

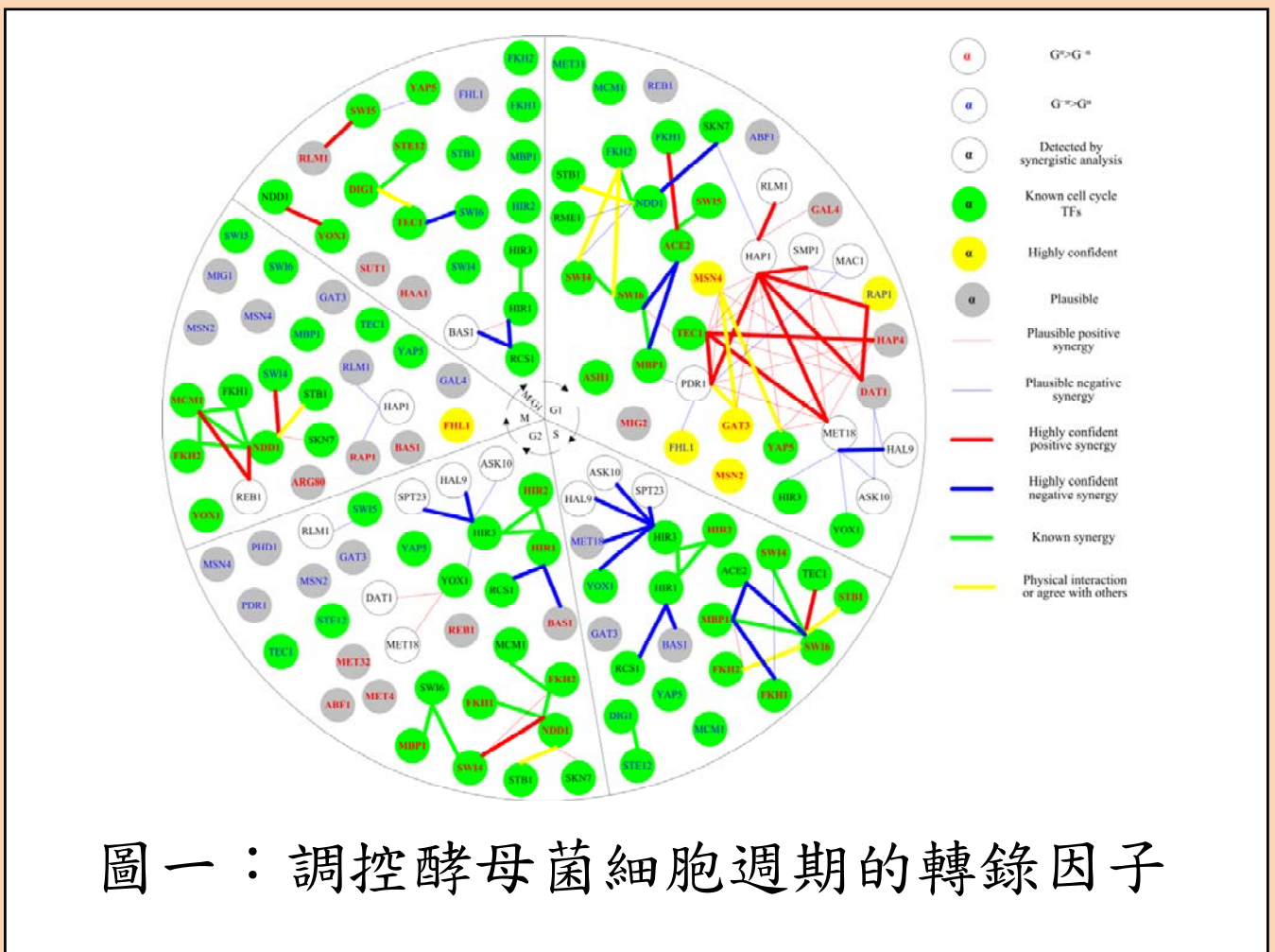
Bioinformatics approaches in studying the yeast regulatory mechanisms 酵母菌調控機制之生物資訊研究



蔡懷寬 副研究員
中央研究院資訊科學研究所

基因表現係由轉錄因子、染色質與具調控功能之RNA交互調控。其中轉錄因子扮演至為重要之角色，分析其功能與對應的結合位置將有助於瞭解基因轉錄的機制。我們利用染色質免疫沈澱分析晶片資料、生物晶片資料以及酵母菌相近物種之DNA序列，以生物資訊方式，發表多項基因調控的相關研究與應用軟體。

由同一轉錄因子所調控之基因表現通常與不同轉錄因子調控之基因表現有所差異，我們以此為基礎，找出和酵母菌細胞週期相關之轉錄因子(圖一)，進而發展了一套辨識轉錄因子之結合位置的軟體TFBSfinder。TFBSfinder能根據生物晶片資料分析而得轉錄因子之調控基因，並尋找在其啟動子中出現頻率極高、且在相近物種間也高度保留的DNA片段，即為轉錄因子偏好結合之特徵序列。根據群聚分析，我們發現轉錄因子之結合位置經常存在核苷酸的差異，亦即其特徵序列容許部分不完美的位點，即為退化位置。



因此，我們利用序列特徵探勘技術以及位置的同現性來加強辨識位置的可信度，並進一步尋找具非連續序列特徵之結合位置。此方法不僅對非連續序列之結合位置具有高正確率，且對一般結合位置亦有良好之辨識能力、有相當高的可信度與效度。轉錄因子常和不同的轉錄因子有交互作用，因此，我們利用轉錄因子結合位置的資訊，發展了新的轉錄因子交互作用預測演算法 **simTFBS**，其能顯著提高預測之準確度。此外，我們也建立了方便使用者查詢的資料庫平台 **MYBS**，提供轉錄因子結合位置的動態資訊。使用者可設定不同的染色質免疫沈澱分析晶片資料以及演化足跡資訊，以觀察特定基因之轉錄因子結合位置。利用 **MYBS**，我們深入探討轉錄因子結合位置的序列組成與其調控基因表現情形之間的關聯性。我們發現在酵母菌已知的可信轉錄因子結合位置上，有超過三分之一的退化位置與約五分之一的退化位置對具有「功能性」。這表示退化位置的出現確實與基因表現有顯著的關聯性，且有些退化位置需要搭配另一個退化位置，形成退化位置對，才具有調控之作用。

進一步分析這些退化位置序列的特性，我們發現這些具有「功能性」的退化位置相對於其他退化位置，在演化過程中顯著地被保留。因此，我們證實了轉錄因子結合序列中有部份退化位置對基因調控確實具備影響力。另外，我們加入核小體資料，以探討在較老直系同源基因和最近演化而成的基因中，其核小體佔據區間和核小體空乏區間內的轉錄因子結合位置的差異。我們發現核小體佔據區間的轉錄因子結合位置，相較於核小體空乏區間，有著更高比例的變異。

最後，我們系統性地檢驗背離式轉錄基因對共同啟動子中的轉錄因子結合位置。利用轉錄因子剔除實驗，我們進行全面性的軟體模擬分析，結果顯示背離基因對在大多數情況下沒有被共同調控，但在同一種環境條件下傾向被共同調控。我們多方探討並解析轉錄因子及其結合位置的特性與功能，期能開創轉錄調控研究之新頁。



2012.12.19
蔡懷寬老師醫學院演講

