Network analysis and cross species comparison of protein–protein interaction networks of human, mouse and rat cytochrome P450 proteins that degrade xenobiotics

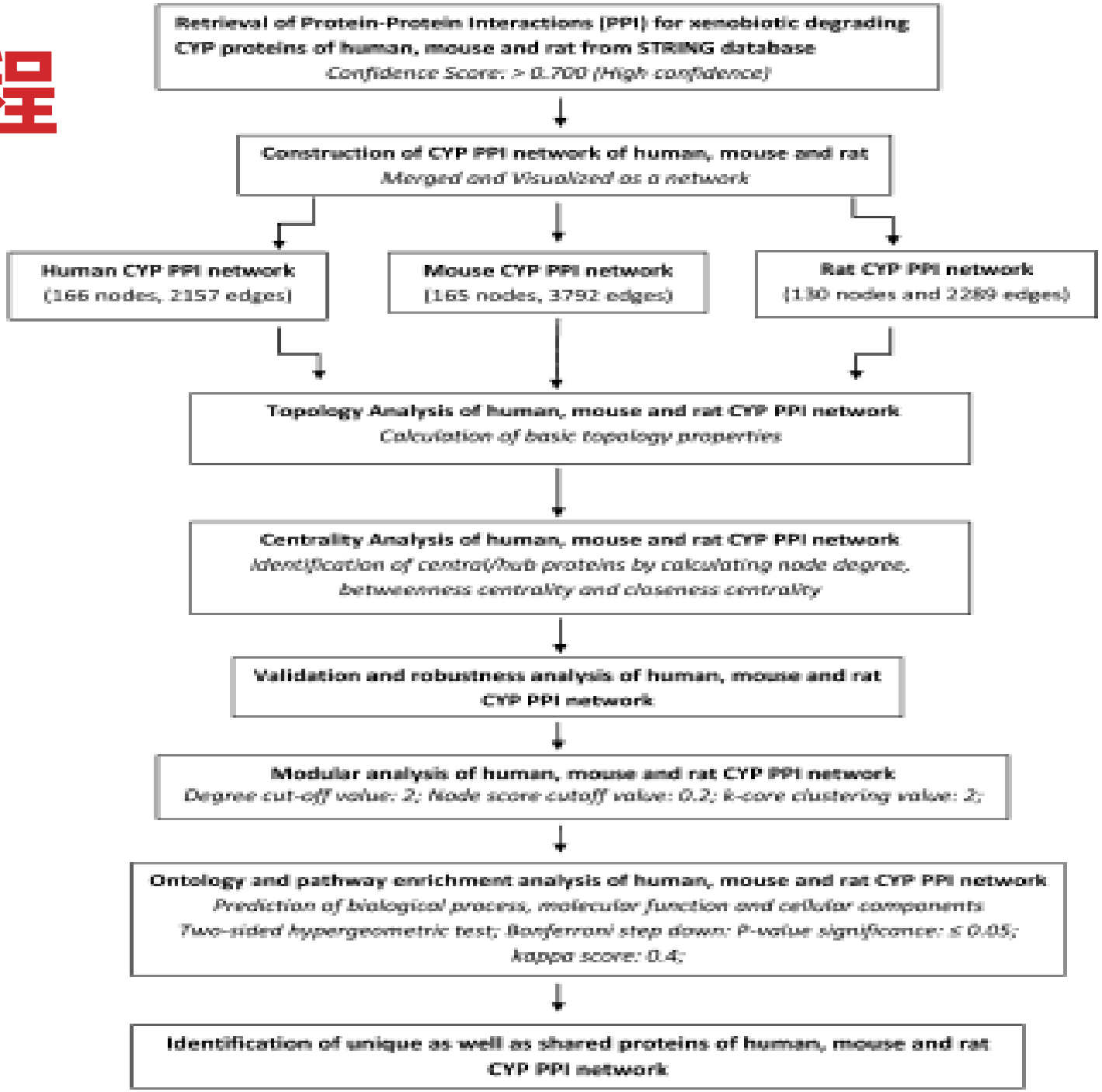
- 演讲：管玮，王景
- 翻译/PPT制作：李蕾，管玮，王景

目的

1.使用网络理论途径研究了解人类，小鼠和大鼠中降解异源物质的CYP（细胞色素P450蛋白酶）的组织结构和功能。

2.在整体水平上研究上述三个物种之间存在的差异。

流程


Retrieval of Protein-Protein Interactions (PPI) for xenobiotic degrading CYP proteins of human, mouse and rat from STRING database
Confidence Score: > 0.700 (High confidence)
Construction of CYP PPI network of human, mouse and rat
Merged and Visualized as a network
Human CYP PPI network
(166 nodes, 2157 edges)
Mouse CYP PPI network
(165 nodes, 3792 edges)
Rat CYP PPI network
(130 nodes and 2289 edges)
Topology Analysis of human, mouse and rat CYP PPI network
Calculation of basic topology properties
Centrality Analysis of human, mouse and rat CYP PPI network
Identification of central/hub proteins by calculating node degree, betweenness centrality and closeness centrality
Validation and robustness analysis of human, mouse and rat CYP PPI network
Modular analysis of human, mouse and rat CYP PPI network
Degree cut-off value: 2; Node score cutoff value: 0.2; k-core clustering value: 2;
Ontology and pathway enrichment analysis of human, mouse and rat CYP PPI network
Prediction of biological process, molecular function and cellular components
Two-sided hypergeometric test; Bonferroni step down; P-value significance: ≤ 0.05;
kappa score: 0.4;
Identification of unique as well as shared proteins of human, mouse and rat CYP PPI network

STEP 1

从STRING数据库检索人类、小鼠和大鼠的CYP异源物质降解蛋白的PPI

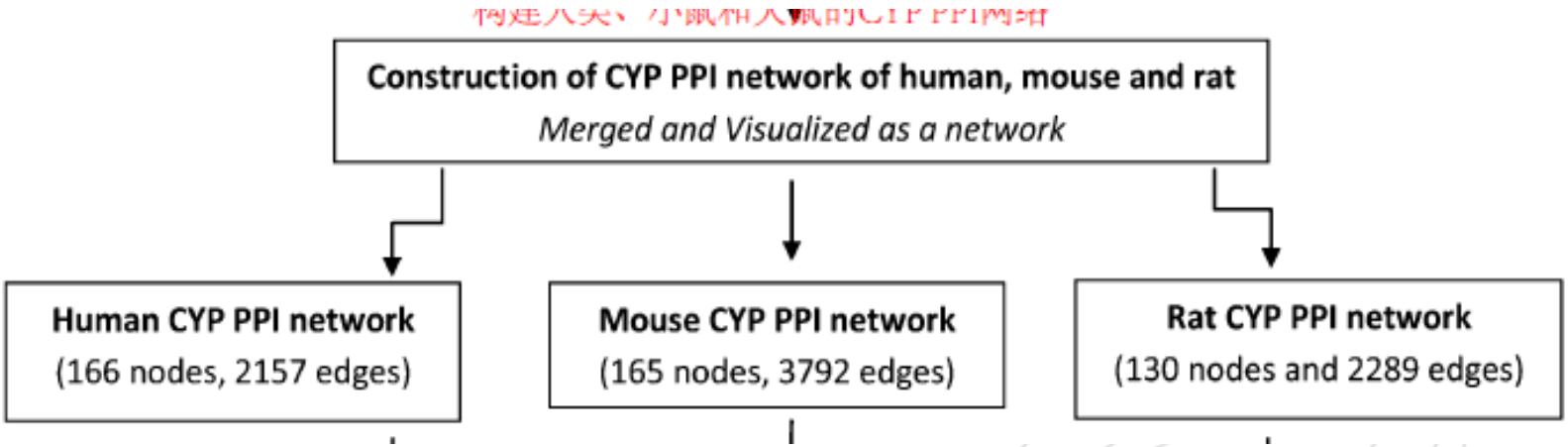
Retrieval of Protein-Protein Interactions (PPI) for xenobiotic degrading CYP proteins of human, mouse and rat from STRING database

Confidence Score: > 0.700 (High confidence)

缺点：

STRING is a meta-database that may contain false positives and false negatives.

STEP 2



使用了Cytoscape version 3.2.0软件构建人类，小鼠和大鼠的蛋白质互作网络

Table 1 Topological parameters of the CYP PPI networks of human, mouse and rat

Parameters	Human	Mouse	Rat
Number of nodes	166	165	130
Number of edges	2157	3792	2289
Average clustering coefficient	0.725	0.693	0.707
Average number of neighbors (or) mean degree	25.988	45.952	35.215
Shortest paths	27 390 (100%)	27 060 (100%)	16 770 (100%)
Characteristic path length	2.095	1.937	1.879
Network diameter	4	5	4
Network radius	2	3	2
Network centralization	0.393	0.358	0.423
Network density	0.158	0.280	0.273
Network heterogeneity	0.720	0.645	0.668
Connected components	1	1	1
Isolated nodes	0	0	0
Number of self-loops	0	0	0
Multi-edge node pairs	0	1	0

STEP 3

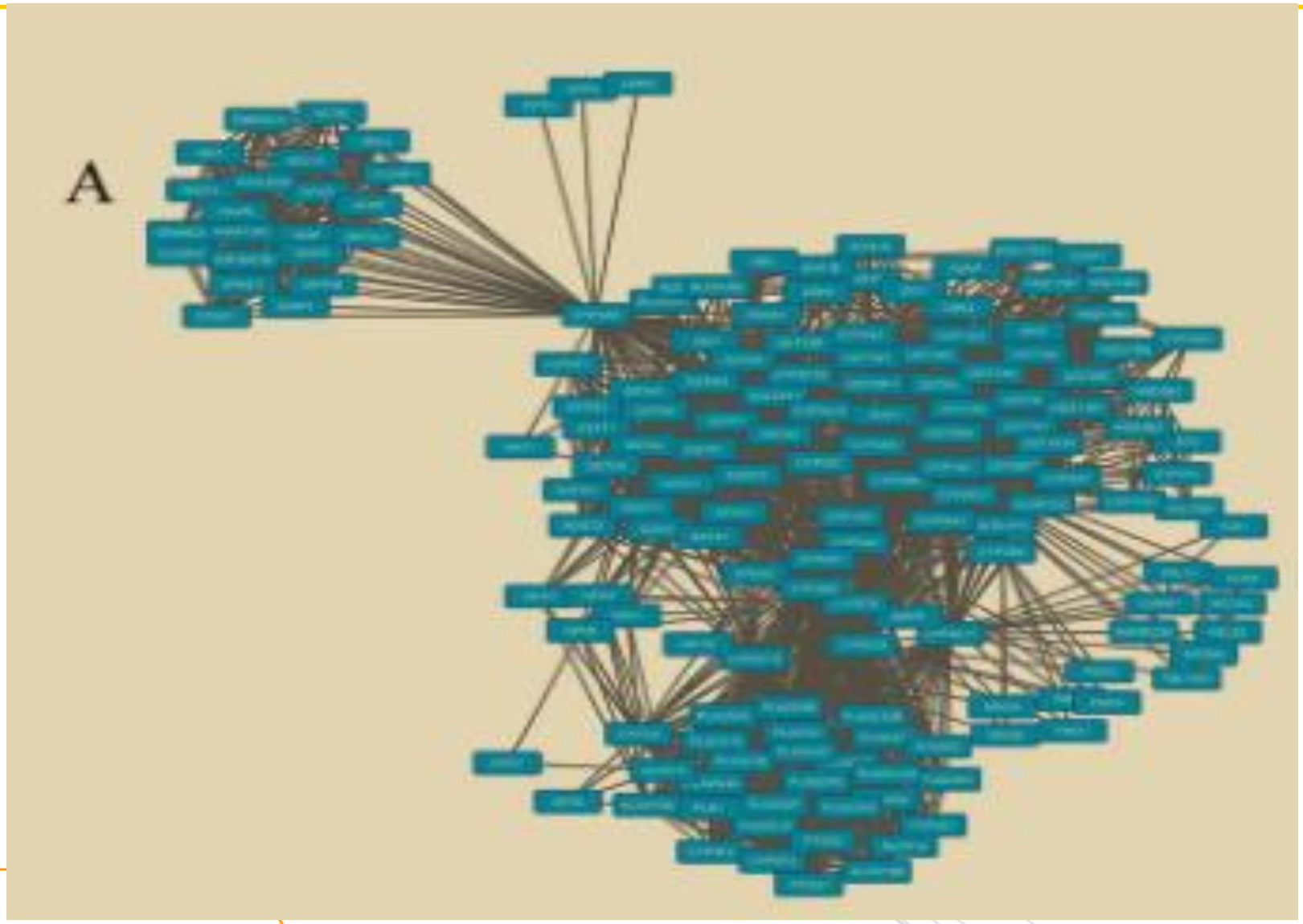
Topology Analysis of human, mouse and rat CYP PPI network

Calculation of basic topology properties

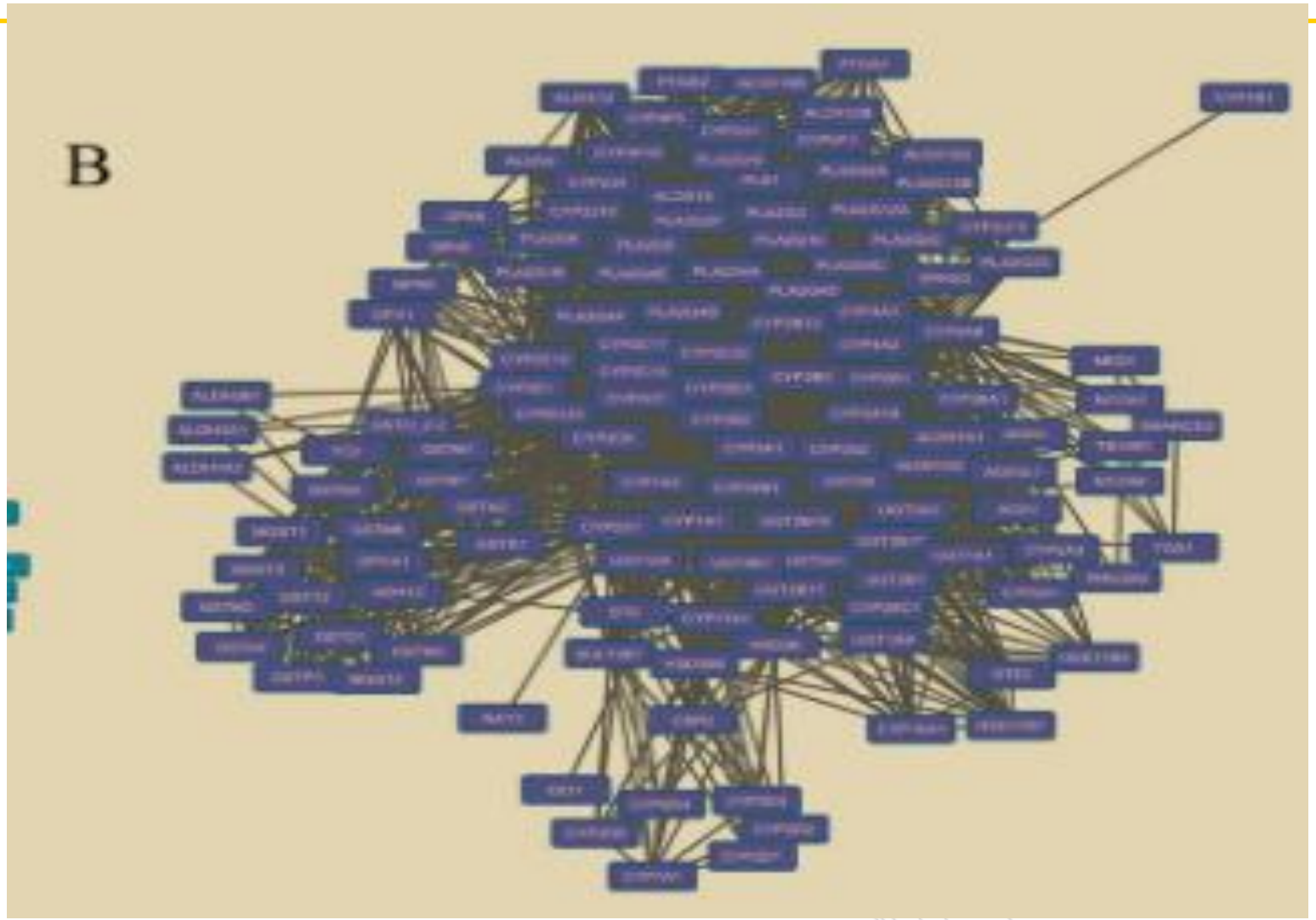
1.计算了聚类系数C(n): $$C(n) = \frac{2e_n}{(k_n(k_n - 1))}$$

2.计算了三个物种的度分布
3.使用网络展示三个物种的拓扑属性（见文章）
4.计算了路径长度分布

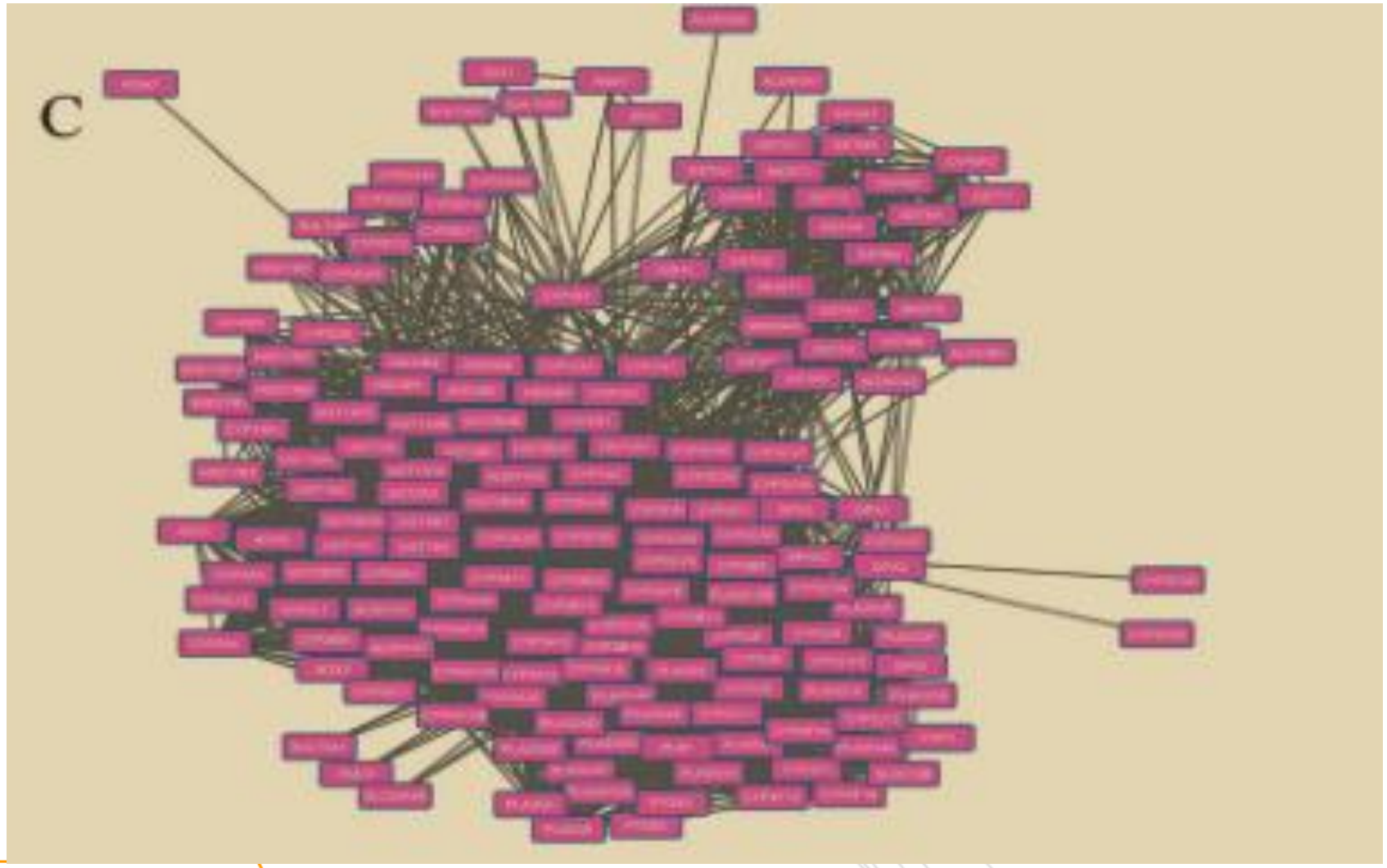
人类CYP PPI 网络（独立节点均已移除）



小鼠CYP PPI 网络



大鼠CYP PPI 网络



作用：

生物网络的拓扑和中心分析有助于揭示细胞系统的功能组织和他们的基本设计原则。

它有助于在全球网络结构的特征使用定量措施(网络拓扑和中心统计数据),进一步帮助在目标选择,识别重要的节点以及无标度性质的表征。

STEP 4

Centrality Analysis of human, mouse and rat CYP PPI network

Identification of central/hub proteins by calculating node degree, betweenness centrality and closeness centrality

1.介度中心性BC(n)

$$\mathrm{BC}(n) = \sum_{s \neq n \neq t} \left(\frac{\sigma_{st}(n)}{\sigma_{st}} \right),$$

2.接近中心性CC(n)

$$\mathrm{CC}(n) = \frac{1}{\mathrm{avg}(L(n,m))},$$

以人类为例

Table 2 Network centrality parameters of hub proteins of the human CYP PPI network

S. no.	Protein	Degree	Betweenness centrality BC(n)	Closeness centrality CC(n)
1	CYP3A4	90	0.08760567	0.68464730
2	CYP1A2	86	0.07765945	0.67346939
3	CYP2E1	86	0.07459228	0.67346939
4	CYP1A1	83	0.09792858	0.66532258
5	CYP2B6	83	0.04589389	0.66532258
6	CYP2A6	79	0.27210067	0.64960630
7	CYP2C8	77	0.03946531	0.59139785
8	CYP2C9	76	0.04954342	0.64705882
9	CYP4A11	71	0.07355107	0.63218391
10	CYP3A5	61	0.01297470	0.61111111
11	CYP2C19	54	0.01527446	0.54098361
12	CYP1B1	48	0.01299541	0.51562500
13	CYP2C18	47	0.01551213	0.53225806

STEP 3/4

Topology and centrality analyses of our predicted CYP PPI networks for human, mouse and rat were performed to identify the central proteins in the network,as well as to elucidate the internal organization of the network using various properties such as degree, diameter, shortest path, average clustering coefficient, betweenness centrality and closeness centrality

STEP 5

Validation and robustness analysis of human, mouse and rat CYP PPI network

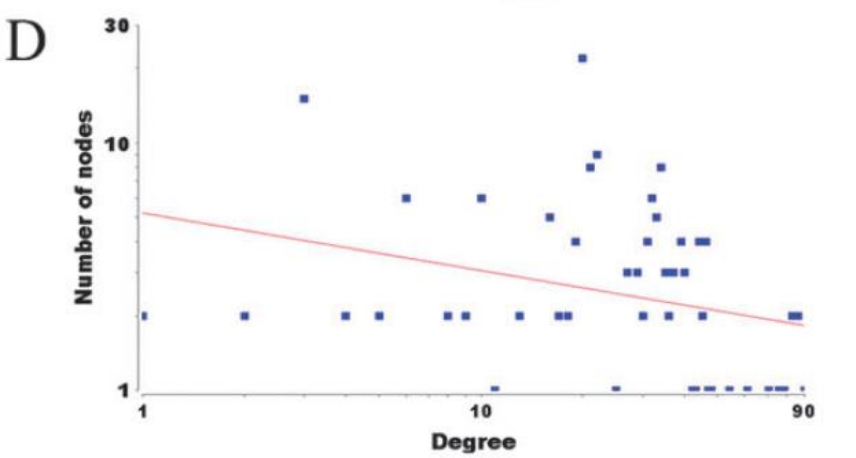
作用：

验证和描述网络的无尺度特性

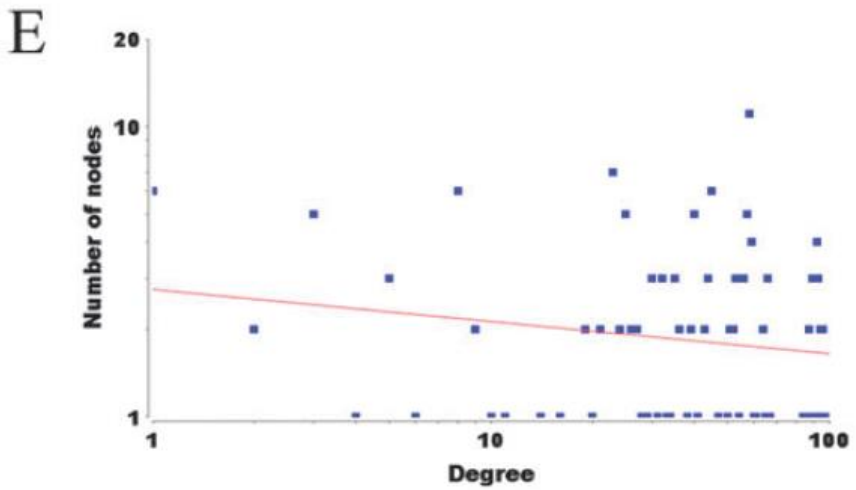
确定hub蛋白在维持网络连通性中的作用

使用Cytoscape的Random Network插件进行验证和鲁棒性分析

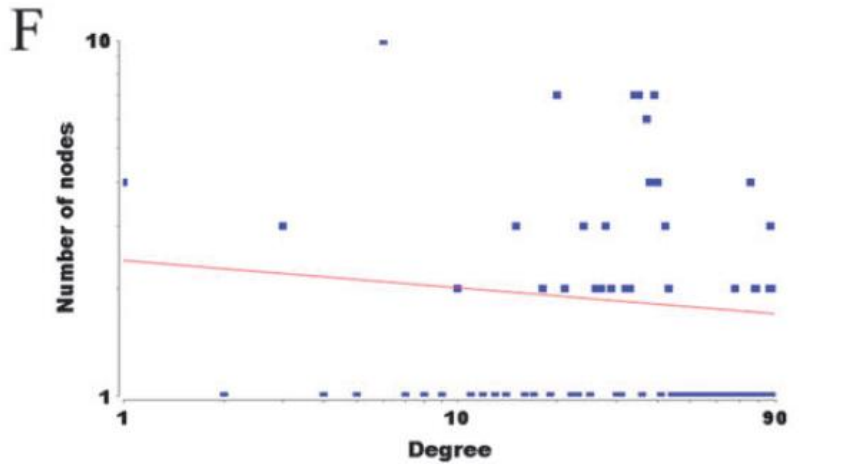
$$Y=a*X^{b}$$



(D) human
y = (5.243)x0.234
(R2 = 0.094)



(E) Mouse
y = (2.724)x0.107
(R2 = 0.028)



(F) Rat
y = (2.405)x0.076
(R2 = 0.011)

Table 5 The average clustering coefficient of the predicted CYP PPI networks and the randomized CYP PPI networks of human, mouse and rat

Organism	Mean degree	Average clustering coefficient of the predicted PPI network	Average clustering coefficient and standard deviation (σ) of random networks
Human	25.988	0.725	0.160 ± 0.0027873
Mouse	45.952	0.693	0.282 ± 0.0018963
Rat	35.215	0.707	0.276 ± 0.0028759

This result indicates that associations among proteins of our predicted human, mouse and rat CYP PPI networks are not random, but biological and scale-free. Hence, our human, mouse and rat CYP PPI networks possess a modular architecture.

这个结果表明蛋白之间的关联在我们预测的人类、小鼠和大鼠**CYP PPI**网络中并不是随机的,而是生物和无标度得。因此,我们人类、小鼠和大鼠**CYP PPI**网络拥有一个模块化的架构。

HUB蛋白的两种分析方法

- The first method is the complete removal of hub proteins

结果：导致聚集系数的平均值减小，节点的特征路径长度增长，网络的结构和连通性遭到破坏。

- The second method is the step-by-step removal of hub proteins in the order of highest to lowest degree

结果：导致聚集系数的值逐步减小，节点的特征路径长度逐步增长

从这个分析推断,人类、小鼠和大鼠CYP PPI网络的鲁棒性会维持着，即使移除一个或两个中心蛋白质

STEP 6

Modular analysis of human, mouse and rat CYP PPI network

Degree cut-off value: 2; Node score cutoff value: 0.2; k-core clustering value: 2;

作用：

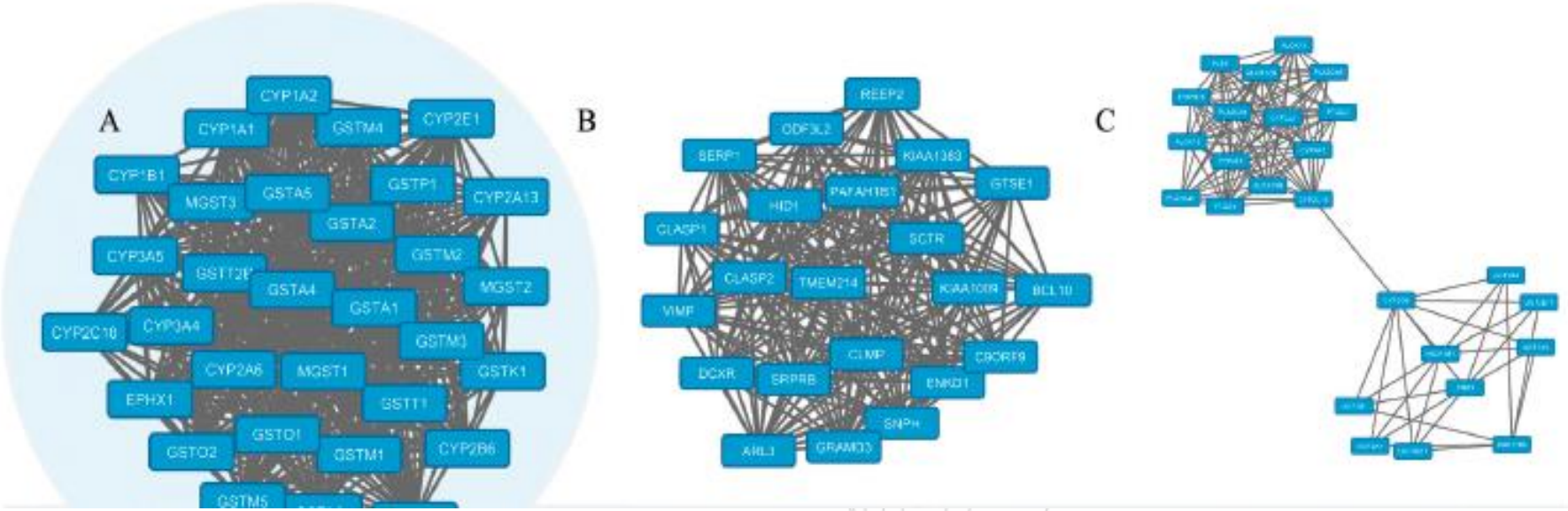
1、eliminate noise；
2、identify the cluster of proteins that are functionally related；
3、identify functionally important proteins that cluster together.

使用Cytoscape的MCODE插件进行模块化分析，选择默认参数：

- 度的截断值是2；
- 节点评分截断值是0.2
- k-core集群值是2。

前三个集群基于聚类分数,节点数量和边数目，从结果中选出the top scoring sub-network 。

人类网络的结果


A
CYP1A2
CYP1A1
GSTM4
CYP2E1
CYP1B1
MGST3
GSTA5
GSTP1
CYP2A13
GSTA2
CYP3A5
GSTT2P
GSTM2
GSTA4
MGST2
GSTA1
CYP2C18
CYP3A4
GSTM3
CYP2A6
MGST1
GSTK1
EPHX1
GSTT1
GSTO1
GSTO2
GSTM1
CYP2B6
GSTM5
B
REEP2
ODF3L2
SERPI1
KIAA1363
GTSE1
PAFAH1B1
HID1
CLASP1
SCTR
CLASP2
TMEM214
KIAA1009
BCL10
VIMP
CLMP
C8ORF8
DCXR
SRPRB
ENKD1
SNPH
ARL3
GRAMD3
C

STEP 7

Ontology and pathway enrichment analysis of human, mouse and rat CYP PPI network

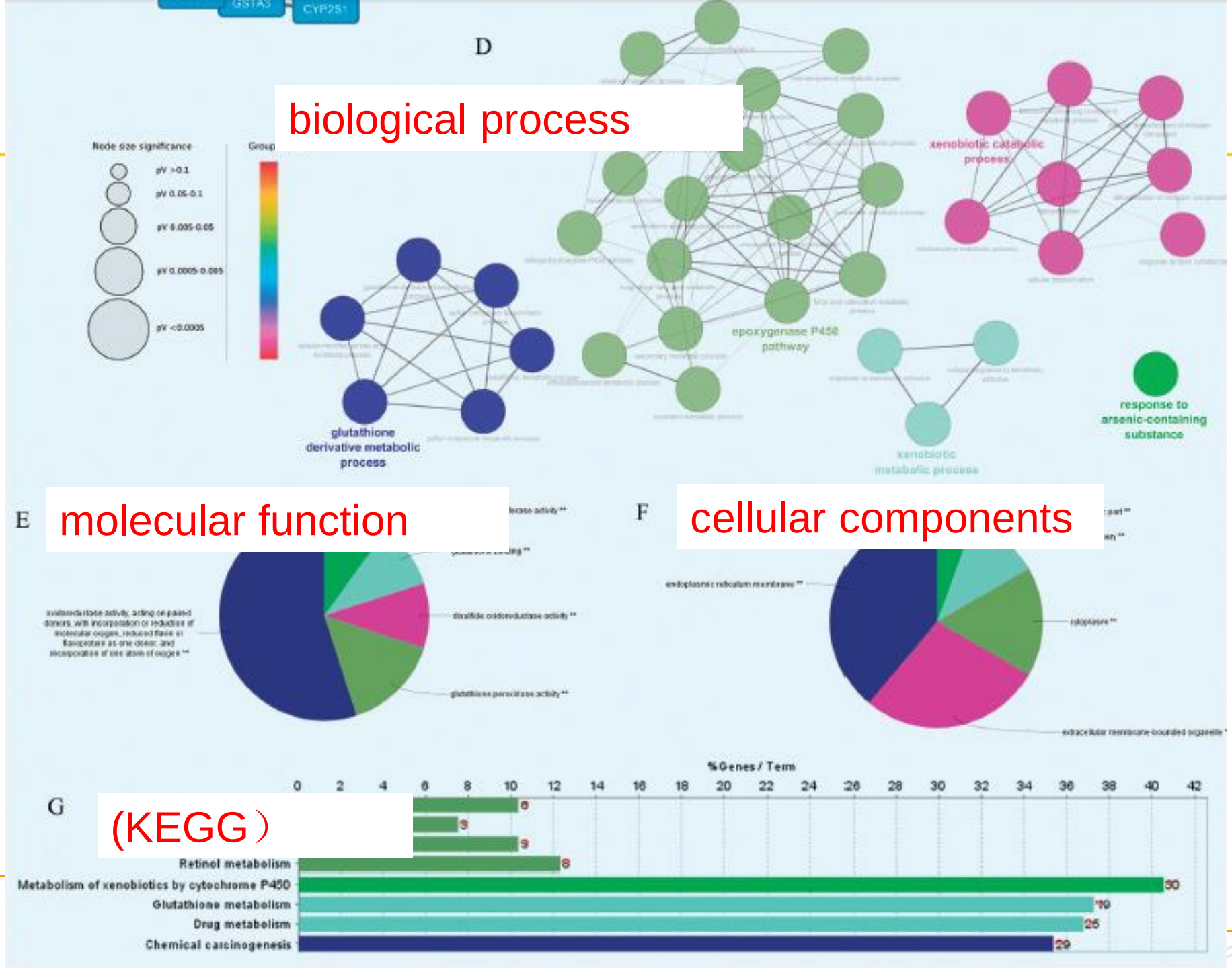
Prediction of biological process, molecular function and cellular components

Two-sided hypergeometric test; Bonferroni step down: P-value significance: ≤ 0.05;

kappa score: 0.4;

作用： Ontology and pathway analyses of these clusters are helpful in predicting gene ontology (GO) terms such as biological process,molecular function, cellular components and Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathways that are overrepresented in the form of groups.

使用Cytoscape的ClueGO插件进行GO和通道富集分析


GSTA3
CYP25↑
D
biological process
Node size significance
pV >0.1
pV 0.05-0.1
pV 0.005-0.05
pV 0.0005-0.005
pV <0.0005
Group
glutathione derivative metabolic process
epoxygenase P450 pathway
xenobiotic catabolic process
xenobiotic metabolic process
response to arsenic-containing substance
E
molecular function
F
cellular components
G
(KEGG）
%Genes / Term
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42
Retinol metabolism
Metabolism of xenobiotics by cytochrome P450
Glutathione metabolism
Drug metabolism
Chemical carcinogenesis
30
29

STEP 8

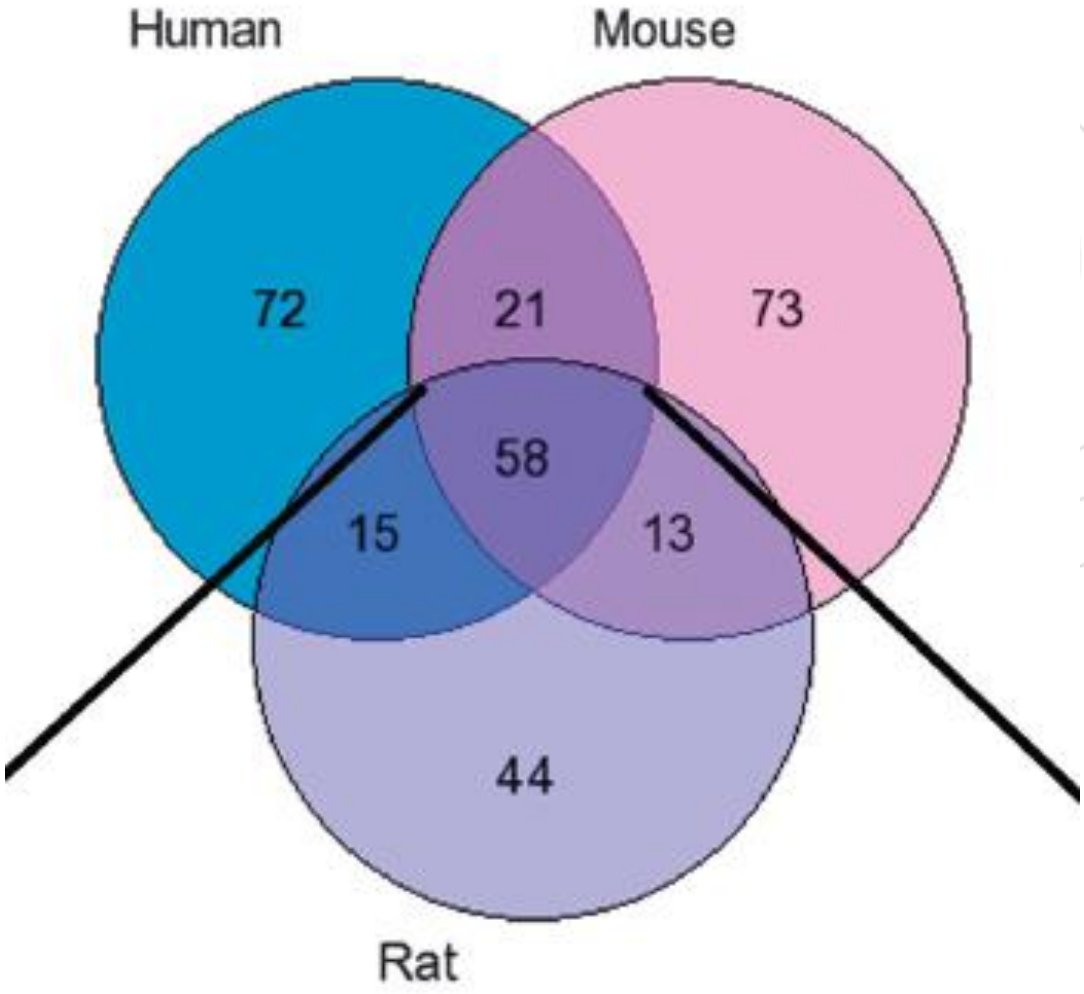
Identification of unique as well as shared proteins of human, mouse and rat CYP PPI network

使用的工具： the GeneVenn web application

使用GeneVenn web应用程序进行鉴别：

- 在这三个生物中重叠的蛋白质；
- 在两个物种中重叠的蛋白质；
- 在其他物种中不存在的非重叠蛋白质。

STEP 8 结果



结论

1、拓扑分析结果显示三个PPI网络都是生物的、无尺度的和强健的网络（不是随机的网络）；

2、三个网络都具有连通性，其中人类的CYP PPI网络中蛋白质的连通性是最高的；

3、人类的CYP PPI网络的集群高度富集在与异源物质/药物代谢关联的途径上；

4、小鼠，大鼠的CYP PPI网络富集的途径被内源信号通路控制。

总之

这些发现：

1、有助于在整体水平上理解人类、小鼠和大鼠CYP酶的相互作用网络和功能性结构组织。

2、此外,在系统水平上跨物种比较人类、小鼠和大鼠CYP PPI网络，揭示了物种之间的串扰和差异。

3、对于CYP的代谢和物种外推的更深层次的理解,简化方法和整体方法的联合是必需的。

THANKS!