

Constructing transcriptional regulatory networks

王 开：文献翻译，PPT展示
李心舒：文献翻译，PPT制作
郑芮钦：文献翻译，PPT制作

CONTENTS

1

Abstract

2

Properties of
biological networks

3

Constructing
transcriptional
regulatory networks



4

elucidating
transcriptional
regulatory networks

5

Deconvoluting
transcriptional
regulatory networks

6

Applying complementary
approaches: regulatory
networks in mammalian cells

7

Future prospects



1. Abstract

主要内容：

1. 简要介绍转录调控网络的性质。
2. 介绍一些时兴的分析网络结构，描述其动态行为的方法。

*目前，主要采用结合ChIP，基因表达谱，还有计算模拟等方法来构建复杂的生物学过程的起始和维持方式的蓝图，包括细胞周期进展、生长停滞以及分化。而分子生物学、计算机生物学的最新进展可以为研究复杂调控转录网络、描述基因表达的调节输入的函数、研究特定蛋白质和脱氧核糖核酸之间的相互作用 等方面提供支持。



2. Properties of biological networks

2.1 biological network

1. 生物网络大多数是由节点和边组成的。

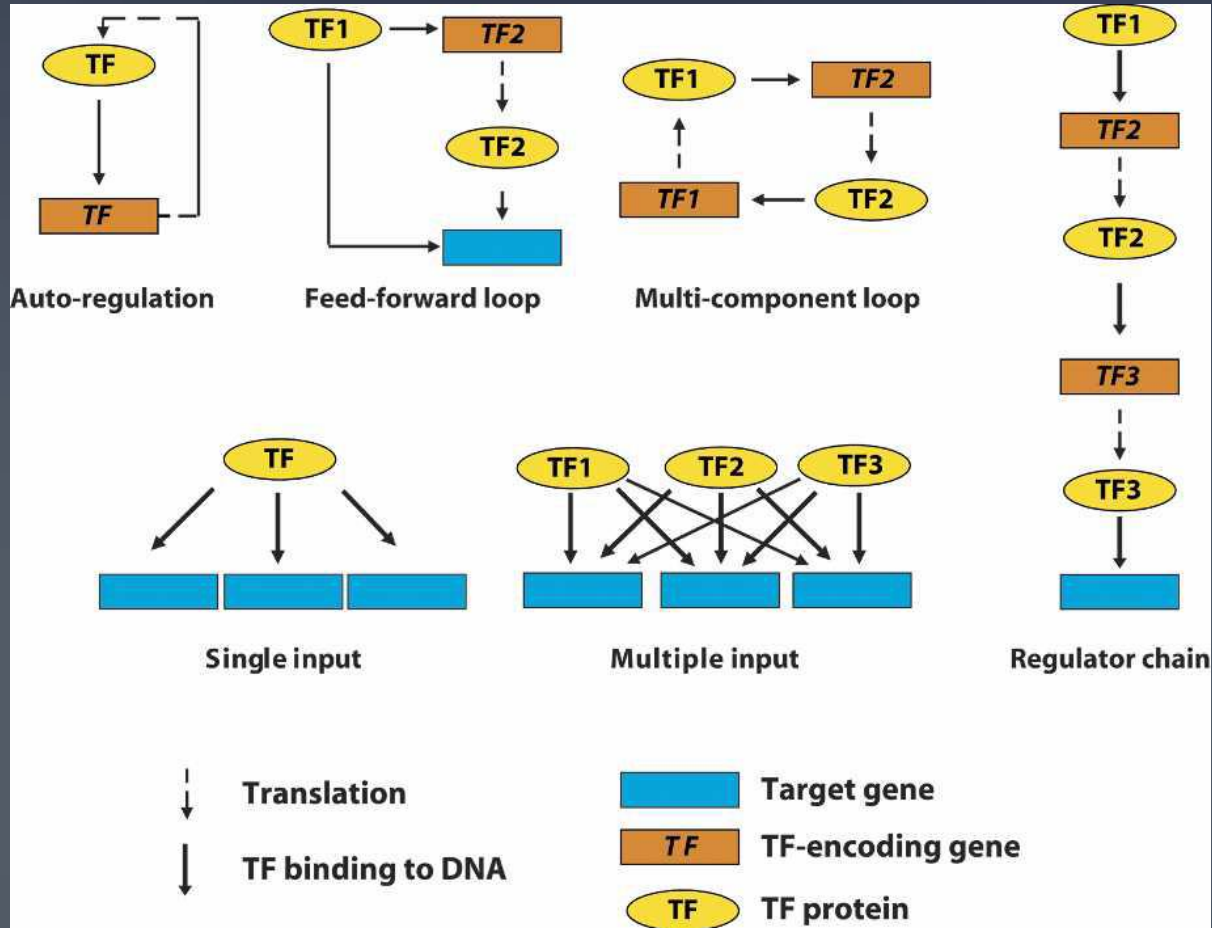
节点通常为蛋白质、基因、酶、传输外界信号的底物；

边通常是代表分子间互作、调控相互作用或共享的功能特性。

2. 生物网络的一个重要特点是其无标度结构。

与大量节点相连的中心点比相连关系少的节点的数目少。

2.2 Network motif



普通的网络更多用于贯彻宏观的所有情况下的调节，但是network motif更加注重在细节的模块中的研究。

这种模块的节点在功能上相连或者直接相连。节点可以表示共享一个共同的调节因子或在同一特定条件下表达的基因的集合。通过考虑其二级结构可以简化网络。

Network motif描述了节点之间是怎样连结的

2.3 Networks undergo condition-specific rewiring

在酵母中，使用一个通用网络的动力学参数在电脑上进行模拟，多数TF中心只在特定情况下活跃，少部分一直活跃。内源外源也被引入，用来描述在不同方面的调控过程中的网络组成

内源子网监管复杂固有的细胞活动（包括细胞周期和产孢），其特征为调控网络的多级结构。TF节点中心的靶基因数量少，往往趋于与其他TF局部相连。这些枢纽通常距离这些生物学过程的效应器比较远，并被一些节点隔开，这些性质说明生物学过程的复杂、耗时间长。同时外源子网对刺激回应速度更快，这种网络和较少的TF相连，但靶基因多，通常在细胞响应刺激时有“终端效应”

在很多不同的实验条件下，对酵母转录因子进行定位分析，分析目标基因在不同情况下的变化。

根据识别其目标的能力，TF分为四种：

1. 条件不敏感 任何条件 结合同目标基因
2. 条件激活 特定条件下 结合目标基因
3. 条件扩张 特定情况 结合多个目标基因
4. 条件绑定 特定条件 结合不同的目标基因



3. Constructing transcriptional regulatory networks

为了了解转录调控网络调节生物学过程，比如细胞周期或者分化的拓扑结构和动力学，对：

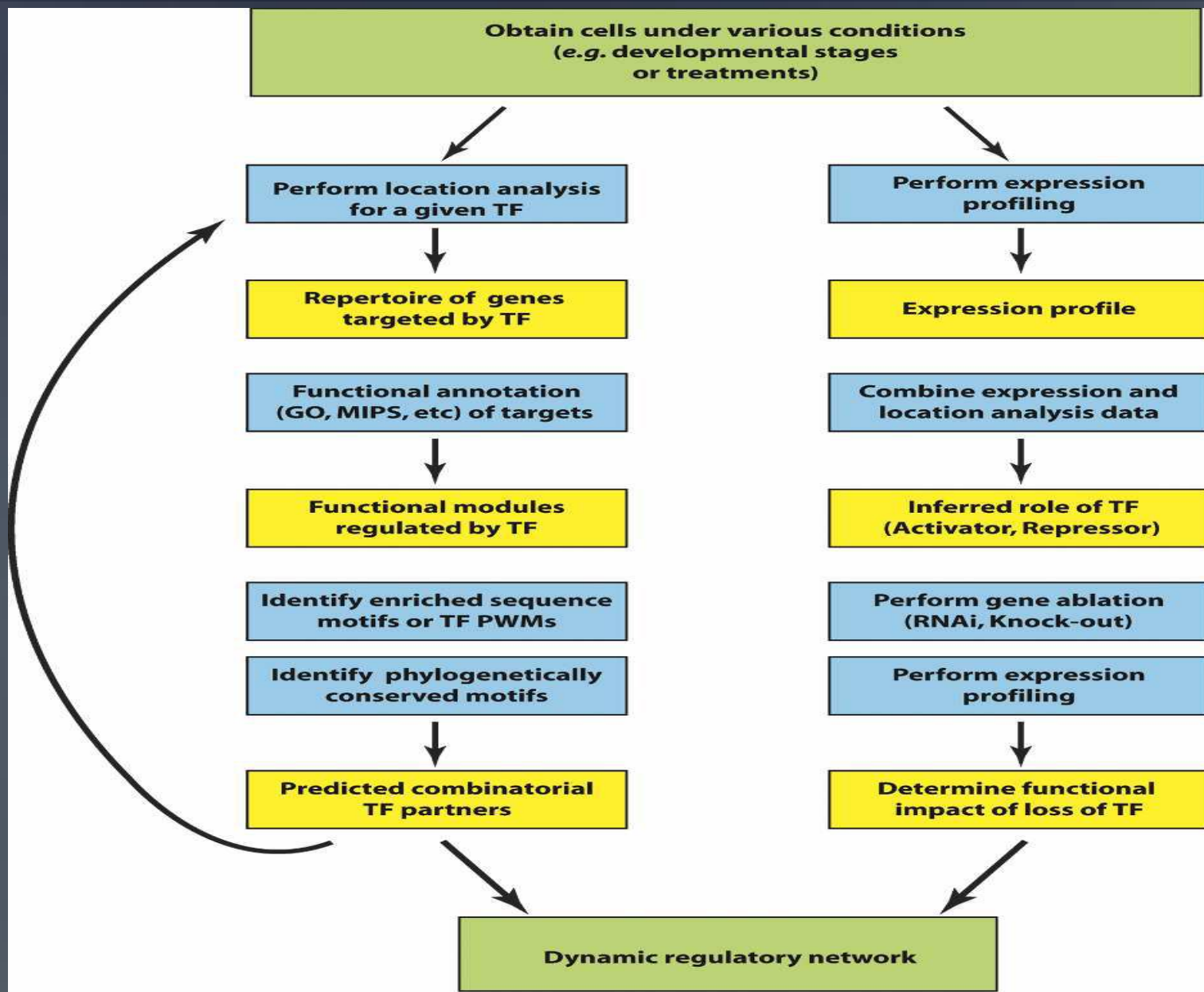
- (1) 互作结点的识别和表达水平
- (2) 互作随时间的改变而改变（例如，在细胞周期中或在分化过程中）
- (3) 破坏关键节点对表型的影响。

这三个方面进行了评估。

由于基因组学的进展，有两种方法已经在很大程度上解释调控网络：

1. 全基因组表达谱
2. 染色质免疫沉淀（芯片）与启动子 DNA 芯片
（称为CHIP-on-chip, CHIP chip或位置分析）的结合

这两种方法可在一组给定的条件下识别直接靶基因。第三种方法是全基因组 RNA 干扰 (RNAi) 筛选。



用互补的方法来解释转录调控网络。蓝色方框表明实验方法，黄色方框表明获得的信息。箭头直接指向表示执行这个实验的顺序。（PWMs）权重矩阵代表TF结合位点。



4. Elucidating transcriptional regulatory networks

4.1 CHIP-on-chip

即染色质免疫共沉淀（CHIP）与基因芯片方法相结合的技术

CHIP-on-chip分析具有一些特性功能：

1. 在其最简单的形式中，使用表达蛋白原水平的活细胞，而不是使用过度表达的活细胞，从而避免特异性的潜在损失。
2. 该方法侧重于调节器与目标之间的直接相互作用。这种特性很重要，因为它允许我们在TF和其目标之间确定中间体（中间节点）的数量。

但此方法也存在一定的弊端。CHIP-on-chip也被限制，如：

1. 抗体能否结合在抗原表位，以及在实际实验中出现的否定的结果，通常是不可解释的。
2. TF在目标启动子上的精确定位的知识没有提供关于它的功能的信息。

 **Table 1.** *Comparison of different methods aimed at identifying genomic targets*

Method	Benefits	Caveats	References
ChIP on CpG island arrays	Microarrays are easy to produce	Targets often ill-defined; poor annotation; limited by antibody performance/availability; low resolution; somewhat biased	(Weinmann et al. 2002; Wells et al. 2003; Kondo et al. 2004)
ChIP on proximal promoter arrays	Straightforward	Limited by antibody performance/availability; low resolution; somewhat biased	(Ren et al. 2000, 2002; Iyer et al. 2001; Cam et al. 2004; Odom et al. 2004)
ChIP on genome tiling arrays	High resolution; less biased	High cost; requires more material or more amplification	(Cawley et al. 2004)
DamID and biotinylation methods	Circumvents the need for an antibody	Exogenously expressed factor may not behave exactly like the endogenous factor	(van Steensel and Henikoff 2000; Orian et al. 2003; Bianchi-Frias et al. 2004; Viens et al. 2004)
STAGE or SABE	Circumvents the need for a microarray; unbiased	Subtractive hybridization may be necessary; high-throughput sequencing needed	(Kim et al. 2004; Chen and Sadowski 2005)

识别基因组 TF 结合位点的几种方法

4.2 Attention

给定的 TF 的位点的信息不提供在主导条件下有关实际上因子是否调节附近基因的信息。有许多例子表明TF的募集与其靶基因转录状态（即，感应或抑制）无关。这种现象可以由辅激活子或辅阻遏物蛋白的补充额外的TFs或者组合的调控解释。因此，额外的功能分析总是需要补足CHIP-on-chip的数据，从而达到对调控网络有更精确的描述。

4.3 gene knock-outs and RNAi

基因敲除采取了两种方法来了解因子与其靶基因之间的功能上的相互作用：

1. 改变因子在靶启动子中的结合位点。

其优点是：1. 其他因子的靶基因保持完整

2. 不太可能在基因敲除中产生不良的次级效应

其缺点是：很难辨别直接效应和次级效应

2. 借助计算模拟。

其优点是不需要辨别直接效应和次级效应，只需要通路中假定的TFs被依次切除。

RNAi在以下几点优于基因敲除：

1. 对阐明大量的调控网络的高产出量的筛选的适应性

2. RNAi的敏锐切除可以避免互补的机制。

4.4 Computational approaches

1. 两个解释调控网络的基于DNA测序的方法：

1.1 依赖于已知的TF结合位点选择参数的信息

1.2 不依赖于已知的TF结合位点选择参数的信息

2. De novo 的Motif查找算法

发现因子结合位点的方法尝试识别出出现在一组启动子中的短序列。这个方法不依赖于已知的TF结合位点的序列，其中包含了Align ACE, MDScan, MEME, REDUCE算法。

3. 用基因表达数据来将基因分成不同的共表达模块，并假定这些模块的调节器也被转录调控。

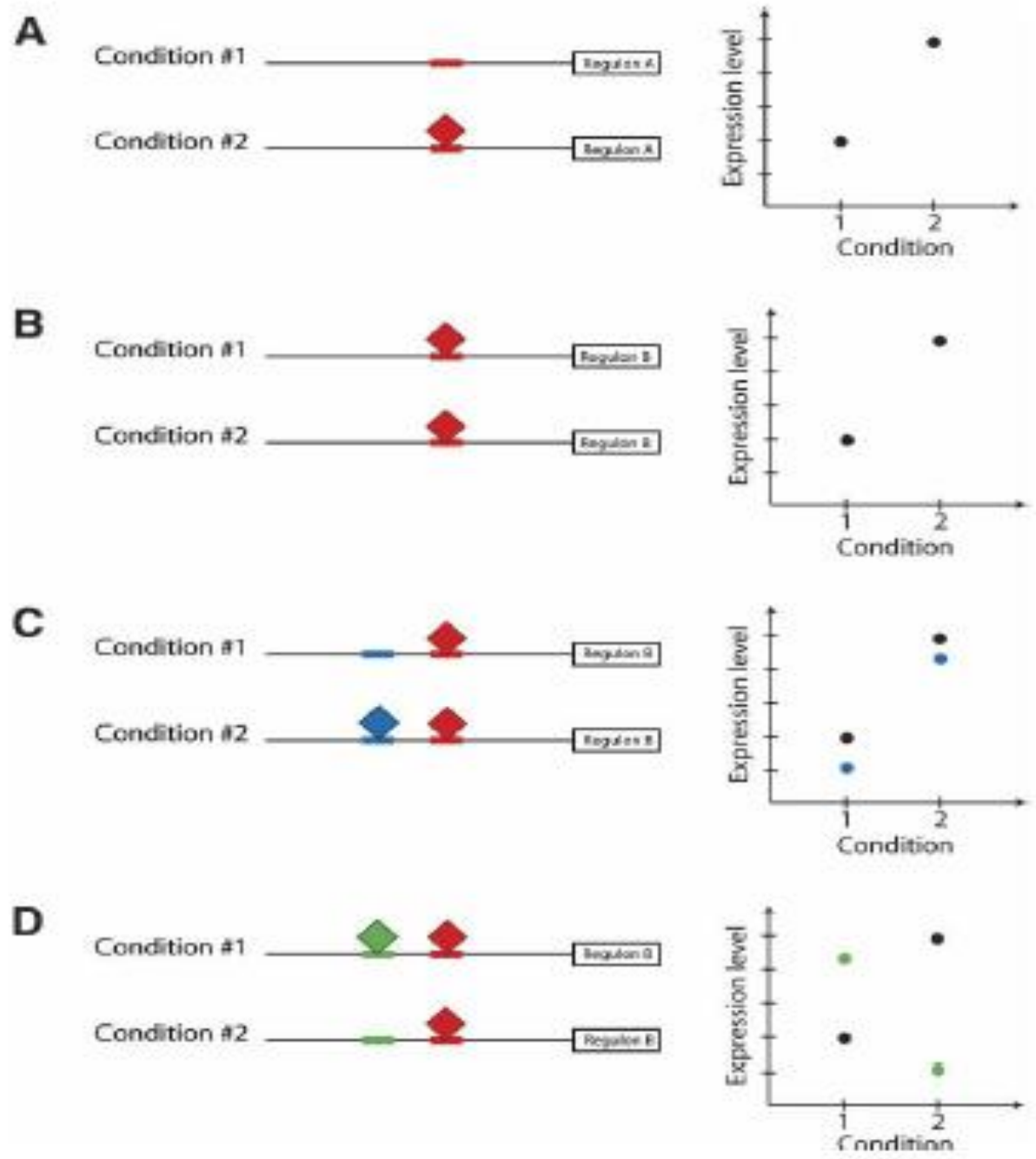
此方法从大量的表达谱数据中将基因与其对应的调节子分组，找到过表达序列，并使用贝叶斯网络来推断表达谱与序列之间的关系。

4.5 Integrated approaches

一些可以优化这两种基于测序的计算模拟，并提高精确度的方法：

1. 几个TFs联合作用；
2. 考虑这些结合位点中系统发生的可能的保守位点；
3. CHIP-on-chip与基于测序的方法相结合。

此外，对限制了特定的因子的启动子的精确分析，可能揭示出额外的TFs的结合位点的存在，从而表明组合调控合作模式。



CHIP-on-chip实验（左）和
表达谱（右）并行实验



5. Deconvoluting transcriptional regulatory networks

细胞周期的转录控制是受多种因素共同作用的结果，并用一种概率算法将基因分组，这种算法基于其表达模式和它们一组因子的共同结合位点。这允许多输入motif精细识别来共表达（MIM-CE）。

这些motif形成在功能上互相链接的小的调控单元，而且在转录网络上加一个循环结构更能反映出细胞周期，主要是因为：

1. 一些TF家族在不同细胞周期阶段调控；
2. 一些TF家族是被在细胞周期中早些表达的因子调控的。

基于MIM-CE方法的GRAM算法被用于发现酵母基因组中的基因模块，并构建一个更大的基因调控网络

全基因组表达谱和位点分析的一个主要缺点是：它们不能够进行多细胞生物的实验。

但在单细胞中却可以绘制出复杂的调控网络

运用了多个工具来识别TF家族与其靶基因的结合位点，

并估算这个结合位点的重要性：

1. 大规模的扰动分析
2. 利用系统发育的保守结合位点的预测和干扰
3. 使用定量化PT-PCR或者cDNA巨型阵列与消减杂交的耦合



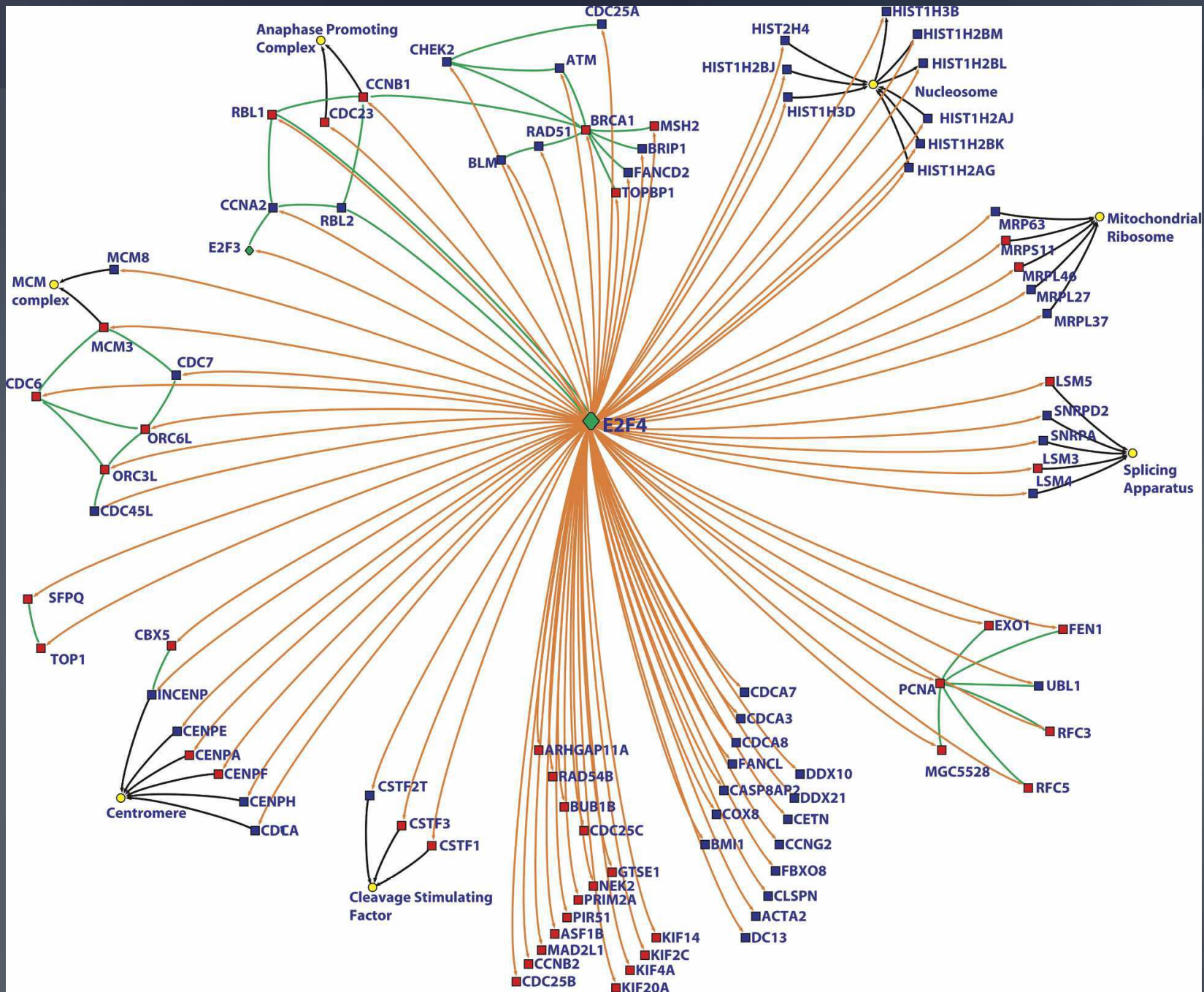
6. Applying complementary approaches: regulatory networks in mammalian cells

6.1 Transcriptional control of cell cycle progression

在哺乳动物细胞中，由E2F和视网膜母细胞瘤TF蛋白家族（pRB和相关p107和p130多肽，统称为口袋蛋白质）强迫转录控制在细胞周期进展中起主要作用。然而，对有关pRB和E2F，以及处理细胞周期阻滞仍缺乏基因调控机制的全面了解。

结合基因表达谱和计算模拟的对上述问题的解决方法如下：

- （1）了解在E2F，Pocket蛋白质，及在一个基因组中它们的靶基因之间的函数关系式
- （2）进一步阐明在细胞周期中，由E2F和Pocket蛋白质控制的网络。



构建的一个“细胞周期阻遏模块”，它在转录调控网络中是一个必要的元件，用于促进或者维持细胞循环的退出。



7.Future prospects

在构建转录调控网络中，提升实验方法，利用计算模拟，有全基因组位点分析结果的数据库等等各种方法的结合可以提高构建的速率与精确度，并且也有可能阐明更加复杂的转录调控网络。

在以后的研究中，调控网络需要解释在基因表达、蛋白蛋白互作和细胞内间隔中的时间变化。这些三维重建无疑是非常复杂的。但越复杂，就越接近它们在活细胞中发生的动态变化。

[THANKS]