



مروری بر بیوانفورماتیک و کاربرد آن در کشاورزی

*فاضلی نسب، بهمن^۱، مهربانی، علی اشرف^۲

۱- کارشناس ارشد بیوتکنولوژی کشاورزی، Fazelie58@yahoo.com

۲- استادیار دانشکده کشاورزی دانشگاه ایلام

چکیده

در دهه گذشته، تحقیقات در زمینه های بیولوژی سلولی و مولکولی و بیوتکنولوژی از رشد بسیار بالایی برخوردار بوده و هرساله بر حجم آن افزوده می شود. آن زمان شاهد تحقیقات در زمینه پروژه های ژنومی انسان، برنج، آرابیدوپسیس و .. بوده ایم که حجم زیادی از اطلاعات را در اختیار محققین قرار داده و باعث افزایش تحقیقات دیگر ژنومی شده و راه را برای تکنیکهایی مثل میکروآرای و پروتئومیک که عصر جدیدی از تحقیقات ژنومی را بوجود آورده اند، باز کرده است. با این همه اطلاعات سلولی و مولکولی، وجود بانکهای اطلاعاتی، ابزارهای جستجوگر و آنالیزی جهت بررسی پروتئینهای و ژنهای احتمالی کشف شده نیاز است. در کشاورزی با کمک تکنیکهای QTL، BSA، پروتئومیک، مارکرهای مولکولی و ... با استفاده از بانکهای اطلاعاتی و با بررسی پروژه های ژنومی EST، STS، HTC، HTG و ... ژنها و پروتئینهای جدید عامل صفات کیفی و کمی را کشف و یا ارتباط بین مارکرهای ژنتیکی با مارکرهای مورفولوژیکی را پیدا کرد و همچنین به کمک پروژه های ژنومی انجام شده به تهیه پرایمرهایی جهت تکثیر ژنها و یا قسمتهای خاصی از DNA برای تعیین چند شکلی بین گیاهان و جانوران و احتمالاً به کلون کردن این ژنها به موجودات جدید اقدام کرد.

واژه های کلیدی: بیوانفورماتیک، بانکهای اطلاعاتی، محصولات زراعی و باغی، کشاورزی



مقدمه

بیوانفورماتیک به مجموعه ای از توالیها، بانک های اطلاعاتی و دانش بیولوژیکی گفته می شود که اطلاعات را در جای ذخیره کرده تا دستیابی و آنالیز بهتر و سریعتری روی آنها صورت بگیرد و یا به مجموعه ای از روشهای آماری، ریاضی، کامپیوتری گفته می شود که با استفاده از توالیهای پروتئین ها و DNA به حل مسائل بیولوژیکی می پردازد. توالی یابی پروتئینها و بانکهای پروتئینی در سال ۱۹۶۰ مصادف با زمانیکه اینترنت برای اولین بار بوجود آمد شروع شد ولی تا سال ۱۹۶۵ این توالیها بقدری کم بودند که در یک فلاپی جای می گرفتند اما بعدها با بوجود آمدن تکنیکهای مختلف (الکتروفورز دو بعدی، میکروآرای پروتئینها، پروتئومیک و ...) این توالیها بقدری زیاد شدند که نیاز به فضای زیادی جهت ذخیره داشتند (بانکها). در سال ۱۹۶۹ مقالات زیادی بصورت الکترونیکی (اینترنتی) منتشر شد و در سال ۱۹۸۱ محققین توانستند برای اولین بار اطلاعات خود را به شکل الکترونیکی مبادله نمایند (ایمیل). جهانی شدن اینترنت در سال ۱۹۸۲ صورت گرفت. ولی بانکهای DNA در سال ۱۹۸۰ تاسیس گردیدند (به علت مشکلاتی که بر روی عامل وراثت و خود ساختار DNA وجود داشت). امروزه بانکهای مختلفی (NCBI، DDBJ، EMBL، Pdb و ...) وجود دارند که توالیهای DNA و پروتئینها را بطور مختلط یا جداگانه نگهداری می کنند و زمینه را برای مطالعات بیشتر ژنومی، پروتئینی، تغییرات بعد از ترجمه، اثرات متقابل پروتئینها و ... فراهم می آورند (۱، ۲، ۳ و ۴)

امروزه با بوجود آمدن تکنیکهای مختلف مولکولی (پروتئومیک، میکروآرای DNA، میکروآرای پروتئینها، STS، EST و ...) و سرعت تولید سخت افزارها و تحقیقات مختلف مولکولی نتایج زیادی بدست آمده که اینترنت بهترین و سریعترین ابزار جهت بررسی و استفاده از این نتایج می باشد (۳).

اصولی که پایه و اساس بیوانفورماتیک بر آنها استوار است (۲، ۳ و ۴)

۱- DNA Micro array - ۲ Functional genomic - ۳ Structural genomic - ۴ Comparative genomic
مطالعه ژنوم کامل برای فهمیدن اختلافات و تبادلات ژنها بین گونه های مختلف ۵-Proteomic

اهداف بیوانفورماتیک (۳)

- ۱- قرار دادن اطلاعات در ساده ترین شکل برای استفاده محققین و دستیابی آنها به این اطلاعات و سعی در ارائه اطلاعات جدید توسط آنها
- ۲- توسعه ابزارها و منابعی که به آنالیز و بازیابی داده ها کمک کند مثل Fasta و Blast
- ۳- استفاده از این ابزارها جهت آنالیز داده ها و یکی کردن نتایج حاصل از آزمایشات مختلف

کاربرد بیوانفورماتیک (۱، ۲ و ۴)

- ۱- نقشه یابی و شناسایی ژنها، مناطق برش، منطق ORF و ...
- ۲- مقایسه توالیهای بدست آمده با توالیهای دیگر از سایر بانکها جهت شناسایی مناطق Matching و Mismatching و ...
- ۳- طراحی پرایمر و طراحی دیجنریت پرایمرها
- ۴- پیش بینی عملکرد واقعی ژن
- ۵- مدلینگ مولکولی و بیولوژی ساختاری (یعنی پیشبینی ساختارهای ثانویه و ثالثیه و ..)
- ۶- مدیریت داده های وسیعی که از میکروسکوپ الکترونی، پروژه های ژنومی و پروتئینی و .. بدست می آید



۷- ترجمه توالیهای DNA موجود در EMBL به پروتئین و مقایسه با توالیهای کشف شده و موجود در بانکهای مختلف جهت تایید این توالیها

تاکنون ۱۷ نوع طبقه بندی برای بانکها گزارش شده و با توجه به سرعت روز افزون تکنیکهای مختلف ژنومی و پروتئینی و افزایش توالیهای کشف شده، بانکها را با توجه به این تکنیکها تقسیم بندی می کنند. ابتدا بانکها را بر اساس باکتریایی، ویروسی، گیاهی، جانوری تقسیم بندی میکردند ولی بعدها به شکل پروتئینی، DNA، کربوهیدرات، 3D STRUCTURE در آمدند و امروزه بر اساس EST، STS، HTC، HTG، SNP و ... تقسیم بندی می گردند. البته بانکهای هم هستند که اطلاعات بسیار مفیدی را در مورد یک یا چند گونه مختلف گیاهی یا جانوری و .. در اختیار محققین قرار می دهند(۷).

بانکهای ژنومی

بانکهای هستند که اطلاعات زیادی در مورد ژنوم یک یا چند گونه خاص را در بردارند. مثلا راجع به محصولات کشاورزی اطلاعاتی در مورد پروتئین ها و ژنهای عامل صفات کمی و کیفی که در تحت شرایط خاصی مثل شوری، خشکی، دما، آفات و بیماریها و ... از گونه هایی مثل گندم، جو، پنبه، درختان میوه، گیاهان دارویی و ... منتشر می کنند

<http://ars-genome.cornell.edu> پنبه، گندم، جو، یونجه، چاودار، برنج
مورد ۲۰۰۰ گونه مختلف
<http://www.sp2000.org> ذرت
<http://www.agron.missouri.edu> پنبه، گندم، جو، یونجه، برنج
<http://ars-genome.cornell.edu>

بانکهای پروتئینی

برای اولین بار در سال ۱۹۶۰ بوجود آمدند و بهترین نوع بانکها هستند زیرا محصول نهایی ژن می باشند و امروزه با استفاده از مارکهای پروتئینی، ایزوآیمی، تکنیک میکروآری پروتئینها گیاهان (گندم، جو، سویا و ...) را تحت شرایط خاصی مثل شوری، خشکی، آفات و بیماریها و ... قرار داده و پروتئینهای که در این شرایط بیان می شوند را توالی یابی می کنند و با استفاده از تکنیکهای Mass Spectrometry، Mass Fingerprinting و ... تغییرات بعد از ترجمه، اثرات متقابل پروتئینها و ... مشخص کرده و می توانند با استفاده از توالی پروتئین مشخص شده، توالی DNA کد کننده آن (ORF) را مشخص کنند و اقدام به کلون کردن نمایند(۱ و ۵).

1- Pir <http://pir.georgetown.edu>
2- Swiss-prot <http://ca.expasy.ch>
<http://us.expasy.org>

بانکهای DNA

چون ماده اصلی وراثتی اسیدهای نوکلئیک می باشند و جهت تهیه موجودات ترانسژنیک باید عمل کلونینگ ژنها صورت گیرد پس وجود بانکهای DNA برای بررسی و تجزیه و تحلیل پروژه های ژنومی و مشخص کردن ژنهای عامل صفات کمی و کیفی اساسی می باشد. مشخص نمودن توالی واقعی ژنها بخاطر داشتن مناطق اینترون و گزون و تغییرات بعد از ترجمه در سطح mRNA (ویرایش) مشکل است اما امروزه به کمک تکنیکهای EST، STS، SCAR، SNP، SAGE، MICROARRAY و ... می توان وجود ژن، سطوح بیان، توالی آنها را مشخص کرد. و با استفاده از



تکنیکهای MICROARRAY و PROTEOMICS و ... می توان ژنهای عامل صفات مختلف را مشخص و اقدام به تعیین توالی و کلون کردن آنها نمود(۸).

- 1- NCBI www.ncbi.nlm.nih.gov
 - 2- DDBJ www.ddbj.nig.ac.jp
 - 3- EBI (Embl) www.ebi.ac.uk
- سایت بیوانفورماتیک ایران <http://ibc¹.ut.ac.ir>

بحث کلی

در رشته های کشاورزی از جمله زراعت و اصلاح نباتات، بیوتکنولوژی، علوم دامی، گیاه پزشکی، باغبانی و ... با توجه به هدف مورد نظر(مقاومت به آفات و بیماریها، علف کشها، خشکی، شوری، عملکرد و ...) و با کمک بانکهای اطلاعاتی که حاوی توالی تمام پروتئینها و ژنهایی که تا بحال کشف شده اند می باشند(NCBI, DDBJ, EMBL, PDB, SWISS-PROT و ...) می توان ژنها و پروتئینهای عامل این صفات کشف شده را پیدا کرد و اقدام به سنتز این قطعات با روشهای مصنوعی جهت کلونینگ نمود و یا به تعیین نواحی موتیف، تعیین نواحی حفظ شده، طراحی پرایمر و طراحی دیجنریت پرایمرها نمود و حضور این ژنها را در محصولات زراعی ایران(گندم، جو، برنج، پنبه، سویا، ذرت و ...)، محصولات باغی (سیب، گردو، پسته، انگور و...)، گیاهان دارویی(کاسنی، گل گاو زبان، کاسنی و ...) و ...مورد بررسی قرار داد و اقدام به برنامه ریزی مفیدی جهت مشخص نمودن ژنهای عامل صفات مختلف در محصولات زراعی، باغی و ...ایران کرده و عمل کلونینگ این ژنها و تولید واریته های مناسب شرایط ایران نمود.

¹ IRAN's BIOINFORMATIC CENTER



منابع:

- 1- Andreas D. Baxeavanis (2002); The Molecular Biology Database Collection: 2002 update. Publisher by oxford university press. Nucleic Acids research. Vol. 30. No. 1
- 2- Bioinformatics (2002) edit by: S.C. Rastogi
- 3- Bioinformatics (2002) edit by: Mohammad Ali Malbobi
- 4- Bioinformatics (2003) edit by: Andreas D. Baxeavanis
- 5- Burkhard Rost (2001). Protein Structure Prediction in 1D, 2D, and 3D. European Molecular Biology Laboratory, Heidelberg, Germany
- 6- Dennis A. Benson (2002), Ilene Karsch-mizrachi; GenBank. Nucleic Acids Research, Vol. 30. No. 1
- 7- Genbank: the Nucleotide Sequence Database (2003) by: Ilene Mizrachi
- 8- Gregory D. Schuler (1997). Sequence mapping by electronic pcr
- 9- Magali Lescot, Patrice Dehais (2002). Plant CARE, a database of plant cis-acting regulatory element and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research. Vol. 30. No. 1
- 10- Ron D. Appel and Denis F. Hochstrasser (2002). Computer Analysis of 2-D images. Vol 112: 2-D Proteome Analysis Protocols
- 11- Pennington s.r and Dunn m.j (2001): Proteomics from protein sequence to function. BIOS scientific publishers. Oxford. P 43-60 and 131-146
- 12- Stephan F. Altschul, Warren Gish (1990). Basic Local Alignment Search Tool. J. mol. Biol. 215. 403-410
- 13- www.ncbi.nlm.nih.gov/ A science primer SNPs: Variation on a theme (2003)
- 14- www.ncbi.nlm.nih.gov/ A science primer ESTs: Gene Discovery Made Easier (2003)
- 15- www.ncbi.nlm.nih.gov/ A science primer Micro arrays: chipping away at the mysteries of science and medicine (2003)



A review to bioinformatics and Application to agriculture

***Fazelienasab, Bahman1. Mehrabi, Ali-Ashraf2**

1-Msc in Agricultural biotechnology. Fazelie58@yahoo.com

2-Assistant Professor at Ilam University

Abstract

In previous decade, that was high grows of investigation in field molecular and cellular biology and biotechnology and every year more information add to it. That year, we witnessed to research genomic projects in human, rice, Arabidopsis and ... that has given much content information to researcher, has added other enquiry of genomic, and has opened ways for many techniques such as DNA micro array and proteomics. With much information, databanks would need for analysis of discovered new sequence protein and genes. In agricultural sciences with assist of techniques QTL, proteomics, and molecular marker and ... by use of databases and with analysis of genomic projects such as EST, STS, and SNP and... to discovering new sequences of protein and genes and to finding relationship between genomic and morphological markers and also to assist projects of genomic to prepare primer result in gene amplification or specific region of DNA for finding polymorphism in plant and animal species, gene cloning and transformation. Till 2001 there is 281-335 data bank in world and will add to them and till 2002 has related genome sequence more 800 organism and have placed 3.5 million sequence to databanks

Key words: bioinformatics, databanks, molecular markers, agriculture sciences