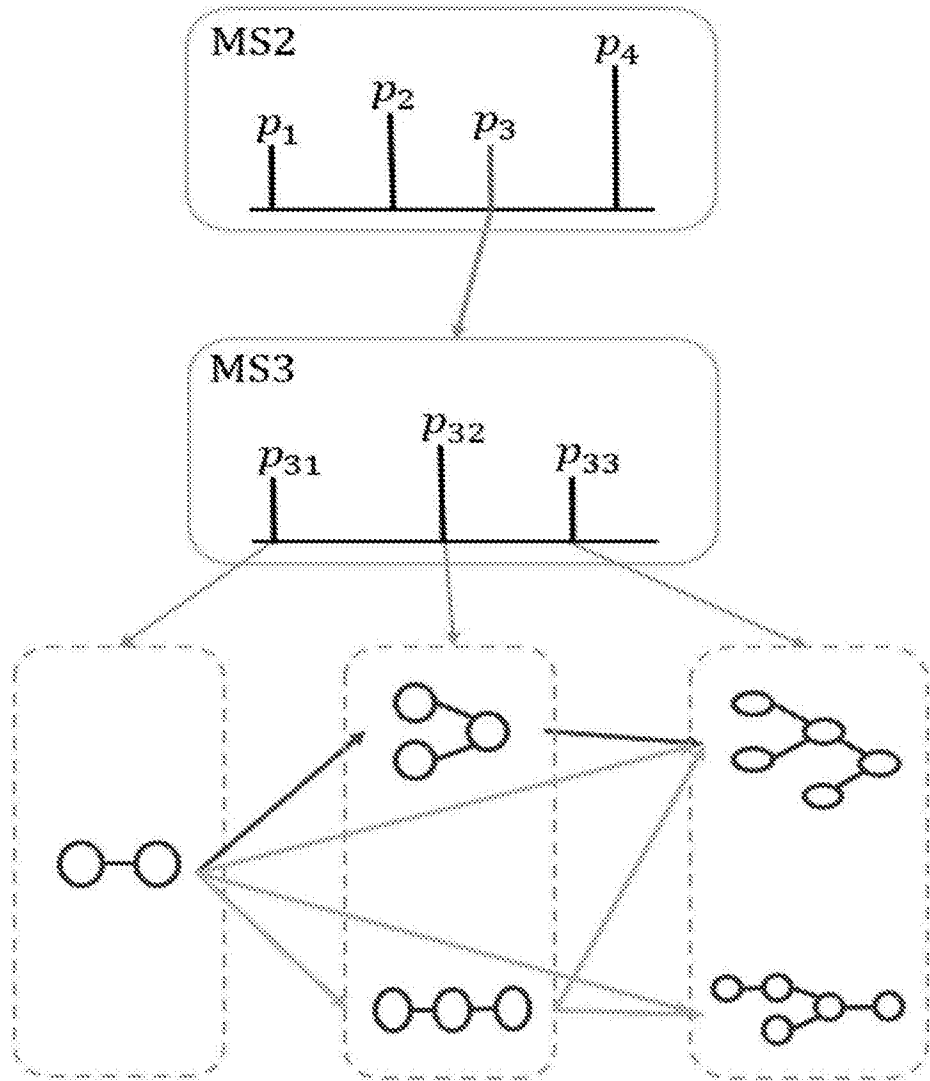


图6

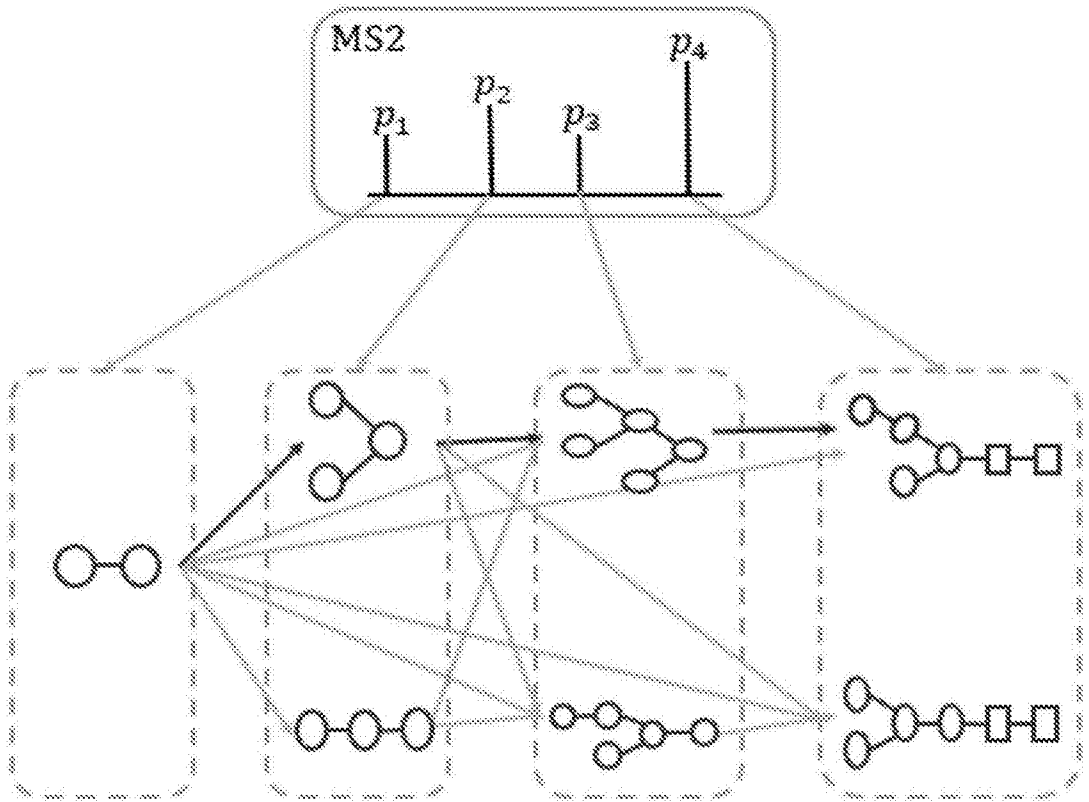


图7