

11. KÜTLE SPEKTROMETRESİ VERİLERİNİN BİYOİNFORMATİK ANALİZİ VE PROTEİNLERİN KİMLİKLENDİRİLMESİ TANIMLANMASI(MASCOT-MATRIX SCIENCE)

A. BİLGİ

Omik teknolojilerinin gelişmesiyle, kullanılan yöntemler sonucunda çok fazla bilgi üretilmektedir ve bu bilgilerin işlenmesi için üst düzey bilgisayarların kullanılması günümüzde bir gereklilik değil zorunluluk olmuştur.

Protein ifade profillerinin çıkarılmasında ve protein tanımlamasında ilk adım genelde 2D-PAGE ile başlar ve böylece proteinler izoelektrik noktalarına ve moleküler ağırlıklarına göre ayrılmış olurlar. Ayrılan proteinler gümüş, Coomassie Blue veya florasan boyalarla boyanarak görünür hale getirilirler ve genelde jel içinde spesifik proteazlarla (örneğin tripsin) peptidlere parçalanırlar. Elde edilen peptidler jel parçalarından ekstrakte edilir ve kütle spektrometresi (ör: MALDI-TOF MS) ile analizlenir. Bir proteine ait peptidlerden elde edilen kütle spektrumu, o proteinin parmak izini oluşturur. Bu yöntemle proteinler femtomol seviyesinin altında duyarlılıkla tanımlanabilirler.

Deneysel olarak elde edilen kütle profilleri, veri tabanlarında bulunan aynı proteinin aynı enzimle bilgisayar simülasyonu ile hazırlanmış teorik kütle profilleri ile karşılaştırılır. Veri tabanında bulunan proteinler, eşleşen peptid oranlarına göre belirli bir kütle hata toleransıya sıralanırlar. Ancak bir proteinin parmak izi beklenen tüm teorik peptid kütlelerini içermez. Ayrıca tripsin otolizinden ve keratin kontaminasyonundan hatalı pikler oluşabilir. Başarılı bir protein tanımlaması için, MALDI piklerinin doğruluğunun yüksek olması ve zengin bir veri tabanı kullanılması gerekmektedir. Bir proteinin başarı ile tanımlanması için, beş ya da daha fazla peptidin 30 ppm'den düşük hassasiyet ile belirlenmesi, deneysel ve teorik dizinlerin % 15 oranında eşleşmesi gerekmektedir (Canas et al 2006).

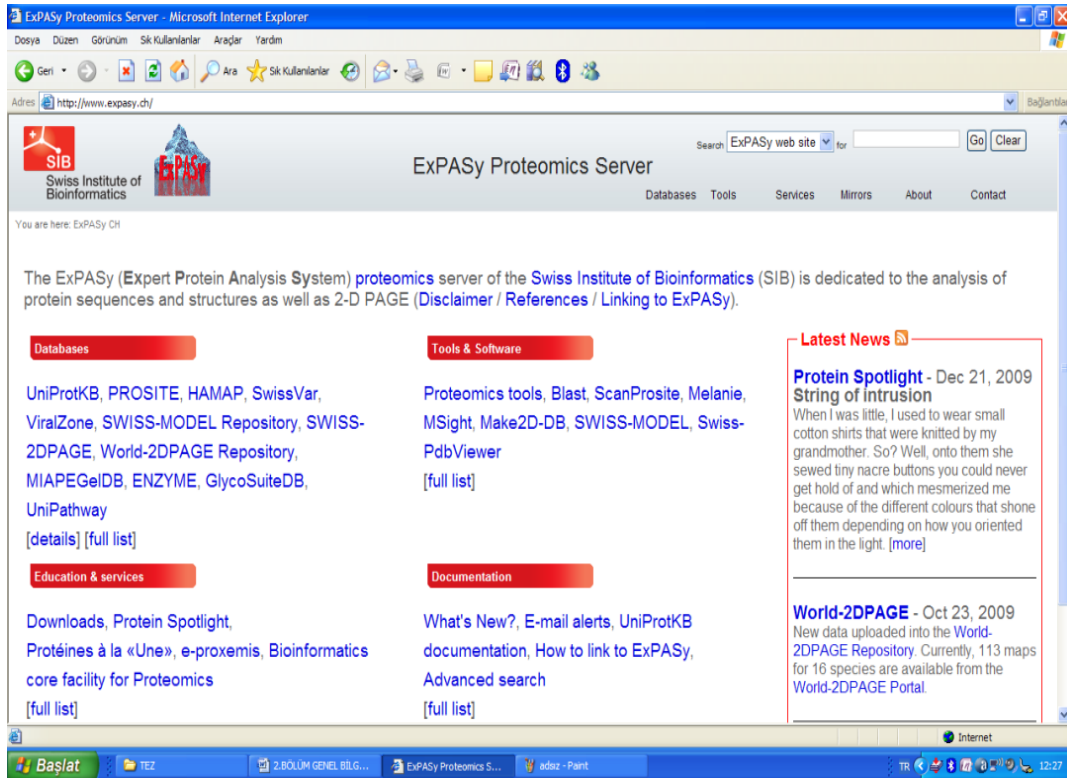
1993 yılından bu yana, peptid kütle parmak izi (PMF) ve MALDI-TOF kombinasyonu en hassas, hızlı ve güvenilir protein tanımlama yöntemi olmuştur. Son yıllarda tanımlanan proteinlerin çoğu bu yöntem kullanılarak tespit edilmiştir. Ancak daha sonra gelişen MS/MS sistemleriyle günümüzde daha fazla peptit üzerinden daha yüksek bir yüzdeyle tanımlamalar gerçekleştirilmektedir. Ayrıca bu sistemlerin öncesinde kullanılan kromatografik ayırım basamağı jel dökülmeden kalitatif ve kantitatif analiz yapılmasına olanak sağlamaktadır.

PMF ile protein tanımlanmasında iki yaklaşım vardır: (1) "bottom-up", (2) "top-down". "Bottom-up" yaklaşımında peptitlerden yola çıkılarak veritabanları ayırdımıyla protein bilgisine ulaşılır ki günümüzde en çok tercih edilen yaklaşımdır. "Top-down" yaklaşımında ise direkt olarak proteinin kendisinden yola çıkılır.

Verilerin analizinde protein veritabanlarından faydalanılır. UniProtKB/Swiss-Prot ve NCBI altında bu amaçla kullanılabilecek veritabanı mevcuttur ve tanımlamalar buradaki bilgiler

ışığında gerçekleştirilir. Programlar ve arama motorları bu veritabanlarındaki bilgileri kullanırlar.

Swiss-Prot geliştiricileri tarafından hazırlanan ExPASy (Expert Protein Analysis System) (<http://www.expasy.ch>) adresinde UniProtKB/Swiss-Prot veri tabanındaki protein bilgisinde arama yapılabilmekte ve arama sonucunda protein hakkında çok detaylı bilgilere ulaşıp sanal kesim yapılabilir. Tüm internet kullanıcılarına açık olan bu platformda faydalı programların yanı sıra, farklı örneklerden elde edilen 2 boyutlu jel görüntülerini bu bunlar üzerinden tanımlanan proteinleri gösteren interaktif bir uygulama da mevcuttur. Veri tabanında haptoglobulin proteininin triptik peptid dizileri, kütleleri, pI ve Ma değerleri şekilde gösterilmiştir.



ExPASy internet ana sayfası görünümü.

You have selected HPT_HUMAN (P00738) from UniProtKB/Swiss-Prot:

Haptoglobin precursor [Contains: Haptoglobin alpha chain; Haptoglobin beta chain]
Signal in positions 1-18 has been removed.

- Chain Haptoglobin at positions 19 - 406 [Theoretical pI: 6.13 / Mw (average mass): 43349.01 / Mw (monoisotopic mass): 43321.52]

mass	position	#MC	modifications	peptide sequence
3817.6486	346-379	0		YQEDTCYGDAGSAFAVHDLE EDTWYATGILSFDK
3292.5174	19-49	0		VDSGNDVTDIADDGCPKPPE IAHGVEHSVR
2848.3028	83-108	0		LPECEADDGCPKPPEIAHGY VEHSVR
2679.3923	179-202	0		MVSHHLTTGATLINEQWLL TTAK
2115.0361	326-345	0		SPVGVQPILNEHTFCAGMSK
1795.0112	236-251	0		VVLHPNYSQVDIGLIK
1650.7978	298-311	0		YVMLPVADQDQCIR
1458.7335	203-215	0		NLFLNHSENATAK
1439.6648	119-131	0		TEGDGVYTLNNEK
1311.6063	60-71	0		TEGDGVYTLNDK
1290.7303	216-227	0		DIAPTLTLVVGK
1288.6242	380-391	0		SCAVAEYGVVVK
1273.6279	142-153	0		LPECEAVCGKPK
1203.6368	392-401	0		VTSIQDWVQK
1146.5425	312-321	0		HYEGSTVPEK
987.5365	262-270	0		VMPICLPSK
980.4948	278-286	0		VGYSVSGWGR
923.5308	162-170	0		ILGGHLDK
920.4624	171-178	0		GSFPWQAK
895.4744	154-161	0		NPANPVQR
858.4931	229-235	0		QLVEIEK
809.3788	271-277	0		DYAEVGR
760.3988	292-297	0		FTDHLK
703.3733	256-261	0		VSVNER
688.3777	73-77	0		QWINK
688.3777	132-136	0		QWINK
593.3042	287-291	0		NANFK
587.2824	54-57	0		NYK
587.2824	113-116	0		NYK

UniProtKB/Swiss Prot veri tabanında Haptoglobin proteini için peptid dizileri, kütleleri, pI ve Ma değerleri

Kütle spektrometresi verilerinin analizlerinde çeşitli firmaların geliştirdiği lisanslı programlar mevcuttur. Bu programlardan biri olan ProteinLynx Global Server (PLGS, Waters) ile elde edilen spektrumlar farklı parametrelerle direkt olarak analiz edilebilir ve ayrıca sisteme özel veritabanı yüklenip sadece bu veritabanı üzerinden de arama yapılabilir.

Lisanslı programların dışında, üniversitelerin veya bazı kurumların geliştirdiği ve internet üzerinde erişilebilen bazı arama motorları da mevcuttur. Bunlardan önemli olan bazıları şunlardır:

Arama motorunun adı	Adresi	Yapılabilen analiz
MASCOT	http://www.matrixscience.com/	MS, MS/MS
Aldente	http://www.expasy.ch/tools/aldente	MS
PepMAPPER	http://www.nwsr.manchester.ac.uk/mapper/	MS
Profound	http://prowl.rockefeller.edu/prowl-cgi/profound.exe	MS
Protein Prospector	http://prospector.ucsf.edu/prospector/mshome.htm	MS, MS/MS
OMSSA	http://pubchem.ncbi.nlm.nih.gov/omssa/	MS/MS

B. UYGULAMA

MASCOT ve Aldente ile PMF veri analizi.

Swiss-Prot'tan özel veritabanı oluşturulması, PLGS'e yüklenmesi ve arama yapılması.

Kaynaklar

Özel Duygu Demiralp, İğci Naşit, Peker Selen, Ayhan Beycan, 2014, Temel Proteomik Stratejiler, Ankara Üniversitesi Yayınevi, ISBN: 978-605-136-148-2, 2014

Özel Duygu Demiralp, İğci Naşit, Peker Selen, Ayhan Beycan, 2013, Temel Proteomik Stratejiler, Ankara Üniversitesi Yayınevi, ISBN: 978-605-136-117-8, 2013.