

東京大学大学院新領域創成科学研究科
情報生命科学専攻

平成 17 年度

修士論文

キメラ化 Sip4 転写因子による下流標的遺伝子の網羅的同定

2006 年 3 月提出
指導教員 伊藤 隆司 教授

46909 谷 祐治

「目的」

出芽酵母のゲノム解読の結果、多数の新規転写因子(候補)が発見された。これらの転写因子の機能解析の鍵は標的遺伝子群の網羅的同定であり、転写因子破壊株のトランスクリプトーム解析やゲノムワイドのクロマチン免疫沈降(ChIP)解析が行われている。しかし転写因子の多くはその機能発現(核移行、DNA 結合能の獲得、転写活性化)に、様々な生育条件等を反映した上流からのシグナル伝達を必要とする。したがって、活性化刺激を先験的に知ることが出来ない新規転写因子では、破壊株や単純な過剰発現株の遺伝子発現解析や ChIP 解析では標的遺伝子を捕らえることが困難である場合も少なくない。

そこで私は、転写因子のキメラ化による恒常的活性化によって発現が誘導される遺伝子の網羅的検索に基づいて標的遺伝子を同定する戦略に着目し、その効率化を目指している。

「方法・結果」

出芽酵母の転写因子の中で最も大きなファミリーを形成し、しかもドメイン構造と機能の関連が解明されているZn₂Cys₆型転写因子を対象に、そのDNA結合ドメインをVP16 転写活性化ドメインとキメラ化することで恒常活性化する方法を用いた(図1)。今回は、このファミリーのうちSip4 を対象に取り上げた。この転写因子は遺伝子破壊では表現型の変化が見られないが、配列の相同性より非発酵性炭素源条件下で機能する転写因子Cat8 と同様の機能を持つと予想されている。実際、Sip4 はCat8 と同様にいくつかの遺伝子のCSRE (Carbon Source-Responsive Element)に*in vitro*で結合し得ることが報告されている。しかし*in vivo*におけるSip4 の標的遺伝子やCat8 との機能分化などはまだよく理解されていない。

Sip4 の coiled-coil ドメインを COIL プログラムで予測し、それより下流領域を λ cl リプレッサーと VP16 で置換したキメラ転写因子を作成した。これを発現させた細胞を用いてマイクロアレイ実験を行ない、対照細胞に比較して発現が4倍以上上昇した遺伝子を42個見出した。これらの遺伝子群を GO term finder で解析すると、糖代謝関連の遺伝子が有意に濃縮されていることが分った。また、それらの上流配列を解析したところ 20 遺伝子が CSRE 様配列を持っていることも分った。

CSRE を持っていない 22 個に関しては二次的な標的、つまり Sip4 により発現が誘導される遺伝子により転写が活性化される遺伝子群もしくは更にその下流遺伝子、あるいは非特異的に発現が誘導された遺伝子であると考えられる。これらの遺伝子は転写因子の機能を正確に理解するうえでノイズとなる。

そこで直接の標的と副次的なものを識別する必要がある。ひとつの手段としてキメラ転写因子に対して ChIP-Chip を行う実験的なものであり、もうひとつのアプローチとしてバイオインフォマティクスのものである。後者に関して CSRE の有無と発現強度に単純な相関がないためこの実験データだけから直接の標的遺伝子を抽出することは不可能である。そこで発現量以外の情報を利用した Sip4 の直接標的の抽出を試みた。

まず、抽出した遺伝子群に対して種々のマイクロアレイデータに基づくクラスタリングを行い、標的遺伝子の濃縮を試みた。これは”同一転写因子により発現が制御される遺伝子群は様々な環境下で同一の挙動を示す” という仮定に基づいている。その結果、Sip4 の標的であれば持っていると考えられる CSRE を上流に含んでいる遺伝子の濃縮が見られた。(表 1,2)

次により正確に上流シスエレメントを決定するために統計学的に上流配列を解析するアルゴリズムに、生物学的知見を加味したプログラムの作成を行った

まず、発現量以外の情報としてシスエレメント候補配列の翻訳開始点からの位置情報を取り入れた。これは転写因子がシスエレメントに結合し、RNA ポリメラーゼと相互作用することにより転写活性を制御するというメカニズムを考慮すると、出芽酵母の転写因子の結合可能な領域が標的遺伝子に対して限定的であろうと考えたため

ある。また、同一配列を持つ遺伝子の発現パターンの類似度を発現量以外の情報として使用した。

その結果、生物学的知見を加味することによりコンセンサス配列抽出の精度が向上することが認められた。(図.3)

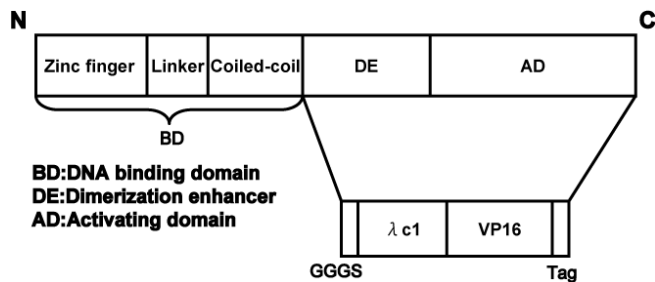


図 1. Zn₂Cys₆型転写因子の共通構造、及びキメラ化転写因子

Name	ID	Ratio	FBP1	YLR377C	7.00	YDR338W	YOR338W	4.90
PCK1	YKR097W	21.31	GRE2	YOL151W	7.00	BEM1	YBR200W	4.88
ICL1	YER059C	19.89	SPC1	YJR086W	8.78	YLR312C	YLR312C	4.84
YAT2	YER024W	18.75	MDH2	YOL126C	8.73	YGR050C	YGR050C	4.85
YGL157W	YGL157W	16.70	YDL241W	YDL241W	8.68	YDR044W	YDR044W	4.82
CAT2	YML042W	16.46	YLR004C	YLR004C	8.32	RK1	YOR095C	4.58
YAT1	YAR039W	12.00	AQR1	YNL065W	8.30	YLR311C	YLR311C	4.49
BTN2	YGR142W	10.10	YNR064C	YNR064C	8.08	YML131W	YML131W	4.34
YAR04C	YAR040C	8.58	MLS1	YNL117W	8.00	YML193W	YML193W	4.36
YLR154C	YLR154C	8.01	ATOS	YDR384C	5.93	YPL222W	YPL222W	4.28
YAR037W	YAR037W	7.95	SP1	YER150W	5.80	UBC8	YEL012W	4.25
CDA1	YLR307W	7.67	HXT12	YIL170W	5.50	YML005W	YML005W	4.20
YPL038W	YPL034W	7.60	ERG13	YML126C	5.40	YFR1	YDR368W	4.11
DCP1	YOL149W	7.21	YDR222W	YDR222W	5.25	SCL1	YNR034W	4.08
			CIT2	YCR005C	5.10			

Ratio: 対照遺伝子と比べた発現強度

表 1 キメラ化 Sip4 により発現が誘導された遺伝子

Name	ID	Ratio	YAR040C	YAR040C	30.46%	BEM1	YBR200W	23.07%
YAT2	YER024W	75.92%	BTN2	YGR142W	38.46%	YOR044W	YOR044W	23.07%
PCK1	YKR097W	69.23%	SP1	YER150W	38.46%	AQL1	YNL065W	23.07%
FBP1	YLR077C	69.23%	YAR037W	YAR037W	38.46%	TRM12	YML009W	23.07%
ICL1	YER069C	69.23%	YML131W	YML131W	38.46%	CDA1	YLR307W	15.38%
MLB1	YNL117W	61.53%	HXT12	YIL170W	38.46%	GRE2	YOL151W	15.38%
SPC1	YJR086W	61.53%	YDR222W	YDR222W	30.76%	YDL241W	YDL241W	15.38%
CAT2	YML042W	61.53%	YCR338W	YCR338W	30.76%	RK1	YOR095C	15.38%
MDM2	YOL126C	53.84%	ATOS	YDR384C	30.76%	ERG13	YML126C	15.38%
YAT1	YAR039W	46.15%	RH403	YLR154C	30.76%	YNR064C	YNR064C	15.38%
YLR311C	YLR311C	46.15%	YPL222W	YPL222W	30.76%	CIT2	YCR005C	7.69%
YML193W	YML193W	46.15%	DCP1	YOL149W	30.76%	YFR1	YDR368W	0%
YLR312C	YLR312C	46.15%	YPL038W	YPL038W	30.76%	SCL1	YNR034W	0%
YGR050C	YGR050C	38.46%	UBC8	YEL012W	30.76%			
YLR004C	YLR004C	38.46%	YGL157W	YGL157W	23.07%			

*Ratio: トップクラスターに選ばれる確率

表 2 クラスタリングを用いた標的遺伝子濃縮

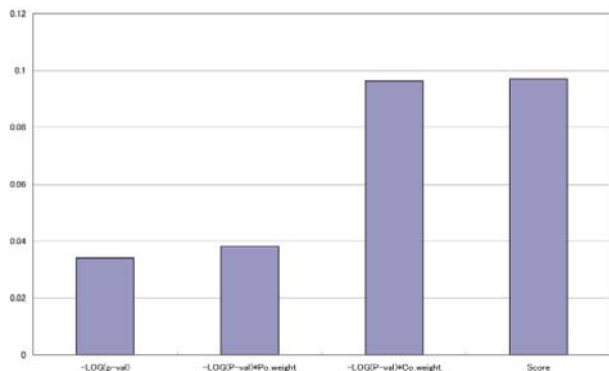


図.2 コンセンサス配列抽出精度

一番左は統計学的手法のみ、その右隣は翻訳開始点からの距離を重みとしている。さらにその右隣は遺伝子の発現パターンの類似度を重みとしている。右端は二つとも重みにしている。

キーワード: 酵母、転写因子、Sip4、キメラ化、コンセンサス配列、標的遺伝子