بسم الله الرحمن الرحیم

وزارت علوم، تحقیقات و فناوری
پژوهشگاه ملی مهندسی ژنتیک و زیست فناوری

پایان نامه کارشناسی ارشد
رشته زیست شناسی سلولی و ملکولی

روش محاسباتی جدید برای پیش‌گویی میان‌کنش بین پروتئین‌ها

نگارش
سعیده خدایی

استاد راهنما
دکتر مهدی صادقی

اسفند ۱۳۹۰

حق استفاده از مفاد پایان نامه برای پژوهشگاه ملی مهندسی ژنتیک و زیست فناوری محفوظ است.

تقدیم به تمامی عزیزانم :

پدر بزرگوار و مادر مهربانم

و

خواهران و برادران دوست داشتنیم

سپاس نامه

تشکر و سپاس خود را تقدیم تمامی عزیزانی می کنم که جز با حمایت ها و راهنمایی هایشان قادر به نوشتن و انجام این پایان نامه نبودم:

خانوادهٔ عزیزم که همواره مشوق و پشتیبان من بوده اند.

استاد راهنمای گرامیم آقای دکتر صادقی که با دانش ارزنده و راهنمایی های بی دریغ خود علاوه بر کمک به انجام این پایان نامه، دید تازه ای در من در ارتباط با علم زیست شناسی ایجاد کردند.

دوستان عزیزم شهرزاد ونگین که در انجام تمامی مراحل این پایان نامه همراه من بودند.

آقای دکتر مرعشی و آقای شجاعی که با پیشنهادات و نظرات ارزنده مرا در نوشتن و ویرایش این مجموعه یاری رساندند.

و همچنین تمامی دوستانی که هرکدام به طریقی همراه و راه گشای من بودند.

چکیده

درک میان‌کنش بین پروتئین‌ها در بررسی انتقال پیام داخل سلولی، مسیرهای متابولیسمی و مدل‌سازی مجموعه‌های بزرگ و پیچیدهٔ پروتئینی مهم است. با افزایش تعداد ژنوم‌هایی که به صورت کامل تعیین توالی شده‌اند، شناسایی میان‌کنش بین پروتئین‌ها از روی اطلاعات ژنومی یکی از زمینه‌های مهم تحقیقاتی در پروتئومیکس کارکردی را به خود اختصاص داده است. همزمان با روش‌های آزمایشگاهی، روش‌های محاسباتی مختلفی نیز به منظور پیش‌گویی روابط عملکردی بین دو پروتئین ارائه شده‌اند. اساس این روش‌ها استفاده از اطلاعات به دست‌آمده از توالی ژن‌ها و روابط تکاملی ژنوم‌ها می‌باشد. در سال‌های اخیر مجموعه داده‌های حاصل از روش‌های مختلفی برای ساخت شبکه‌های میان‌کنش بین پروتئین‌ها به کار رفته‌اند. تلاش‌ها برای یافتن ویژگیهای جدید ژنومی که بتوان برای پیش‌گویی روابط عملکردی از آن استفاده کرد، در جریان است. در این پژوهش، یک روش محاسباتی جدیدی براساس ارتباط بین تعویض رشته‌ای دو ژن که به لحاظ عملکردی به هم وابسته هستند، پیشنهاد شده است. در طول تکامل، ممکن است که جای‌گیری در ارتباط با هم دو ژن بر روی رشته‌های رهبر و پیرو DNA، به دلیل ارتباط عملکردی بین پروتئین‌های متناظرشان باشد. روش محاسباتی جدید ما (SwInt) با استفاده از این فرض و محاسبهٔ مقدار اطلاعات دو طرفه بین تغییر مکان دو ژن بر روی دو رشتهٔ DNA، به بررسی میان‌کنش بین پروتئین‌های متناظر با این ژن‌ها می‌پردازد. این روش برای مجموعهٔ بزرگی از جفت‌پروتئین‌های *اشریشیا کولی* K۱۲، *هلیکوباکتر پیلوری* ۲٦٦۹۵ و *ویبریوکلرا* به کار رفته‌است. میان‌کنش‌های پیش‌گویی شده با کمک این روش به صورت معناداری با میان‌کنش‌های گزارش شده در پایگاه‌داده‌های شناخته‌شدهٔ میان‌کنش پروتئین-پروتئین مطابقت دارد. ما معتقدیم که مطالعهٔ ما می‌تواند به عنوان قدم اول به سوی پیش‌گویی در مقیاس وسیع میان‌کنش‌های عملکردی در ارگانیسم‌های مختلف به حساب آید.

کلید واژه‌ها:

میان‌کنش بین پروتئین‌ها، پروتئومیکس کارکردی، ارتباط عملکردی، روش‌های آزمایشگاهی، روش-های محاسباتی، شبکه‌های میان‌کنش، اطلاعات ژنومیک، تعویض رشته، اطلاعات دوطرفه

فهرست

۱- مقدمه ۲

۲- مروری بر منابع ٦

۲-۱ تعریف میان‌کنش بین پروتئین‌ها..... ٦

۲-۲ روش‌های تجربی برای شناسایی میان‌کنش پروتئین-پروتئین ۷

۲-۲-۱ روش دورگه‌سازی دوگانه مخمر ۷

۲-۲-۱-۱ مزایا و معایب روش Y2H ۹

۲-۲-۲ روش MS-TAP..... ۱۱

۲-۲-۲-۱ مقایسهٔ بین TAP-MS و Y2H ۱۲

۲-۲-۳ رسوب‌دهی ایمنی هم‌زمان ۱۲

۲-۲-٤ روش کشندگی مصنوعی ۱۲

۲-۲-۵ نمایش میان‌کنش‌های خاص ۱۳

۲-۳ روش‌های محاسباتی پیش‌گویی میان‌کنش پروتئین- پروتئین ۱۳

۲-۳-۱ روش‌های محاسباتی براساس ساختار ۱۳

۲-۳-۲ روش‌های محاسباتی براساس اطلاعات ژنومیک..... ۱٤

۲-۳-۲-۱ روش‌های مکان‌یابی توأم..... ۱٤

۲-۳-۲-۱-۱ روش خوشه‌ژنی ۱٤

۲-۳-۲-۱-۲ روش همسایگی ژن‌ها ۱٤

۲-۳-۲-۱-۳ روش ادغام ژنی ۱٦

۲-۳-۲-۲ تکامل همزمان ۱۸

۲-۳-۲-۲-۱ پروفایل فیلوژنتیک ۱۸

۲-۳-۲-۲-۲ روش In silico two- hybrid ۲۱

۲-۳-۲-۲-۳ روش شباهت بین درخت‌های فیلوژنتیک ۲۲

۲-٤ روش‌های حد واسط ۲٤

۲-٤-۱ بازسازی متابولیک ۲٤
۲-٤-۲ پروفایل‌های بیان همزمان ژن‌ها ۲۵
۲-۵ پایگاه‌داده‌ها و ابزارهای آنالیز میان‌کنش پروتئین-پروتئین ۲٦
۲-۵-۱ پایگاه‌دادهٔ STRING ۲٦
۲-۵-۲ پایگاه‌دادهٔHPRD ۲۷
۲-۵-۳ پایگاه‌دادهٔ DIP ۲۸
۲-۵-٤ پایگاه‌دادهٔ ProLink ۲۹
۲-٦ ابزارهای آنالیز و تجسم شبکه ۲۹
۲-٦-۱ ابزار تجسم Cytoscape ۲۹
۲-٦-۲ ابزار تجسم Pajek ۲۹
۲-۷ آنالیز ساختار شبکه ۳۰
۲-۷-۱ پارامترهای شبکه ۳۰
۲-۷-۱-۱ اتصال یا درجه ۳۱
۲-۷-۱-۲ توزیع درجه ۳۱
۲-۷-۱-۳ ضریب خوشه‌بندی ۳۱
۲-۷-۱-٤ طول مسیر مشخصه ۳۲
۲-۷-۱-۵ قطر شبکه ۳۲
۲-۷-۱-٦ میانه ۳۲
۲-۷-۲ توپولوژی شبکه ۳۲
۲-۷-۲-۱ شبکه‌های تصادفی ۳۳
۲-۷-۲-۲ شبکه‌های جهان کوچک ۳۳
۲-۷-۲-۳ شبکه‌های مستقل از مقیاس ۳٤
۲-۸ روش پیشنهادی ما (SwInt) ۳۵
۲-۸-۱ فرضیات ۳۵

۲-۸-۱-۱ انتقال ژن‌ها در طول تکامل ۳۶
۲-۸-۱-۱-۱ ارتباط بین تکرارهای معکوس و تکامل ساختار ژنوم باکتری ۳۶
۲-۸-۱-۱-۲ معکوس شدگی و فرایند همانندسازی ۳۶
۲-۸-۱-۲ ارتباط بین انتقال ژن‌ها در طول تکامل ۳۷
۲-۸-۲ اهداف پژوهش حاضر ۳۸
۳- مواد و روشها ٤۰
۳-۱ جمع‌آوری داده‌ها: ٤۰
۳-۲ محاسبهٔ اطلاعات دوطرفه ٤۱
۳-۲-۱ آنتروپی ٤۲
۳-۲-۱-۱ آنتروپی توأم ٤۲
۳-۲-۱-۲ آنتروپی شرطی ٤۳
۳-۲-۲ اطلاعات دوطرفه: ٤۳
۳-۳ محاسبهٔ اطلاعات دوطرفهٔ بین دو ژن ٤۵
۳-۳-۱ تاثیر اندازهٔ نمونه بر روی اطلاعات دوطرفه ٤٦
۳-٤ بررسی کارایی روش ٤٦
۳-٤-۱ مقایسه با پایگاه‌دادهٔ COG ٤۷
۳-٤-۲ مقایسه با پایگاه‌داده‌های شناخته شدهٔ میان‌کنش پروتئین- پروتئین ٤۷
٤- نتایج و بحث ۵۲
٤-۱ بررسی ارتباط عملکردی جفت‌های با امتیاز Z-score بالا ۵۲
٤-۲ مقایسه با پایگاه‌داده‌های استاندارد میان‌کنش پروتئین-پروتئین ۵۷
٤-۲-۱ نمودار ROC برای مقایسهٔ روش SwInt با روش‌های تجربی ۵۷
٤-۲-۲ نمودار ROC برای مقایسهٔ روش SwInt با پایگاه‌دادهٔ STRING ۵۸
٤-۲-۳ نمودار ROC برای مقایسهٔ روش SwInt با روش‌های محاسباتی ۵۹
٤-۳ بررسی میان‌کنش عملکردی بین پروتئین‌های موجود در سایر ژنوم‌ها ٦۰

٤-٤ پیش‌گویی‌های با امتیاز بالا ٦٠
٤-٤-١ میان‌کنش‌های شناخته‌شده ٦١
٤-٤-٢ پیش‌گویی میان‌کنش‌های جدید ٦٢
٤-٤-٢-١ میان‌کنش بین فعال‌کنندهٔ رونویسی rfaH با پروتئین‌های ریبوزومی L_1 و L_{10} ٦٢
٤-٤-٢-٢ میان‌کنش بین فعال‌کنندهٔ رونویسی rfaH و زیرواحد بتای RNA پلیمراز ٦٢
٤-٤-٢-٣ میان‌کنش بین پروتئین‌های yadB و yeiL ٦٢
٤-٤-٢-٤ میان‌کنش بین پروتئین‌های crp و yadB ٦٣
٤-٤-٢-٥ میان‌کنش بین پروتئین‌های rhmD و ilvD ٦٣
٤-٤-٢-٦ میان‌کنش بین پروتئین‌های gmk و rpoZ در *هلیکوباکتر پیلوری* ٦٣
٥- جمع‌بندی ٦٦
٦- پیشنهادات ٦٩
٧- منابع ٧١

فهرست جدول‌ها

جدول ۲-۱ روش‌های مختلف آزمایشگاهی برای شناسایی میان‌کنش‌های پروتئینی ۸
جدول ۲-۲ پایگاه‌داده‌های میان‌کنش بین پروتئین‌ها ۲۸
جدول ۲-۳ مثال‌هایی از ابزارهای تجسم شبکه‌ها و میان‌کنش‌های بین پروتئین ها ۳۰
جدول ٤-۱ بررسی کارایی روش‌های مختلف در پیش‌گویی میان‌کنش‌های عملکردی..... ۵۵
جدول ٤-۲ تعداد میان‌کنش‌های عملکردی پیش‌گویی‌شده به تفکیک گروه‌های عملکردی COG..... ۵٦
جدول ٤-۳ محاسبهٔ حساسیت روش‌های محاسباتی در مقایسه با روش‌های آزمایشگاهی ٦۰
جدول ٤-٤ مقدار AUC نمودار ROC حاصل از مقایسهٔ میان‌کنش‌های پیش‌گویی شده در ژنوم‌های *هلیکوباکترپیلوری* و *ویبریوکلرا* با پایگاه‌داده‌های مختلف..... ٦۱

فهرست شکل‌ها

شکل ۲-۱: تصویر شماتیک روش‌های اصلی آزمایشگاهی ۱۰
شکل ۲-۲ شبکهٔ جهان کوچک ۳٤
شکل ۳-۱ رابطهٔ بین آنتروپی و اطلاعات دوطرفه ٤٤
شکل ۳-۲دسته‌بندی داده‌ها ٤۹
شکل ۳-۳ منحنی پایهٔ ROC..... ۵۰
شکل ٤-۱ ارزیابی احتمال تعلق دو پروتئین به یک گروه عملکردی COG مشترک..... ۵۳
شکل ٤-۲ منحنی‌های ROC پیش‌گویی میان‌کنش‌های عملکردی توسط روش پیشنهادی ما و سایر روش‌ها. ۵٤
شکل ٤-۳ منحنی ROC مقایسهٔ روش پیشنهادی ما با پایگاه‌دادهای تجربی ۵۷
شکل ٤-٤ نمودار ROC مقایسهٔ روش پیشنهادی ما با پایگاه‌داده STRING..... ۵۸
شکل ٤-۵نمودار ROC مقایسهٔ روش پیشنهادی ما با پایگاه‌دادهٔ ProLink ۵۹

بخش اول

مقدمه

۱- مقدمه

با گسترش روش‌های تعیین توالی ژنوم در سطح وسیع، زیست‌شناسان به حجم وسیعی از اطلاعات ژنومی مربوط به ارگانیسم‌های مختلف دسترسی پیدا کردند. مسئله اصلی که زیست‌شناسان در حال حاضر با آن مواجه هستند، شناسایی عملکرد ژن‌های مختلف و درک فرایندهای زیستی درون سلول با کمک این اطلاعات می‌باشد. در این میان، میان‌کنش بین پروتئین‌ها به دلیل نقش مهمی که در این فرایندها دارند، بخش مهمی از تحقیقات پساژنومیک را به خود اختصاص داده‌اند. شناسایی میان‌کنش بین پروتئین‌ها و ساخت شبکه‌های میان‌کنش در درک مکانیسم پیام‌های خارج سلولی، بررسی مسیرهای متابولیسمی و بررسی ساختار مجموعه‌های پروتئینی بزرگ اهمیت دارند.

تاکنون روش‌های تجربی متعددی به صورت گسترده برای شناسایی میان‌کنش بین پروتئین‌ها و ساخت شبکه‌های بزرگ میان‌کنش استفاده شده‌اند. انواعی از این روش‌ها شامل دورگه‌سازی مخمری[۱] (Y2H) (Ito *et al.*, 2001)، طیف‌سنجی جرمی[۲] همراه با خالص‌سازی تمایلی همزمان[۳] (MS-TAP) (Rigaut *et al.*, 1999)، ریزآرایه‌های پروتئین و DNA (Jansen *et al.*, 2002) و سیستم فاژی[٤] (Smith, 1985) می‌باشند.

علاوه بر این روش‌ها، روش‌های محاسباتی برای پیش‌گویی میان‌کنش عملکردی[٥] بین پروتئین‌ها گسترش پیدا کرده‌است. اصطلاح میان‌کنش عملکردی به مشارکت پروتئین‌ها در میان‌کنش فیزیکی مستقیم، در ساختار مجموعه‌های بزرگ و یا در یک مسیر بیولوژیکی مشترک اطلاق می‌شود.

تعدادی از این روش‌های محاسباتی از اطلاعات بیولوژیکی نظیر بیان همزمان ژن‌ها[٦] برای پیش‌گویی استفاده می‌کنند. این روشها از نتایج روش‌های تجربی ریزآرایه‌ها که فعالیت هر ژن خاص را در ژنوم در شرایط متفاوت سلول و در زمان‌های مختلف نشان می‌دهند، استفاده می‌کنند. به احتمال زیاد ژن هایی که بیان آنها در شرایط مختلف سلول و در زمان خاصی با همدیگر صورت می گیرد از نظر عملکردی با هم در ارتباط هستند (Jansen et al., 2002, Kemmeren et al., 2002, Troyanskaya et al., 2002).

---

[۱] Yeast two- hybrid

[۲] Mass spectroscopy

[۳] Tandem affinity purification

[٤] Phage display

[٥] functional interaction

[٦] Coexpression

بعضی از روش‌های محاسباتی برای پیش‌گویی میان‌کنش‌های عملکردی بین پروتئین‌ها از اطلاعات ساختاری موجود در پایگاه‌داده‌هایی نظیر PDB استفاده می‌کنند. البته به علت تعداد محدود پروتئین-هایی که ساختار سه‌بعدی آنها شناسایی شده است، این روش‌ها در سطح وسیع به کار نمی‌روند (Aloy and Russell, 2002, Aloy and Russell, 2003).

روش‌های محاسباتی که در سال‌های اخیر توسعه پیدا کرده‌اند، از اطلاعات موجود در توالی ژنوم و همچنین روابط تکاملی ژن‌ها در ژنوم‌های مختلف برای پیش‌گویی وابستگی عملکردی ژن‌ها استفاده می‌کنند. نیروی انتخاب طبیعی باعث می‌شود که ژن‌هایی که به لحاظ عملکردی با هم در ارتباطند، در مجاورت هم قرار بگیرند. نمونهٔ بارز آن در سلول‌های پروکاریوتی و در مورد اپران‌ها صورت گرفته-است. روش خوشهٔ ژنی[۷] بر اساس فواصل درون‌ژنی[۸] ساختار این اپران‌ها را در ژنوم‌های مختلف شناسایی می‌کند (Ermolaeva et al., 2001, Bowers et al., 2004) .

با وجود رخدادهای تکاملی نظیر انتقال افقی ژن‌ها، ممکن است ژن‌های درون اپران‌ها و برخی از ژن‌هایی که درون سلول‌های یوکاریوت با هم تنظیم می‌شوند، نزدیکی خود را در ژنوم‌های مختلف حفظ کنند. بر این اساس، روش همسایگی ژنی[۹] وابستگی عملکردی ژن‌ها را بر مبنای مجاورت آنها در ژنوم‌های مختلف ، پیش‌گویی می‌کند (Dandekar et al., 1998, Galperin and Koonin, 2000, Rogozin et al., 2002). همچنین، ادغام پروتئین‌ها در طی تکامل به منظور کاهش بار تنظیم همزمان پروتئین‌هایی که با هم میان‌کنش دارند، صورت می‌گیرد. روش ادغام ژنی[۱۰] جستجوی پروتئین‌هایی را انجام می‌دهد که در برخی از گونه‌ها از هم جدا هستند و در یک یا تعداد بیشتری از ارگانیسم‌ها، دومین‌های[۱۱] یک پروتئین چند دومینی با فعالیت متعدد را تشکیل می‌دهند (Marcotte et al., 1999, Marcotte and Marcotte, 2002).

دو ژنی که عملکردشان در ارتباط با هم باشد، ممکن است با هم تکامل پیدا کنند. در طی تکامل ژنوم، برخی از ژن‌ها ایجاد و برخی دیگر از بین می‌روند. در روش پروفایل فیلوژنتیک[۱۲]، پروفایل‌های جداگانه برای هر کدام از پروتئین‌ها بر اساس حضور یا عدم حضورشان در یک ژنوم خاص ساخته

---

[۷] gene cluster

[۸] intergenic distances

[۹] gene neighborhood

[۱۰] gene fusion (Rosetta stone)

[۱۱] Domaines

[۱۲] phylogenetic profile

می‌شود و با مقایسهٔ این پروفایل‌ها با هم و شناسایی پروتئین‌هایی که پروفایل شبیه به هم دارند، میان-کنش بین آنها پیش‌گویی می‌شود (Pellegrini et al., 1999, Date and Marcotte, 2003).

حالت دیگری از بررسی تکامل همزمان ژن‌ها، بررسی شباهت بین درخت‌های فیلوژنتیک مربوط به آنها می‌باشد. شباهت بین درخت فیلوژنتیک[۱۳] (درخت آینه‌ای[۱۴]) دو پروتئین، نشانهٔ وابستگی عملکردی بین آنها می‌باشد (Pazos and Valencia, 2001, Pazos and Valencia, 2002, Jothi *et al.*, 2005).

همچنین ممکن است از تغییرات (جهش‌های) همزمان توالی‌های دو ژن، تکامل همزمان آنها استنباط شود. روش in silico two-hybrid با محاسبهٔ جهش‌های همبسته[۱۵] برای هر جفت از اسیدهای آمینهٔ بین دو پروتئین، میان‌کنش‌ها را پیش‌گویی می‌کند (Pazos and Valencia, 2002).

در این پایان‌نامه، روش محاسباتی جدیدی برای پیش‌گویی میان‌کنش عملکردی بین پروتئین‌ها معرفی می‌شود. اساس روش پیشنهادی ما (SwInt) ارتباط بین تعویض‌های رشته‌ای[۱۶] ژن‌هایی است که عملکردشان به هم مرتبط است. تعویض‌های رشته‌ای به انتقال ژن‌ها از روی یک رشته DNA به روی رشتهٔ دیگر در طول تکامل اشاره می‌کند. یکی از مکانیسم‌های مهم این انتقال در پروکاریوت‌ها، نوترکیبی بین تکرارهای معکوس است. اگر تکرارهای معکوس بر روی رپلیکور[۱۷] یکسانی باشند، نوترکیبی بین آنها در هنگام همانندسازی، توالی حدفاصل این تکرارها را از روی یک رشتهٔ DNA به روی رشتهٔ دیگر انتقال می‌دهد. بنابراین در طول تکامل، توالی ژن‌هایی که در حدفاصل این تکرارهای معکوس هستند از روی رشتهٔ رهبر به روی رشتهٔ پیرو و بالعکس انتقال پیدا می‌کنند (Achaz et al., 2003). اگر این تعویض‌های رشته‌ای برای دو ژن وابسته به هم صورت گیرد، می‌توان نتیجه گرفت که رابطه عملکردی بین پروتئین‌های آنها وجود دارد. SwInt از این فرض برای پیش‌گویی میان‌کنش بین دو پروتئین استفاده کرده است. اطلاعات دوطرفه به عنوان یک شاخص کمی برای محاسبهٔ ارتباط بین تعویض‌های رشته‌ای دو ژن به کار رفته‌است. بر اساس این فرض، جفت‌های ژنی با مقدار اطلاعات دوطرفهٔ بالا، کاندیداهای مناسب برای بررسی میان‌کنش عملکردی بین پروتئین‌های‌شان می-باشند.

---

[۱۳] phylogenetic tree

[۱۴] mirror tree

[۱۵] correlated mutations

[۱۶] strand switches

[۱۷] replicore

فصل دوم

مروری بر منابع

۲- مروری بر منابع

تحقیقاتی که در دهه‌های اخیر در زمینه زیست‌شناسی سلولی و مولکولی، بیوشیمی و بیوفیزیک صورت گرفته، اطلاعات باارزشی از خصوصیات مولکولی، ساختاری و عملکردی پروتئین‌ها فراهم کرده است. امروزه این اطلاعات در پایگاه‌داده‌هایی نظیر UniProt در دسترس همگان قرار گرفته است. باید توجه داشت که پروتئین‌ها به ندرت به صورت مجزا از هم فعالیت می‌کنند. بیشتر فعالیت-هایی که درون سلول زنده انجام می‌شود حاصل فعالیت گروهی پروتئین‌ها با هم و ارتباطات فیزیکو-شیمیایی فعالی است که این مولکول‌ها با یکدیگر و با سایر مولکول‌ها برقرار می‌کنند.

نقشه‌برداری از میان‌کنش‌هایی که بین پروتئین‌های درون یک موجود زنده رخ می‌دهد، از مهمترین مراحل شناسایی و آشکارسازی روابط مولکولی بین مجموعه‌های درگیر در فرایندهای زیستی می‌باشد (De Las Rivas and Fontanillo, 2010).

قبل از معرفی روش‌هایی که تاکنون برای شناسایی میان‌کنشهای پروتئینی مورداستفاده قرار گرفته‌اند، لازم است که اصطلاح «میان‌کنش بین پروتئین‌ها» تعریف شود.

۲-۱ تعریف میان‌کنش بین پروتئین‌ها

به طور معمول اصطلاح میان‌کنش بین دو پروتئین به عنوان ارتباط فیزیکی مستقیم بین آنها، که حاصل اتصالات مولکولی بین اسیدهای آمینه‌شان می‌باشد، تعریف می‌شود. این سؤال که آیا دو پروتئین «ارتباط عملکردی» با یکدیگر دارند از این پرسش که آیا دو پروتئین به صورت فیزیکی مستقیماً با هم میان‌کنش دارند متفاوت است. به عنوان مثال، پروتئین‌هایی که در مجموعه ریبوزوم هستند و یا پروتئین‌هایی که در فرایند رونویسی دخیل هستند به لحاظ عملکردی با یکدیگر در ارتباطند، اما تمامی پروتئین‌هایی که در یک مجموعه خاص هستند و یا در یک فرایند سلولی با هم شریکند لزوماً به صورت مستقیم با یکدیگر میان‌کنش ندارند. بنابراین میان‌کنش بین پروتئین‌ها علاوه بر اینکه در مورد پروتئین‌هایی که "ارتباط فیزیکی" مستقیمی با هم دارند استفاده می‌شود، در مورد آن دسته از پروتئین‌هایی که با یکدیگر "ارتباط عملکردی" دارند (در یک مجموعه چندزیرواحدی هستند و یا در مسیرهای زیستی با هم مشارکت دارند) نیز به کار می‌رود.

در میان کنش بین پروتئین‌ها، پیوند فیزیکی بین پروتئین‌ها باید اختصاصی باشد و حاصل تماس تصادفی پروتئین‌ها با همدیگر نباشد. همچنین تمام میان کنش‌هایی که یک پروتئین در زمان ساخت، تعیین ساختار و شکستن تجربه می‌کند، باید حذف شود. برای مثال، همهٔ پروتئین‌ها در یک مرحله از حیاتشان با پروتئین‌های ریبوزومی تماس دارند و تعداد زیادی از آنها با چاپرون‌ها تماس پیدا می‌کنند.

میان‌کنش فیزیکی بین دو پروتئین به صورت دائمی یا گذرا است. میان‌کنش بین پروتئین‌هایی که مجموعه‌های بزرگ پروتئینی را می‌سازند، مانند ATP سنتتاز (شامل هشت پروتئین مختلف در پستانداران) و یا سیتوکروم اکسیداز (شامل سیزده پروتئین مختلف در پستانداران)، پایدار هستند. پروتئین‌هایی که در چنین مجموعه‌هایی نقش دارند به عنوان «زیرواحد» شناخته می‌شوند. سایر اجتماعات پروتئینی مثل فرایند بیان ژن که با اتصال فاکتورهای رونویسی بر روی ناحیهٔ پروموتر ژن فعال می‌شود، فقط برای انجام اعمال موقتی سلول شکل می‌گیرد و گذرا است.

یکی دیگر از عناصر مهم دیگر برای تعریف میان‌کنش بین پروتئین‌ها، اطلاعات زیستی است. همهٔ میان‌کنش‌های ممکن، در هر سلول و در هر زمانی رخ نمی‌دهند. در عوض، میان‌کنش‌ها به نوع سلول، فاز چرخهٔ سلولی، مرحلهٔ تمایز، تغییرات پروتئین (مانند فسفریلاسیون)، حضور کوفاکتورها و دیگر شریک‌های اتصال بستگی دارند (De Las Rivas and Fontanillo, 2010) .

۲-۲ روش‌های تجربی برای شناسایی میان‌کنش پروتئین-پروتئین

میان‌کنش‌های پروتئین بوسیلهٔ روش‌های فیزیکی، ژنتیکی و بیوشیمیایی مختلف آنالیز شده‌اند. تعدادی از این روش‌ها، قادر به شناسایی میان‌کنش گروه بزرگی از پروتئین‌ها در سلول هستند. نمونه-ای از این روش‌ها سیستم دورگه‌سازی دوگانهٔ مخمر (Y2H) طیف‌سنجی جرمی همراه با خالص-سازی تمایلی همزمان (TAP- MS) ریزآرایهٔ DNA و پروتئین، کشندگی مصنوعی[۱۸] می‌باشد. روش-های دیگر بر روی نمایش و تعیین خصوصیات بیوشیمیایی و فیزیکوشیمیایی مجموعه‌های پروتئینی تمرکز می‌کنند. برخی از روش‌ها تنها میان‌کنش مستقیم بین دو پروتئین را شناسایی می‌کنند، در حالی-که روش‌های دیگر علاوه بر میان‌کنش‌های مستقیم می‌توانند میان‌کنش‌های غیرمستقیم بین یک مجموعهٔ پروتئینی را شناسایی کنند. در جدول ۲-۱ اسامی تعدادی از روش‌های آزمایشگاهی و نوع میان‌کنش‌هایی که شناسایی می‌کنند نشان داده شده است.

۲-۲-۱ روش دورگه‌سازی دوگانه مخمر

روش دورگه‌سازی دوگانه مخمر (Y2H) بر اساس ساختار فعال‌کننده‌های رونویسی در سلول‌های یوکاریوتی است. بیشتر این فعال‌کننده‌ها از دو دومین مجزا ساخته شده‌اند، دومین متصل‌شونده به ناحیه پروموتری توالی DNA (DBD) و دومینی که فعال‌کننده رونویسی است و به صورت مستقیم با توالی DNA در ارتباط نیست (AD). برای فعال‌سازی رونویسی نیاز به اتصال فیزیکی (اتصال غیرکوالانسی) دو دومین است و هیچ دومینی به تنهایی قادر به فعال‌کردن رونویسی نیست. در روش

[۱۸] Synthetic lethality

جدول ۲-۱ روش‌های مختلف آزمایشگاهی برای شناسایی میان‌کنش‌های بین پروتئینی (Shoemaker and Panchenko, 2007a)

روش	در مقیاس وسیع	روش سنجش	نوع میان‌کنش	نوع خصوصیات
دورگه‌سازی مخمر	+	در سطح موجود زنده	فیزیکی(دوتایی)	شناسایی
MS-TAP	+	در سطح آزمایشگاه	فیزیکی(مجموعه)	شناسایی
ریزآرایهٔ DNA	+	در سطح آزمایشگاه	عملکردی	شناسایی
ریزآرایه‌های پروتئین	+	در سطح آزمایشگاه	فیزیکی(مجموعه)	شناسایی
کشندگی مصنوعی	+	در سطح موجود زنده	عملکردی	شناسایی
سیستم فاژی	+	در سطح آزمایشگاه	فیزیکی(مجموعه)	شناسایی
کریستالوگرافی اشعهٔ X طیف سنجی NMR	-	در سطح آزمایشگاه	فیزیکی(مجموعه)	خصوصیات ساختاری وزیستی
انتقال انرژی تشدید فلورسانس	-	در سطح موجود زنده	فیزیکی(دوتایی)	خصوصیات زیستی
تشدید پلاسمون سطح	-	در سطح آزمایشگاه	فیزیکی(مجموعه)	جنبشی، خصوصیات پویا
میکروسکوپ نیروی اتمی	-	در سطح آزمایشگاه	فیزیکی(دوتایی)	مکانیکی، خصوصیات پویا
میکروسکوپ الکترونی	-	در سطح آزمایشگاه	فیزیکی(مجموعه)	خصوصیات ساختاری وزیستی

در ستون دوم، روش‌های در مقیاس وسیع با مثبت نشان‌داده شده، در ستون سه نوع سنجش در سطح موجود زنده یا درسطح آزمایشگاه مشخص شده است. در ستون چهارم نوع روش شناسایی میان‌کنش فیزیکی مستقیم، بر اساس اینکه میان‌کنش فیزیکی بین دو پروتئین (دوتایی) شناسایی شود یا میان‌کنش بین گروهی از پروتئین‌ها (مجموعهٔ با هم)، مشخص شده است. همچنین روش‌های تعیین‌کنندهٔ میان‌کنش عملکردی هم نشان داده شده است. نوع خصوصیات میان‌کنش پروتئین‌ها در ستون آخر ارائه شده است.

Y2H، دورگه‌ای شامل یکی از دو پروتئین موردنظر با دومین فعال‌کنندهٔ رونویسی و دورگهٔ دیگر با پروتئین دوم و دومین متصل‌شونده به DNA ساخته می‌شود. پروتئینی که با دومین متصل‌شونده به DNA هیبرید می‌شود به نام پروتئین «طعمه» و پروتئین دورگه دیگر به عنوان پروتئین «شکار» خوانده می‌شوند. این پروتئین‌های دورگه داخل یک پلاسمید بیانی کلون می‌شوند و به درون یک سلول مخمر ترانسفورم می‌شوند. در صورتی که این دو پروتئین با هم میان‌کنش داشته باشند دو دومین به هم متصل شده و ژن گزارشگر فعال می‌شود (شکل۲-A۱). سیستم‌های تک هیبریدی یا سه هیبریدی

برای شناسایی میان‌کنش‌های بین پروتئین‌ها با ماکرومولکول‌های DNA و RNA توسعه پیدا کرده‌اند (Van Criekinge and Beyaert, 1999, Fashena *et al.*, 2000, Causier, 2004).

برای شناسایی میان‌کنش بین پروتئین‌ها در سطح وسیع روش Y2H به دو صورت روش شبکه‌ای و روش کتابخانه‌ای به‌کار می‌رود.

۱- روش شبکه‌ای[۱۹] در این روش یک شبکه از کلون پروتئین‌های طعمه ساخته می‌شوند. هر کلون، یک پروتئین طعمه خاص را بیان می‌کند و داخل یک چاهک از ظرف قرار می‌گیرد. سپس هرکدام از پروتئین‌های دورگه شکار با این آرایه از پروتئین‌های طعمه برخورد داده می‌شوند. میان‌کنش بین دو پروتئین دورگه براساس بیان ژن گزارشگر و جایگاه آن بر روی ظرف شناسایی می‌شود.

۲- روش کتابخانه‌ای[۲۰] ابتدا کتابخانه‌ای از پروتئین‌های دورگه شکار ساخته‌می‌شود. این پروتئین‌ها قطعاتی از cDNA تصادفی و چهارچوب‌های باز خواندنی هستند. هر پروتئین طعمه در مقابل یک کتابخانه از پروتئین‌های شکار ناشناخته آشکارسازی می‌شوند. دوتایی‌ها مثبت بر اساس توانایی رشدشان در یک سوبسترای خاص انتخاب می‌شوند و پروتئین‌هایی که با هم میان‌کنش دارند با روش توالی‌یابی DNA شناسایی می‌شوند (Shoemaker and Panchenko, 2007a)

۲-۲-۱-۱ مزایا و معایب روش Y2H

از آنجا که این روش در سطح موجود زنده انجام می‌شود، حساسیت آن به نسبت روش‌های آزمایشگاهی دیگر بالاتر است و میان‌کنش‌های ضعیف و گذرا قابل آشکارسازی هستند. همچنین نیازی به خالص‌سازی پروتئین‌ها نیست. با این وجود، استفاده از روش Y2H محدودیت‌هایی دارد. برای مثال پروتئین‌هایی که خودشان به تنهایی قادر به فعال‌کردن رونویسی هستند، نمی‌توانند نمونه-های مناسبی برای این روش باشند. اتصال پروتئین هدف با هر یک از دومین‌ها ممکن است موجب تغییر در ساختار پروتئین هدف شده و بر روی درستی نتایج اثر بگذارد. به علاوه، ساختار فضایی و تغییرات پس از ترجمه پروتئین هدف در سلول مخمر ممکن است با سایر موجودات متفاوت باشد و این مسئله شناسایی میان‌کنش بین پروتئین‌های پروکاریوتی و پستانداران را در سلول مخمر با مشکل مواجه می‌کند. و بالاخره، به دلیل اینکه روش Y2H در درون هسته انجام می‌شود، آشکارسازی میان-کنش بین پروتئین‌هایی که در شرایط فیزیولوژیکی سلول در بخش دیگر از سلول فعالیت می‌کنند (برای مثال پروتئین‌های ترشحی و گیرنده‌های سطح سلول) امکان‌پذیر نیست. برای برطرف‌کردن

[۱۹] matrix approach

[۲۰] laboratory approach