

Univerzita Karlova v Praze

Přírodovědecká fakulta

Studijní program: Biologie

Studijní obor: Biologie



Martina Nováková

Fosforylace a její vliv na vznik rakoviny – bioinformatický pohled
Phosphorylation and its effect on cancer development – bioinformatics view

Bakalářská práce

Školitel: Mgr. Marian Novotný, PhD.

Praha, 2013

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 22. 8. 2013

Podpis

Poděkování

Chtěla bych poděkovat především svému školiteli Mgr. Marianu Novotnému, Ph. D. za odborné rady, konzultace a motivaci. Neméně děkuji svým rodičům, kteří mě neustále podporují. A slečně Petře Štěpánové, že se odvážíla hledat chyby tam, kde já tu odvahu neměla.

Obsah

Poděkování.....	3
Abstrakt.....	5
Abstract.....	6
Seznam zkratek.....	7
Úvod.....	8
Posttranslační úpravy.....	8
Fosforylace.....	8
Kinázy a fosfatázy.....	10
Kinázy.....	10
Fosfatázy.....	12
Fosforylace jako přepínače:.....	13
Rozdělení kináz a fosfatáz.....	14
Nádorová onemocnění.....	17
Možnost léčby pomocí kinázových inhibitorů.....	18
Získávání a ukládání dat.....	19
Data experimentální.....	19
Hmotnostní spektrometrie.....	19
2D elektroforéza.....	20
Databáze.....	21
Databáze posttranslačních modifikací.....	21
Databáze enzymů.....	25
Databáze týkající se nádorů.....	25
Data predikční.....	26
Analýza fosforylací v nádorových onemocněních.....	28
Úbytek fosforylačních míst a jejich vznik de novo.....	29
Závěr.....	31
Zdroje.....	32

Abstrakt

Nádorová onemocnění jsou častou příčinou smrti a jen v České republice ročně onemocní přes 70 tisíc lidí a více než 25 tisíc zemře. Tato onemocnění způsobují změnu buněčného metabolismu a jedním ze způsobů, jak se toto děje, je změna posttranslačních modifikací, zejména fosforylací.

Fosforylace jsou důležitý regulační mechanismus, který má přímý vliv na aktivitu enzymů. Fosforylované zbytky jsou zpravidla serin, threonin a tyrosin, fosforylacemi na těchto aminokyselinách se budu v bakalářské práci zabývat.

Na problematiku fosforylací ve spojitosti s nádorovými onemocněními se snažím nahlédnout z bioinformatické pozice, zejména se soustředuji na databáze shromažďující data o fosforylacích spojených s nádorovými onemocněními.

Data týkající se fosforylace proteinů se získávají pomocí experimentálních metod, jako je hmotnostní spektrometrie, ale také in silico, jako výsledky složitých výpočtů počítačů.

Databáze se mezi sebou liší typem dat, který shromažďují, velikostí, způsobem anotování a kontrolou obsahu.

V práci podávám přehled nejdůležitějších databází a představuji práce, které bioinformatickými přístupy analyzují data o fosforylacích a nádorech.

Klíčová slova: fosforylace, proteiny, rakovina, databáze, hmotnostní spektrometrie, bioinformatika

Abstract

Cancer is a frequent disease and cause of human death, during one year over 70 000 people suffer and over 25 000 die because of cancer in the Czech Republic. Cancer causes change of cellular metabolism and one way this is done is the change of post-translational modifications, especially phosphorylations.

Phosphorylations are important regulatory mechanism, because they often have a direct effect on the enzyme activity. Phosphorylated residues are mostly serine, threonine and tyrosine and phosphorylation of these amino-acids are considered in this bachelor thesis.

Data about phosphorylations are then stored in the database. The databases differ by the type of data they collect, size, way of annotating and content control.

My thesis deals with the connection between cancer and phosphorylation. The emphasis of this thesis is on bioinformatic approaches to study changes in phosphorylations in cancer cells.

The thesis is a summary of the most important databases. This work also summarizes articles with a bioinformatic approach to analyze data on phosphorylation and tumours.

Keywords: phosphorylation, proteins, cancer, database, mass spectrometry, bioinformatics

Seznam zkratek

ANN	Neuronové sítě (Artificial neural networks)
ATP	Adenosintrifosfát
CDK	Cyklindependentní kináza
EGFR	Receptor epidermálního růstového faktoru
GIST	Gastrointestinální stromatární tumor
HER	Lidský gen pro EGFR
HMM	Skryté Markovovy Modely (Hidden Markov Models)
MAP	Mitogenem aktivovaná proteinkináza
MEK	kináza MAP kinázy
OMIM	Online Mendelian Inheritance in Man
PDB	Protein DataBank
PGDFR	Receptor růstového faktoru odvozeného od krevních destiček
PKA	Proteinkináza A
PKC	Proteinkináza C
PKG	Proteinkináza G
PTM	posttranslační modifikace
PTP	Fosfotyrosin fosfatáza
SNP	Jednonukleotidový polymorfismus (Single nucleotide polymorphism)
SNV	Jednonukleotidová variace (Single nucleotide variation)
TKI	tyrosinkinázový inhibitor

Úvod

Posttranslační úpravy

Po translaci RNA na polypeptid dochází velmi často k jeho dalším posttranslačním modifikacím, než se stane plně funkčním proteinem. Posttranslační modifikace lze rozdělit na vratné a nevratné. Nevratnou posttranslační modifikací je odstranění části proteinu, například vyštěpení části řetězce u trypsinu.

Vratné modifikace jsou spíše regulační povahy, ať již regulují aktivitu, efektivitu a dobu existence proteinu, nebo vlastnosti chemicko-fyzikální, například náboj, hydrofobicitu. Methylace, acetylace, hydroxylace, glykosylace jsou modifikace, které mění spíše chemicko-fyzikální vlastnosti proteinu. Příkladem je modifikace histonů, pokud se na lysiny v N koncové části váží kyselé acetylové skupiny odpuzující DNA, která je namotaná na histonech, nukleosom se tak rozvolní. Acetylace na N konci proteinu (na první aminokyselině) je ireverzibilní a není přesně známa její funkce, ale dlouho se soudilo, že je to modifikace prodlužující životnost proteinu. V poslední době se zjistilo, že tato acetylace může směřovat protein do degradační dráhy (Arnesen 2011). Zásadní vratnou regulační modifikací je fosforylace (a defosforylace) fungující za prvé jako přepínač mezi aktivním a inaktivním stavem proteinu a za druhé slouží v některých případech ke zvýšení či snížení enzymatické činnosti.

Fosforylace

Fosforylace je vratná modifikace pomocí zbytku kyseliny fosforečné. Je možné fosforylovat více substrátů, ale ve své práci se budu zabývat fosforylací proteinů. Fosforylace je možná na kyslíku tří aminokyselin - tyrosinu, serinu a threoninu. Fosforylující enzymy, proteinkinázy, jsou zpravidla dvou typů, kinázy fosforylující na serinech a threoninech a kinázy fosforylující na tyrosinu, ovšem existují i výjimky schopné fosforylace na serinu/threoninu a tyrosinu, takzvané duální proteinkinázy. Proteiny provádějící defosforylaci se nazývají fosfatázy, tyto enzymy mají také různou míru specifity.

Prokaryotické organismy mají také histidinové a aspartátové kinázy, kde ovšem na rozdíl od tyrosinových a serin/threoninových kináz dochází k fosforylaci na dusíku, aspartátové kinázy živočichové nemají. Histidinové a jiné dusíkové kinázy mají i živočichové, ale tyto kinázy nejsou s prokaryotními dusíkovými kinázami příbuzné (Klumpp, Krieglstein 2002).

Během fosforylace je napadena vazba mezi fosforem gama fosfátového zbytku ATP a kyslíkem β fosfátového zbytku. Následně vzniká esterová, kovalentní vazba mezi fosforem z fosfátového zbytku, přeneseným z ATP, a kyslíkem aminokyselinového zbytku na proteinu. Produkty této reakce jsou ADP a fosforylovaný protein.

Během defosforylace dochází k hydrolýze vazby mezi fosfátem a proteinem, vzniká anorganický fosfát.

Fosforylace probíhá v prokaryotních i eukaryotních buňkách po celou dobu jejich života. Pomocí signálních kaskád umožňuje přijímat podněty z okolí i z buňky samotné a pomocí kináz a fosfatáz na ně adekvátně odpovídat, 30 až 50 % proteinů buňky je během jejího života fosforylováno (Pinna, Ruzzene 1996). Fosforylace je klíčová pro buněčné dělení a celý buněčný cyklus, z čehož je patrné, že poruchy a špatná interpretace signálu mohou být fatální nejen pro buňku samotnou, ale i pro celý organismus.

Kinázy a fosfatázy

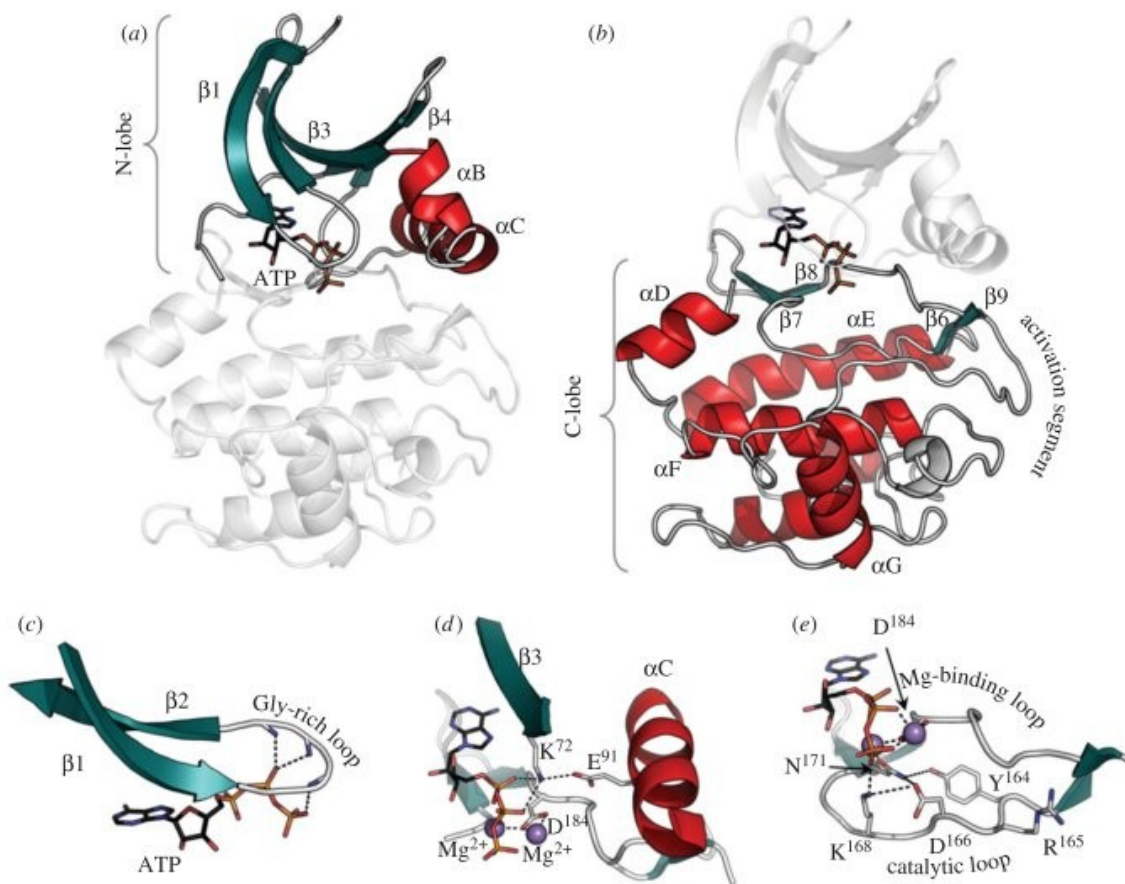
Fosforylaci proteinů zajišťují enzymy proteinkinázy, tyto enzymy se řadí, podle nomenklatury zavedené Mezinárodní biochemickou a molekulárně biologickou unií, mezi transferázy. EC 2.7.10 mají tyrosin proteinkinázy, EC 2.7.11 serin/threonin proteinkinázy, EC 2.7.12 duální proteinkinázy. Zatímco proteinfosfatázy, které katalyzují hydrolýzu vazby mezi fosforem a kyslíkem, se řadí mezi hydrolázy, EC 3.1.3.16 serin/threoninové proteinfosfatázy, EC 3.1.3.48 tyrosin proteinfosfatázy, duální fosfatázy jsou označeny EC 3.1.3.16 a zároveň EC 3.1.3.48.

Kinázy

Kinázová doména eukaryotních proteinkináz má tři úlohy: Vazbu a orientaci fosfátového donoru (ATP nebo GTP) pomocí komplexu s dvojmocnými kationty. Vazbu a orientaci substrátu. Přenesení gama fosfátu z ATP na hydroxylový zbytek substrátu (Hanks, Hunter 1995).

Kinázová doména má ustálené uspořádání, skládá se z N koncového laloku a C koncového laloku. N terminální lalok je tvořen pěti antiparalelními β -hřebeney a dvěma α -helixy, mezi prvním a druhým β -listem je na glycin bohatá smyčka, která zajišťuje správnou orientaci ATP. Mezi třetím a čtvrtým β -hřebenem se nachází jeden nebo dva α -helixy, některé proteinkinázy mají zachovány pouze jeden α -helix (označovaný α C), který je ovšem nezbytný. C koncový lalok je tvořen převážně α -helixy, ale kromě nich také čtyřmi β -hřebeney. Mezi β 6 a β 7 se nachází katalytická smyčka, která směřuje fosfát na substrát a mezi β 8 a β 9 je smyčka, která určuje pozici hořčíkových iontů (Obrázek 1) (Taylor et al. 2012).

Vazebná energie pro fosforylaci na serinu/threoninu je podobná energii pro vazbu fosfátu na tyrosin. Fosfoserin a fosfothreonin hrají spíše strukturní roli, zatím co fosfotyrosin je určen k regulaci. Tyrosinkinázy jsou silně negativně regulované a aktivují se pouze za určitých podmínek. Fosfotyrosinfosfatázy mají vysokou rychlost přeměny, což vede k rychlému odbourání fosfotyrosinu, pokud není chráněn vazbou k SH2 doméně (Hunter 2009).



Obrázek 1: Stavba kinázové domény eukaryotické proteinkinázy (PKA), N koncový lalok (a) a C koncový lalok (b) a důležité funkční motivy: glycin bohatá smyčka (c), konzervované zbytky v $\beta 3$ a αC (d), katalytická smyčka (e). (Taylor et al. 2012)

Vazba mezi SH2 doménou a aminokyselinovým zbytkem je díky aromátovému kruhu velmi silná, což je důvod, proč je tyrosinová fosforylace použita jako regulační prvek.

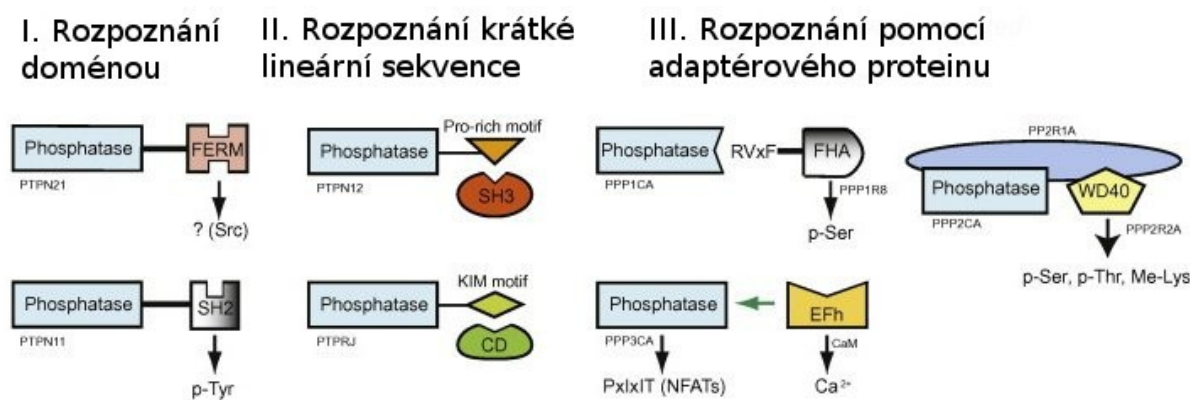
Tyrosinkinázy jsou zpravidla dvojího typu. Receptorové tyrosinkinázy, které mají extracelulární a intracelulární doménu, a volné tyrosinkinázy, které mají domény vázající se na receptory bez enzymatické aktivity. Receptory, ať už s enzymatickou aktivitou či bez ní, jsou aktivovány ligandy z extracelulárního prostředí, například růstovým faktorem. Následně dochází k dimerizaci dvou kináz a vzájemné fosforylaci, pak se přenáší signál dál fosforylací dalších proteinů, často také kináz. Přenos signálu z tyrosin kináz na serin/threoninové kinázy byl původně považován za důvod existence kináz s dvojnou aktivitou.

Kinázy s dvojnou aktivitou, tedy takové, které jsou schopny fosforylovat jak serinový nebo threoninový zbytek, tak i zbytek tyrosinový, nejsou příliš časté. Mnohdy bývá dvojná aktivita buď autofosforylační, nebo probíhá velmi zřídka. Existují kinázy, které jsou tak schopny činit za velmi přísně daných podmínek. Jediná běžně fungující kináza s dvojnou aktivitou je MEK (kináza MAP kinázy), jež fosforyluje tyrosin a threonin v konzervované sekvenci v MAP kináze (Pinna, Ruzzene 1996).

Fosfatázy

Fosfatáz je obecně mnohem méně než kináz, například kvasinky *Saccharomyces cerevisiae* mají ve svém genomu 1,9 % protein kináz, ale pouze 0,23 % proteinfosfatáz (Hunter, Plowman 1997). Což neznamená, že jsou fosfatázy nedůležité. I přes jejich malý počet může absence či nadbytek proteinfosfatáz způsobit poruchu funkce buňky. Pokud by například neexistovala fosfatáza, která by defosforylovala aktivovanou glykogen syntetázu, došlo by k vyčerpání buňky. Regulační charakter tyrosinové fosforylace je zdůrazněn faktem, že přestože je obecně všech fosfatáz méně než všech kináz, specifických tyrosinových fosfatáz je více než tyrosinkináz (Mustelin 2007).

Fosfatázy jsou při pokusech in vitro nespecifické, co se týká substrátu. V živých systémech jsou naopak velmi silně specifické, nejspíše proto, že se v in vitro systémech vyskytují další substráty, které nejsou in vivo dostupné. Fosfatázy používají tři mechanismy pro rozpoznání substrátu (Obrázek 2). V prvním případě obsahují fosfatázy doménu, rozpoznávající cílové proteiny, například SH2 doménu. Druhý způsob je rozpoznání a navázání krátkého lineárního motivu fosfatázy cílovým proteinem. Třetí způsob je nekovalentní vazba na adaptérový protein, který zajišťuje rozpoznání substrátu (Sacco et al. 2012).



Obrázek 2: Způsob rozpoznání substrátu fosfatáz (Sacco et al. 2012)

Fosforylace jako přepínače:

Fosforylace se vyvinula jako regulační mechanismus již u prokaryot, spolu s fosforylací se vyvíjely i kinázy a fosfatázy. Kinázy prokaryot mohou být homologní k eukaryotním kinázám, ale mnoho kináz existujících u prokaryot se nevyskytuje u eukaryot, případně jen u specifické skupiny.

U prokaryotních organismů se často uplatňuje regulace pomocí dostupnosti substrátu či produktu metabolismu, ovšem u eukaryotních organismů se kvůli zvětšení samotné buňky a její kompartmentaci, i postupnému přechodu od jednobuněčných organismů k mnohobuněčnosti muselo přejít na složitější formy regulace.

Fosforylace na tyrosinu se jako regulační prvek běžně vyskytuje u zvířat (kmen *Metazoa*). U kvasinek (*Saccharomyces*) byly nalezeny pouze duální kinázy v kaskádě MAP kinázy, receptové a src-like kinázy nikoliv. U rostlin se také není znám výskyt receptorových a nereceptorových (src-like) kináz. Mezi houbami (*Fungi*) mají pouze hlenky (*Dictyostelium*) protein s SH2 doménou a využívají signalizace pomocí fosfotyrosinu (Darnell 1997).

Fosfátová skupina má poměrně velký náboj, je kyselá a zabírá významně prostor, což jí umožňuje působit na bližší i širší okolí a měnit tak mírně terciární nebo kvartérní strukturu proteinu, což může vést k jeho aktivaci, inaktivaci či možnosti rozpoznání jako substrátu jiného enzymu. Esterová vazba fosfátu má navíc dobré chemické vlastnosti, vysoká energie vazby neumožňuje její spontánní defosforylaci. Vazba mezi proteinem a fosfátem, obecně esterová vazba s fosfátem, je ve vodném roztoku při neutrálním pH vysoce stálá (Hunter 2012), v kinázách a fosfatázách je tato vazba narušena pomocí postranních řetězců aminokyselin a iontů kovu v jádře enzymů.

Nejméně pravděpodobný výskyt regulačních fosforylací je v uspořádaných strukturách, čím neuspořádanější a delší smyčky jsou, tím pravděpodobnější je výskyt regulačních fosforylací. Vysoký výskyt fosforylací je v dlouhých neuspořádaných vláknitých strukturách proto, že jsou aminokyselinové zbytky více vystaveny okolnímu prostředí a snadněji přístupné kinázám a fosfatázám. V dlouhých neuspořádaných vláknech je mnohonásobný výskyt fosforylačních míst a regulace aktivity proteinu je tedy mnohem více variabilní než pouze dvoustavový přepínač mezi aktivním a neaktivním stavem. Příkladem může být retinoblastoma protein (Rb), kde fosforylace ovlivňují jeho vazbu na transkripční faktory a další regulační proteiny. Například jedna z fosforylací snižuje

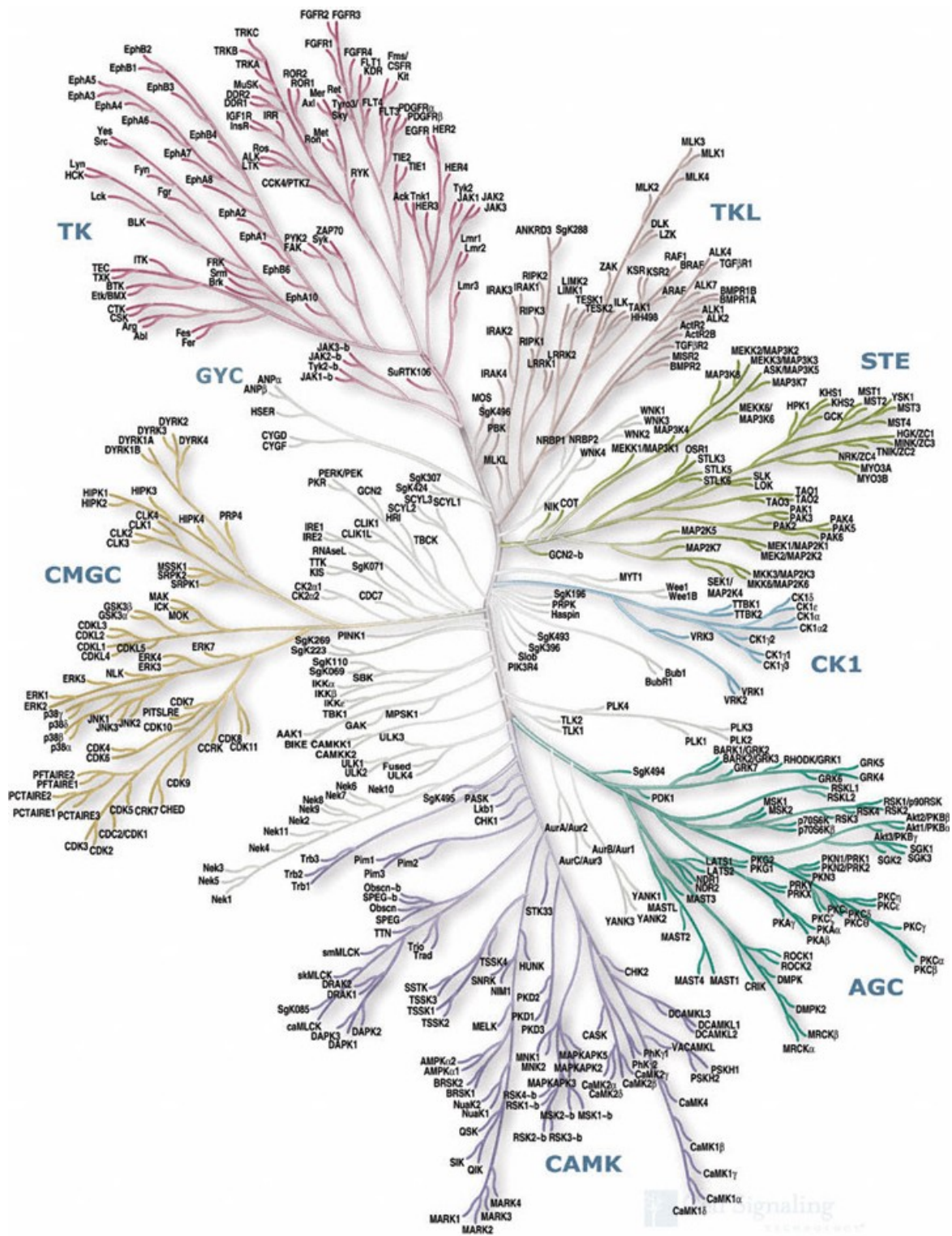
afinitu Rb k transkripčnímu faktoru E2F pomocí přestavby vazebné domény a další fosforylace ve vazebné doméně, zároveň zcela blokuje vazbu na E2F faktor (Tyanova et al. 2013).

Rozdělení kináz a fosfatáz

Nadrodinu (superfamily) eukaryotních proteinkináz rozdělili Hanks a Hunter (Hanks, Hunter 1995) do rodin pomocí analýz aminokyselinových sekvencí katalytické části proteinkináz. Rodiny jsou sdružené ve skupinách takto:

1. A-C-G skupina, jež obsahuje mimo jiné rodinu kináz regulovaných cyklickým nukleotidem (kinázy PKA, PKG), rodinu kináz diacylglycerol aktivovaných/ fosfolipid-dependentních (kináza PKC), rodina kináz příbuzných PKA a PKC (RAC) a rodinu kináz fosforylující s receptory spřažené G - proteiny
2. CAMK skupina obsahuje dvě rodiny, rodinu kináz regulovaných Ca^{2+} /Calmodulin a blízce příbuzné kinázy. Druhá rodina je Snfl/AMPK (AMP aktivovaná proteinkináza)
3. C-M-G-C skupina obsahuje především rodiny CDK a příbuzných kináz, jež obsahuje cyklin-dependentní kinázy a Cdc kinázy (cell division cycle – kinázy buněčného cyklu) a jejich homology, a rodinu MAP kináz (Erk). Dále rodinu glykogen syntetázy aktivujících kináz 3, a Clk rodinu (kináza „Cdc-like kinase“ a jí příbuzné)
4. Konvenční tyrosinové kinázy. Zde je podstatné především dělení na receptorové kinázy a kinázy nereceptorové. Kinázy, podle kterých se následně pojmenovaly rodiny v této skupině jsou často pojmenovány podle patologických stavů, kdy jsou exprimovány, či jsou aktivní. Například Brk rodina je pojmenována podle Brk kinázy, exprimované v nádoru při rakovině prsu.
5. Ostatní, do skupin nezařazené kinázy.

Toto rozdělení bylo rozšířeno o další skupiny (Obrázek 2). Z databází lidských genů byly pomocí skrytých Markovových modelů vybrány ty geny, které odpovídaly profilu eukaryotní proteinkinázové domény. Genů eukaryotických proteinkináz bylo nalezeno v lidských genech 478, došlo také k objevení 13 atypických proteinkináz (Manning et al. 2002).



Obrázek 3: Dendrogram ekaryotických proteinkinázových domén, vyznačeno 8 hlavních skupin. (Manning et al. 2002)

Serin/threoninové kinázy jsou častější nežli tyrosinové, tvoří průměrně 85-90 % kinomu (souboru všech kináz) živočicha (Hunter 2009). Lze je rozdělit podle charakteru okolí místa, které fosforylují, a to takto (Pinna, Ruzzene 1996):

1. Basofilní kinázy rozpoznávají zásadité zbytky v okolí fosforylovaného místa, preferují lysiny, také mohou rozpoznávat zbytky hydrofobní, jejichž význam dokonce někdy převládá nad zbytky basickými. Bazofilní je většina kináz patřících do skupiny ACG kináz a až na pár výjimek celá skupina CaMC kináz.
2. Prolinem naváděné kinázy, vyžadují prolin v pozici n+1. Prolin v pozici n+1 je pro ostatní kinázy zpravidla kritérium vylučující fosforylaci. Prolinem naváděné kinázy mohou mít další požadavky specifické pro rodiny kináz, například přítomnost bazických či kyselých zbytků v okolí. Prolinem naváděné jsou rodiny cyklin dependentních kináz a MAP kináz (CMGC skupina).
3. Acidofilní kinázy a kinázy vyhledávající místa s předchozí fosforylací. Takovýchto kináz není mnoho, ale za to jsou hojně rozšířené. Patří sem kinázy z různých třídění podle Hankse a Huntera, mimo jiné kaseinové kinázy (CK2 rodina z CMGC skupiny, CK1 rodina zařazena v ostatních proteinkinázách)

Rodiny fosfatáz (Tabulka 1) podle Mustelina (Mustelin 2007).

Fosforyl. AK	rodina	podrodina	zástupce
S/T	PPM		PP2C
	FPC		FPC
	PPP		PP1, PP2A, kalcineurin
Y,S	HAD		Eya, CTD, cronophin
Y	Třída I Cys-založených PTP	klasické PTP	CD45, PTP1B
		VH1-like a duální fosfatázy	MKP
	Třída II Cys-založených PTP		CDC25A, CDC25B, CDC25C
	Třída III Cys-založených PTP		LMPTP

Tabulka 1: Rozdělení proteinových fosfatáz do rodin s příklady zástupců.

PPM a FCP rodiny jsou Mg^{2+} dependentní. HAD rodina (Asp-založená) má velmi širokou škálu substrátu - nejen fosfoproteiny, ale například i fosfolipidy, cukry či nukleotidy. Patří sem tyrosinové fosfatázy a serinové fosfatázy. Třída I Cys- založených PTP má dvě části, klasické fosfotyrosinfosfatázy a VH1-like a duální fosfatázy. Klasické PTP - obsahují transmembránové (receptorové) fosfotyrosinfosfatázy a nereceptorové fosfatázy. VH-1 like kinázy a duální fosfatázy jsou zastoupeny fosfatázou MAP kinázy. Třída II Cys-založených PTP je funkčně podobná fosfatázám v třídě I, ovšem strukturně jsou si nepřibuzné. Třída III Cys-založených PTP má v lidském genomu jediného zástupce, a to LMPTP (fosfotyrosinfosfatázu s nízkou molekulovou hmotností) (Mustelin 2007).

Nádorová onemocnění

Nádorové onemocnění je onemocnění způsobené nekontrolovaným buněčným dělením. Je to onemocnění s různým stupněm závažnosti, od méně závažných nezhoubných nádorů na kůži až po velmi závažné, jako například rakovina tlustého střeva. Pro rok 2011 bylo vyjma rakoviny kůže odhaleno celosvětově přes 12 milionů nových případů nádorových onemocnění (Jemal et al. 2011).

Rakovinné buňky s poškozenou regulací dělení mohou vzniknout ze všech buněčných typů. Při rakovinném zvrhnutí buněk epitelů může mít nádor zcela odlišné vlastnosti než nádor na kosti, mohou se mezi sebou lišit i nádory stejných tkání. Jedním z hledisek třídění může být schopnost invazivního růstu a tvorby metastáz; maligní (zhoubné) nádory jsou invazivní a metastázuje, benigní (nezhoubný) nádor nemá schopnost metastázovat a pokud roste, roste pomalu. Dále podle původnosti, je-li nádor primární či vzniklý z metastáz, podle podobnosti nádoru s tkání, která mu dala vzniknout. Mezi faktory napomáhající vzniku rakoviny patří genetické dispozice, kondice, konzumace jídla a jiných látek, chemické faktory jako karcinogeny a mutageny, fyzikální faktory např. UV záření, rentgenové a gama záření a biologické faktory dvou typů - prvním jsou infekce, ať již virové či neviróvé. Druhé jsou pochody tělu vlastní - některé rakoviny vznikají díky stimulaci hormony.

Jedním ze zásadních znaků nádorové buňky je, kromě nekontrolovaného dělení, i necitlivost vůči apoptotickému signálu. Zdravá buňka má vyvážený poměr mezi signály vedoucí k apoptóze a signály antagonistickými. Apoptotický signál přichází zevnitř buňky, například v důsledku nevratných mnohonásobných poškození DNA, i z vnějšku od buněk imunitního systému a z okolní tkáně. Apoptotický signál se předává kaskádou

proteinů, které se účastní i enzymy s kinázovou aktivitou. Rakovinné buňky naruší rovnováhu a oslabí apoptotický signál, což může vést k progresi či vzniku metastáz. Naopak jiné nádorové změny mohou podporovat apoptózu během vícestupňové karcinogeneze (Lowe, Lin 2000).

Možnost léčby pomocí kinázových inhibitorů

Objevy na poli regulace fosfotyrosinem mají praktické důsledky a již dnes existuje možnost léčby pomocí tyrosinkinázových inhibitorů. Ovšem v nádorech často bývá aktivována řada tyrosinkináz s překrývajícími se funkcemi, což může léčbu tyrosinkinázovými inhibitory komplikovat. Pokud nezabírá jeden inhibitor, je možné použít několik dalších, obdobně jako se postupuje při léčbě pacienta s HIV (Hunter 2009). Gleevec (v Evropě distribuován jako Glivec) s účinnou látkou imanitib, tyrosinkináza první generace byl povolen pro léčbu chronické myeloidní leukémie v roce 2001 (Hunter 2009), o rok později k léčbě gastrointestinálního stromatního tumoru (GIST). Inhibuje Bcr-Abl tyrosinkinázu, která vzniká translokací, při které vznikne tzv. Philadelphia chromosom; tato chromozomální abnormalita se vyskytuje při chronické myeloidní leukémii. Další kinázy, které inhibuje, jsou PDGFR a c-kit.

V roce 2006 byl povolen Sutent, s účinnou látkou sunitinibem, zároveň pro léčbu imanitib rezistentních gastrointestinálních stromatických nádorů a pro karcinomy vzniklé z renálních buněk (Wishart et al. 2008, DrugBank: Sunitinib).

Další z léčiv je Herceptin s účinnou látkou trastuzumab. Na rozdíl od výše zmíněných látek, které jsou malé molekuly (mezi 20-30 uhlíky), je trastuzumab monoklonální protilátkou, jedná se tedy o protein a používá se při léčbě rakoviny prsu. Tato protilátka se váže na HER2/neu tyrosinkinázu, protein exprimovaný z genu HER2 (Human EGFR Related gene), lidskému receptoru epidermálního růstového faktoru příbuzný gen, který se v rakovinných buňkách nadměrně exprimuje. Váže se také na EGF-receptor-like protein (Shawver et al. 2002).

Jak nádor roste, začíná podporovat angiogenezi ve svém okolí, na to se zaměřují další léčiva; zaměřují se především na receptor vaskulárního endoteliálního růstového faktoru (VEGFR2). VEGFR mají tyrosinkinázovou aktivitu. Jedním z takovýchto léčiv je Votrient (účinná látka pazopanib), kromě inhibice VEGFR1, 2 a 3 inhibuje PDGFR (receptory růstového faktoru odvozeného z trombocytů) a c-kit (tyrosinkinázový receptor především kmenových pluripotentních krvetvorných buněk v kostní dřeni) (Sleijfer et al. 2009).

Tento lék se používá při léčbě rozvinutého nádoru vzniklého z renálních buněk a po chemoterapii při léčbě nádoru měkkých tkání. Poměrně zajímavým zjištěním je, že pokud se při radioterapii používá inhibice angiogeneze, zvyšuje to její účinnost a zabraňuje tvorbě nových cév jako obrannému mechanismu nádoru (Shawver et al. 2002).

Na pacientech s diferencovanou rakovinou štítné žlázy byl prováděn pokus, který měl ověřit účinky TKI sorafenibu a sunitinibu. Sorafenib byl podáván 13 pacientům, sunitinib byl podán jednomu pacientu hned a druhému potom, co nezabírala léčba sorafenibem.

Zlepšení nastalo ve 12 případech, ať již částečné zmírnění progresu ve třech případech nebo stabilizace choroby ve zbylých devíti. U tří pacientů zůstala progresivní forma rakoviny. Před léčbou pomocí TKI, při konvenční léčbě, byl medián doby bez progresu (PFS – progress-free survival) nádoru 14 měsíců, po léčbě sorafenibem a sunitinibem byl medián PFS 19 měsíců (Cabanillas et al. 2010).

Získávání a ukládání dat

Je vidět, že posttranslační modifikace jsou plnohodnotnou součástí proteinu s dopadem na jeho funkci. Ovšem problémem a zároveň výzvou je, že posttranslační modifikace nelze nijak zaznamenat do genetického kódu jako aminokyseliny proteinu. I když už dnes existují způsoby, jak odhadnout ze sekvence aminokyselin, nebo spíše ze specifity akceptorového místa, jestli a jaká posttranslační modifikace tam bude nejpravděpodobněji. Takovéto odhady stojí především na velkém množství experimentálních dat, uložených v databázích, jako je dbPTM, výpočetní síle moderních počítačů a různých matematických modelech.

Data experimentální

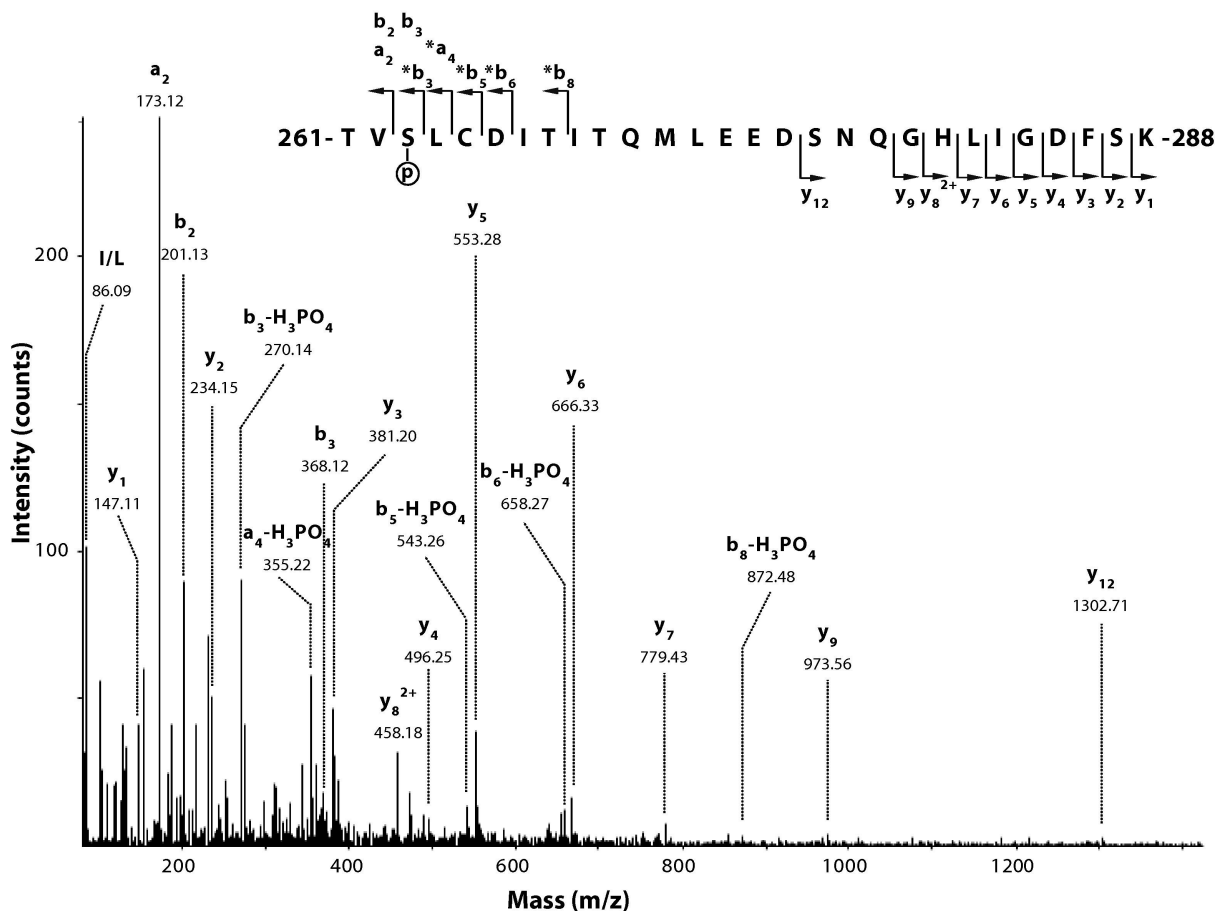
Hmotnostní spektrometrie

K získávání dat o fosfoproteinech pomocí hmotnostní spektrometrie dochází ve třech základních krocích. Během prvního kroku dochází k separaci fosforylovaných proteinů, druhým krokem je tandemová hmotnostní spektrometrie spojená s kapalinovou chromatografií (lc-ms/ms), ve třetí fázi dochází k identifikaci fosforylovaných aminokyselinových zbytků korelací sekvenčních dat a dat z hmotnostní spektrometrie. (Zhou et al. 2001)

Během hmotnostní spektrometrie dochází ke štěpení vzorků na ionty, jejich uspořádání dle hmotnosti a následné detekci.

Při tandemové hmotnostní spektrometrii se v prvním hmotnostním analyzátoru vyselektují požadované ionty, které jsou opět fragmentovány a prochází dalším hmotnostním analyzátozem. Při studiu fosforylací se používají hmotnostní spektrometrie druhého a třetího stupně. (Witze et al. 2007)

Výstupem z hmotnostní spektrometrie jsou spektra, jež ukazují relativní nebo absolutní zastoupení iontů s určitým poměrem hmotnosti a náboje (m/Q) (Obrázek 4).



Obrázek 4: Ukázka spektra z tandemové hmotnostní spektrometrie. Způsob identifikace fosfoserinu 263, pomocí řad iontů. (Esmenjaud-Mailhat et al. 2007)

2D elektroforéza

Pomocí dvoudimenzionální elektroforézy je možné získávat data dvěma způsoby. V prvním případě dochází ke značení proteinů pomocí radioaktivního fosforu ^{32}P nebo ^{33}P prováděné in vivo nebo in vitro, následuje separace pomocí 2D elektroforézy a vizualizace autoradiografií. Tato metoda je vysoce přesná, ale má několik omezení, jako například to,

že není schopna rozeznat, která aminokyselina je fosforylována. Radiace může stresovat buňku natolik, že může dojít k vytvoření artefaktů, také často dochází k zastavení buněčného cyklu.

Dalším způsobem je značení pomocí protilátek proti jednotlivým fosforylovaným aminokyselinám. Protilátky proti fosfotyrosinu jsou poměrně efektivní, ale protilátky proti fosfothreoninu a fosfoserinu nemají tak dobré výsledky; je to způsobeno tím, že jsou mnohem menší než fosfotyrosin a nedojde tak k jejich rozpoznání. Proto se pro jejich identifikaci používají protilátky proti motivu nebo většímu epitopu. Například rozpoznání fosfothreoninu s prolinem v pozici +1, typické pro MAP kinázy a cyklin dependentní kinázy (Kovarova et al. 2003).

Databáze

Experimentální zjišťování posttranslačních modifikací je časově i finančně náročné a proto je třeba věnovat velké úsilí i ukládání těchto dat. Databáze však dnes nejsou jen úložištěm, ale také místem, které umožňuje další zkoumání. Databáze mohou být zaměřeny přímo na posttranslační modifikace, například Phosida, nebo na proteiny jako takové (UniProt/Swiss-Prot, UniProt/TrEMBL), případně jako databáze zaměřená na choroby, například OMIM, kde jsou zaznamenány patologické posttranslační modifikace v souvislosti s jednotlivými chorobami. Další možností je databáze enzymů, například Kinbase, Kinweb nebo HomoKinase, se záznamy o kinázách.

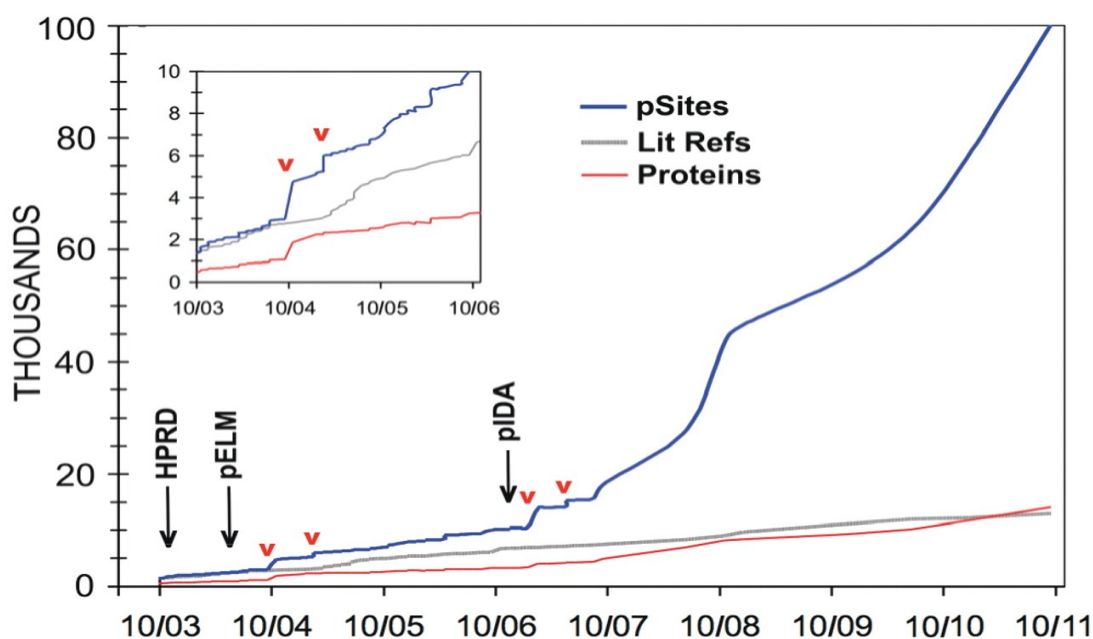
Databáze posttranslačních modifikací

Posttranslační modifikace jsou evidovány v několika databázích. Existují všeobecné databáze i databáze se zaměřením na určitý okruh organismů, například databáze virových posttranslačních modifikací. Databáze specializované na fosforylace jsou z hlediska fosforylace srovnány v tabulce níže (Tabulka 2).

Phospho.ELM (Dinkel et al. 2011) je databáze dat o fosforylaci in vivo i in vitro, jedná se o data experimentální i data z literatury. Obsahuje neredundantní záznamy o fosforylaci na serinech, threoninech a tyrosinech. V této databázi jsou především data z lidských proteinů, dále z modelových organismů (myš, octomilka, háďátko).

PhosphoSitePlus (Hornbeck et al. 2012) je databáze zaměřená na posttranslační modifikace převážně lidských a myších proteinů, nicméně obsahuje i data týkající se jiných organismů. Obsahuje data v první řadě o fosforylaci, ubiquitinaci a acetylaci. Je

ručně sestavovaná a její předností je rozřídění dat podle několika hledisek a tedy snadnější přístup k požadovanému záznamu. Nabízí možnost procházení databáze nejen podle místa fosforylace, ale i podle buněčného typu a podle choroby. Databáze PhosphoSitePlus obsahuje více než dvě stě tisíc fosforylačních míst na více než devatenácti tisících proteinů. Mimo dat získaných pomocí úzce zaměřených studií, která jsou přesná, obsahuje databáze také data získaná pomocí tandemové hmotnostní spektrometrie. PhosphoSitePlus vychází z databáze Phosphosite, která obsahovala data týkající se fosforylací v myších a lidských proteinech, včetně informací o souvisejících kinázách a fosfatázách, její růst můžete vidět níže (Obrázek 7) (Hornbeck et al. 2004).



Obrázek 5: Obsah PhosphoSite sestavený z literatury od spuštění databáze v roce 2003 do roku 2011. Modrá křivka ukazuje počet záznamů o fosforylačních místech, červená počet proteinů a šedá reference. Černé šipky naznačují spuštění dalších databází (Human Protein Reference Database, Phospho.ELM, PHOSIDA). Červenými zobáčky je naznačena publikace některých souborů dat z hmotnostní spektrometrie. (Hornbeck et al. 2012)

Další z databází fosforylací je Phosida (Gnad et al. 2007, 2011), která kromě dat o fosforylací shromažďuje data o N – glykosylaci a acetylací. Zabývá se modelovými organismy biologie a tak kromě myši a člověka obsahuje záznamy o fosforylací i u jiných obratlovců, bezobratlých, jednobuněčných eukaryot i prokaryot. Záznamy z této databáze nejsou zahrnuty v dbPTM databázi.

Phospho3D (Zanzoni et al. 2011) je databáze, která shromažďuje data o 3D struktuře fosforylačních míst, respektive predikce založené na databázi Phospho.ELM, která pro fosforylační místa zaznamenává i zbytky v prostorovém okolí. Phospho3D nabízí také nástroj P3DSCAN, umožňující porovnání lokální struktury proteinu vůči 5387 zmapovaných případů v Phospho.ELM..

MtcPTM (Jiménez et al. b.r.) je databáze myších a lidských fosforylačních míst, s možností porovnání motivů. V jiných databázích jsou fosforylace pozičně vztaženy k proteinové sekvenci, v mtcTPM jsou fosforylace vztaženy k experimentálně určené peptidové sekvenci, což umožňuje snadnější manipulaci s daty, jejich aktualizaci. Kromě toho jsou v databázi také uloženy atomové modely proteinů a je možno je zobrazit.

HPRD (Human Protein Reference Database) (Keshava Prasad et al. 2009), je databáze pouze lidských proteinů, nabízí procházení obsahu podle posttranslačních modifikací, domény, motivu, lokalizace a funkce. Je možno využít nástroj BLAST. Dále nabízí nástroj PhosphoMotif Finder, který vyhledává v uživatelem zadané sekvenci motiv fosforylačního místa.

PTMScout je databáze, která obsahuje data z hmotnostní spektrometrie fosfoproteinů, dále obsahuje data z databází PhosphoSite, Phospho.ELM a Uniprot. Je založena na MySQL databázi, nabízí nástroj pro generování biologických hypotéz, který je ovšem omezen anotacemi metadat (Naegle et al. 2010).

dbPTM (Lee et al. 2006) je souhrnná databáze se záznamy z několika dalších databází, nejvíce záznamů je ze Swiss-Prot, mimo jiné proto, že jsou do databáze zahrnuty i předpokládané posttranslační modifikace. Další velké databáze zahrnuté v dbPTM jsou PhosphositePlus a Phospho.ELM. Dále dbPTM obsahuje záznamy z databází dalších posttranslačních modifikací, například methylací, O-glykosylací nebo ubiquitinací. Obsahuje i databázi kináz z Kinbase. Kromě modifikací předpokládaných v databázi Swiss-Prot obsahuje modifikace predikované pomocí skrytých Markovových modelů.

Swiss-Prot je částí UniProtKB (The UniProt Consortium, 2012) s daty anotovanými a přezkoumána lidmi. Druhá část UniProtKB je TrEMBL, kde jsou data anotována strojově a nejsou přezkoumávána. Swiss-Prot obsahuje data experimentální, ale i predikční. UniProtKB je databáze neredundantní. Pro jeden gen existuje pouze jeden záznam obsahující všechny další informace, jako sestřihové varianty a mutace. Autoři mohou také zahrnout informace o fosforylaci, pokud je známa.

databáze	počet záznamů	organismy	www
Phospho.ELM	8718 sekvencí proteinů 299 kináz	<i>H. sapiens</i> , <i>M. musculus</i> , <i>D. melangoster</i>	http://phospho.elm.eu.org
PHOSIDA	15368 proteinů 45833 fosforylačních míst	<i>H. sapiens</i> , <i>M. musculus</i> , <i>D. melangoster</i>	http://www.phosida.com/
PhosphoSitePlus	19844 proteinů 208901 fosforylačních míst	<i>H. sapiens</i> , <i>M. musculus</i> , <i>D. melangoster</i>	http://www.phosphosite.org
Phospho3D	5387 zmapovaných případů	Viz Phospho.ELM	http://www.phospho3d.org
mtcPTM	22005 fosforylačních míst 11619 proteinů	<i>H. sapiens</i> <i>M. musculus</i>	http://www.mitocheck.org/cgi-bin/mtc
HPDR	30047 proteinů 8160 záznamů o fosforylaci	<i>H. sapiens</i>	http://www.hprd.org
dbPTM	153113 experimentálně zjištěných fosforylačních míst 74365 předpokládaných fosforylačních míst z UniProt/SwissProt 1763242 fosforylačních míst předpokládaných pomocí HMM	<i>H. sapiens</i>	http://dbptm.mbc.nctu.edu.tw
PTMscout	206495 fosforylací	72 různých druhů	http://ptmscout.mit.edu

Tabulka 2: Databáze týkající se fosforylace.

Databáze enzymů

Kinbase je databází kináz různých organismů od člověka, myš, přes kvasinky až po rostlinu. Nachází se na webu www.kinase.com. V této databázi se dá vyhledávat podle kritérií (rodina kináz, organismu, ze kterého kináza pochází, doména přítomná v kináze) a také pomocí nástroje BLAST. Další nástroje, nikoliv vyhledávací, jsou HyperTree, prohlížeč fylogenetických stromů, a u jednotlivých tříd nabízí grafické srovnání zastoupení aminokyselin na jednotlivých pozicích.

PhosphaBase (Wolstencroft et al. 2005) je databáze fosfatáz, nabízí vyhledávání podle kritérií (rodina, název, tkáň, organismus, ze kterého pochází, funkce), vyhledání podle motivu a podle domény. Kromě vyhledávání nabízí také 3D prohlížeč proteinových struktur, nástroj pro práci s alignmentem a alignment fosfotyrosin fosfatáz. Nachází se na webu www.bioinf.manchester.ac.uk/phosphabase/.

Databáze týkající se nádorů

OMIM® (Online Mendelian Inheritance in Man) je databáze lidských genů a genetických poruch. Nabízí vyhledávání záznamů a genovou mapu. K jednotlivým genům nabízí komplexní text zabývající se různými pohledy, od biochemie, přes mapování, po roli v patologiích. V tomto textu jsou citovány články, které se daným genem zabývají.

Další z genomových databází, která se zaměřuje i na posttraslační modifikace, je cBioPortal for Cancer Genomics, která shromažďuje genomová data ze studií o rakovině; v březnu 2012 měla 5000 vzorků nádorů z 20 studií (Cerami et al. 2012), nyní je to přes 11000 vzorků získaných z 33 publikací. Databáze pomocí funkce OnkoPrint nabízí jednoduchý způsob zobrazování změn na úrovni DNA, jako jsou delece, duplikace, SNP, pro jednotlivé geny i pro jednotlivé pacienty. Kromě databáze změn na úrovni DNA jsou v databázi RNA záznamy (miRNA a mRNA) zejména o expresi, proteinová data i se záznamy o fosforylaci, záznamy o metylaci DNA a data týkající se rakoviny (zejména přežití). Pro MATLAB a R má databáze vytvořené vlastní knihovny, které umožňují přímý přístup k datům serveru, kromě těchto programů lze s daty pracovat i jiných programovacích jazycích (Python, Perl...) Celá databáze je Open-Source s licencí GNU Lesser General Public License (<http://www.gnu.org/licenses/lgpl.html>).

ROCK je databáze zaměřená na funkční genomiku rakoviny prsu. Obsahuje data o genech z různých databází (např. OMIM, Uniprot, Ensembl, UniGene, PDB, RefSeq), data o interakcích a data o metabolických drahách, kterých se proteiny v databázi obsažené účastní.

Nachází se zde také záznamy o kinázách a fosfatázách. Poměrně intuitivní rozhraní nabízí přístup k datům pomocí hledání nebo procházení dat podle různých kritérií, kromě dat nabízí i analýzu dat podle zvolených kritérií. Tato databáze byla myšlena jako protiváha k databázím, které se zabývají především sběrem genomických dat z literatury (Breast Cancer Information Core, Breast Cancer Database) (Sims et al. 2010).

Data predikční

Z výše uvedeného je vidět, že experimentální metody jsou náročné na vybavení a přes to neumí odhalit všechna fosforylační místa, proto se velmi brzo začaly používat prediktivní metody, založené na analýze již známých dat. Díky vývoji výpočetní techniky a většímu množství experimentálních dat dochází ke stále přesnějším predikcím. Bohužel je toto téma tak obsáhlé, že by vydalo na samostatnou práci, nicméně je natolik zásadní, že ho musím alespoň částečně zmínit ve své práci.

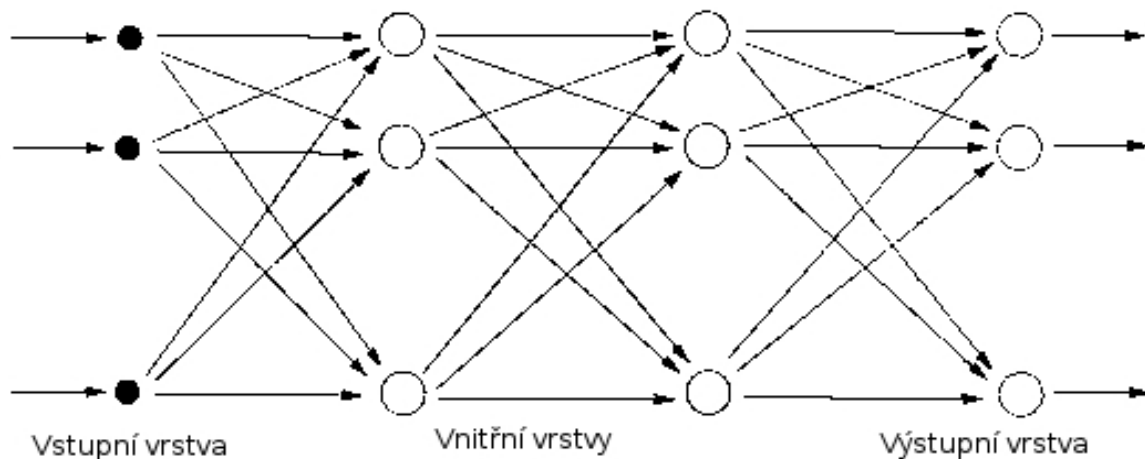
Dříve se k predikci posttranslačních modifikací používaly váhové matice, dnes se přechází postupně od skrytých Markovových modelů k neuronovým sítím. (Blom et al. 2004)

Markovovy modely jsou používány v bioinformatice nejen pro určování posttranslačních modifikací. Tyto výpočetní metody využívají Bayesovy věty a na ní založené statistiky. Tyto početní metody dovolují na základě nových skutečností upravovat, zpřesňovat, podmíněnou pravděpodobnost jevů.

Ve skrytých Markovových modelech je skrytý Markovovský řetězec, což znamená, že nejsou známy ani jednotlivé stavy, ani přechody mezi nimi, známy jsou pouze výstupy, které vznikají při přechodu mezi stavy. (Eddy 2004)

Neuronové sítě (ANN – Artificial neural networks) nejsou na poli výpočetní techniky novinkou, byly objeveny již ve 40. letech 20. století, ovšem jejich běžné využívání roste s výpočetní kapacitou počítačů. Existují dva typy neuronových sítí, dopředné typy bez smyček a typy se smyčkami, které umožňují zpětnou vazbu (Jain, Mohiuddin 1996).

V dopředných neuronových sítích jsou jednotlivé výpočetní body uspořádány ve vrstvách, body jsou propojeny mezi vrstvami, nikoliv uvnitř vrstev, a hrany (spoje mezi výpočetními body) mají váhu. Často bývá bod propojen se všemi body předchozí vrstvy vstupy a pomocí výstupů bývá propojen se všemi body vrstvy následující (Obrázek 6). Vstupní signály bývají různé, dochází k jejich zpracování a na jejich základě vytvoření jednotného výstupního signálu (Blom et al. 2004).



Obrázek 6: Ukázka stavby trojvrstvé neuronové sítě. (Jain, Mohiuddin 1996)

Podobně jako skryté Markovovy modely jsou neuronové sítě trénovatelné. Učení probíhá pomocí iterativního upravování existence a vah hran mezi výpočetními body (Jain, Mohiuddin 1996).

Neuronové sítě se pro predikci posttranslačních modifikací používají proto, že umí analyzovat nelineární strukturu okolí akceptorového místa posttranslační modifikace, tedy berou v potaz okolí prostorové, nikoli pouze okolí sekvenční (Blom et al. 2004).

Často používané nástroje vycházející z popsaných výpočetní metod najdete v tabulce níže

Metoda	Nástroj	Odkaz
ANN	NetPhos	http://www.cbs.dtu.dk/services/NetPhos/
	NetPhosK	http://www.cbs.dtu.dk/services/NetPhosK/
HMM	KinasePhos	http://kinasephos.mbc.nctu.edu.tw/

Tabulka 3: Přehled predikčních nástrojů.

(Tabulka 3).

Analýza fosforylací v nádorových onemocněních

Mutace v DNA jsou podle své závažnosti děleny na řídicí (drivers) a vezoucí se (passenger). Řídicí mutace jsou charakterizovány tím, že se vyskytují v regiorech s vysokým poměrem nesynonymních mutací vůči synonymním a očekává se, že vedou ke změnám vedoucím k rakovinnému zvrhnutí buňky (Torkamani et al. 2009). Řídicí mutace mají tendenci se vyskytovat spíše na povrchu, v blízkosti hydrofilních aminokyselin (Talavera et al. 2010).

Řídicí mutace často souvisí s fosforylací, ať už se vyskytuje v kinázových doménách nebo v sekvenci substrátových proteinů. V substrátových proteinech může dojít k mutaci fosforylovatelných aminokyselin nebo aminokyselin v okolí serinu, threoninu nebo tyrosinu, což může zabránit rozpoznání fosforylačního místa.

Ze 793 vzorků z 8 typů nádorových onemocnění byly shromážděny geny, jejichž jednonukleotidová mutace souvisí s fosforylací, ať už tak, že změní kinázovou doménu nebo fosforylační místo (7 kyselin před a za fosforylovanou aminokyselinou). Tato analýza vedla k označení 775 různých genů, jež obsahují 949 fosforylačně specifických SNV v alespoň jednom typu rakoviny.

Přibližně jedna třetina označených genů s SNV ve fosforylačním místě se vyskytovala ve více typech rakoviny, což může pomoci zviditelnit mechanismus řídicích mutací v rakovině. (Reimand, Bader 2013)

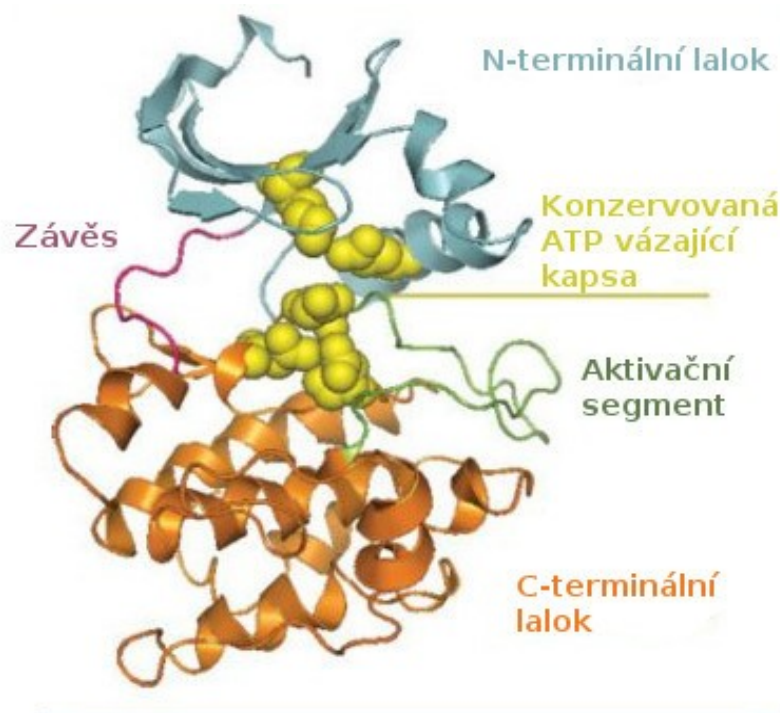
Řídicí mutace objevující se v oblastech kinázové domény se vyskytují statisticky prokazatelně blíže k ATP-vázající oblasti než vezoucí se mutace (Obrázek 7) (Izarzugaza et al. 2009).

Řídicí mutace mohou být i prospěšné. V případě tumorsupresorového proteinu TP53 (v případech rakoviny vaječnicků a glioblastomů), může dojít ke vzniku fosforylačně specifických SNV tak, že pacienti mají delší dobu přežití, v případě glioblastomu, nebo jich přežívá větší počet, v případě rakoviny vaječnicků. (Reimand, Bader 2013)

V rakovině plic dochází k hyperfosforylaci, například na receptoru pro epidermální růstový faktor (Wu et al. 2009), což vede k nadměrné tyrosinové fosforylaci proteinů (Guha et al. 2008). Nadměrná exprese fosfotyrosinových proteinů souvisí s proliferací nádorových buněk a změnami tvaru buňky a tkáně. Čím vyšší exprese fosfotyrosinových

proteinů, tím horší prognóza přežití, i bez dalších patologických faktorů (Gong et al. 2002). Nadměrná exprese nebo zvýšená aktivity src-like kináz a aktivace EGFR jsou nezbytné pro proliferaci a invazi rakovinných buněk (Wu et al. 2010).

Na základě rozdílné fosforylace tyrosinových zbytků lze jasně rozlišit zdravé a nádorové buňky, a to bez ohledu na to, zda byla analýza provedena na velkém nebo malém počtu fosforylačních míst. Toto rozdělení je v podstatě rovnocenné microarray analýze (Wu et al. 2009).



Obrázek 7: Model kinázy, založený na MAP3K1, ukazuje základní uspořádání s dvěma laloky spojenými pomocí závěsu. ATP se váže mezi laloky na 5 vysoce konzervativních zbytků (K74, E96, D171, N176, D190). (Izarzugaza et al. 2009)

Úbytek fosforylačních míst a jejich vznik de novo

Nádorové buňky obsahují řadu mutací. Exprese poškozených genů vede ke vzniku proteinů, které mohou mít nová místa vyhledávaná kinázami, případně může díky SNP (jednonukleotidovému polymorfismu) dojít k záměně aminokyseliny fosforylovatelné za jinou. Stejně tak to může vést k expresi fosfatáz a kináz vyhledávající jiný substrát nebo naopak nefunkčních.

V rakovinných buňkách dochází spíše k mírnému přibývání fosforylačních míst (ztráta fosforylačních míst kolem 12%, zatím co vznik nových míst cca 14%), zatímco v kontrolním datasetu, vzniklém mutací sekvencí *in silico*, docházelo častěji ke ztrátám (ztráta fosforylačních míst 11%, vznik nových míst cca 9%). Tento rozdíl je dán tím, že algoritmus provádí evolučně preferované modifikace, tedy spíše neutrální. Příbytek fosforylací v nádorových onemocněních může mít za důsledek destabilizaci proteinu, přerušení jeho interakcí, narušení jeho enzymatické funkce nebo dokonce ovlivnění aktivity jiných enzymů úplně jiných metabolických drah, což má za následek poruchy v celé buňce (Radivojac et al. 2008).

Závěr

Ve své práci jsem se zabývala fosforylacemi, způsoby získávání dat o fosforylaci a jejich uskladnění s důrazem na vliv fosforylací na nádorová onemocnění.

Fosfotyrosin je používán jako regulační prvek aktivity enzymů a jeho existence je silně regulována, nádorové buňky se vymknou této regulaci a dochází k tvorbě nádoru pomocí nekontrolovatelného dělení buněk, dále k vytváření krevního řečiště kolem nádoru a posléze migraci nádorových buněk a vzniku metastáz. Změnu na proteinové úrovni způsobuje mutace v DNA, za rakovinné zvrhnutí buňky mohou takzvané řídicí mutace.

Zásadním přínosem pro proteomiku bylo použití hmotnostní spektrometrie, která v kombinaci s dalšími metodami umožňuje objevit mnohem větší množství fosforylačních míst. Další možností je predikce fosforylačních míst na základě získaných dat. Získaná data jsou ukládána do databází, aby byla snadno dosažitelná. Databáze se liší svou specializací, redundancí a anotací dat. Kromě ukládání dat nabízí databáze i nástroje pro jejich vyhledávání dle různých kritérií a nástroje pro analýzu, jako je alignment, které mohou sloužit k zjišťování konzervativních motivů napříč rodinami proteinů.

V posledních letech se začala používat biologická léčba nádorových onemocnění, používají se k ní tyrosinkinázové inhibitory. Některé tyrosinkinázové inhibitory interagují s kinázami buněk nádoru a inhibují dělení buněk. Další zabraňuje angiogenezi, pomocí inhibice receptorových tyrosinkináz endotelu a dalších tyrosinkináz krevních buněk. Existují i kinázové inhibitory, které mohou inhibovat více proteinkináz.

Jako jakákoliv jiná léčba mají i tyrosinkinázové inhibitory vedlejší účinky, související s mechanismem jejich působení. Poškozují zdravé i nádorové buňky, ale kvůli rychlejšímu metabolismu a nízké reparaci poškozují nádorové buňky mnohem více. Kombinují se s dalšími léčebnými postupy, ale jejich používání zvyšuje dobu dožití pacientů.

Pomocí výpočetní analýzy dat týkajících se fosforylace lze mnohem rychleji najít látky, které by mohly být biologicky aktivní a reagovat s proteiny rakovinné buňky, zejména s kinázami a fosfatázami. Také je možno mnohem lépe odhalit souvislosti v dynamice fosforylací, což pomáhá identifikovat možné cíle biologické léčby.

Zdroje

- ARNESEN, Thomas, 2011. Towards a functional understanding of protein N-terminal acetylation. *PLoS biology* [online]. vol. 9, no. 5, s. e1001074, [vid. 17. August 2013]. ISSN 1545-7885. Dostupné z: doi:10.1371/journal.pbio.1001074
- BLOM, Nikolaj, Thomas SICHERITZ-PONTÉN, Ramneek GUPTA, Steen GAMMELTOFT and Søren BRUNAK, 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* [online]. vol. 4, no. 6, s. 1633–49, [vid. 4. March 2013]. ISSN 1615-9853. Dostupné z: doi:10.1002/pmic.200300771
- CABANILLAS, Maria E, Steven G WAGUESPACK, Yulia BRONSTEIN, Michelle D WILLIAMS, Lei FENG, Mike HERNANDEZ, Adriana LOPEZ, Steven I SHERMAN and Naifa L BUSAIDY, 2010. Treatment with tyrosine kinase inhibitors for patients with differentiated thyroid cancer: the M. D. Anderson experience. *The Journal of clinical endocrinology and metabolism* [online]. vol. 95, no. 6, s. 2588–95, [vid. 6. May 2013]. ISSN 1945-7197. Dostupné z: doi:10.1210/jc.2009-1923
- CERAMI, Ethan, Jianjiong GAO, Ugur DOGRUSOZ, Benjamin E GROSS, Selcuk Onur SUMER, Bülent Arman AKSOY, Anders JACOBSEN, Caitlin J BYRNE, Michael L HEUER, Erik LARSSON, Yevgeniy ANTIPIN, Boris REVA, Arthur P GOLDBERG, Chris SANDER and Nikolaus SCHULTZ, 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery* [online]. . S.l.: American Association for Cancer Research, vol. 2, no. 5, s. 401–4, [vid. 2. April 2013]. ISSN 2159-8290. Dostupné z: doi:10.1158/2159-8290.CD-12-0095
- DARNELL, James E., 1997. Phosphotyrosine signaling and the single cell: metazoan boundary. *Proceedings of the National Academy of ...* [online]. vol. 94, no. October, s. 11767–11769, [vid. 18. August 2013]. Dostupné z: <http://www.pnas.org/content/94/22/11767.short>
- DINKEL, Holger, Claudia CHICA, Allegra VIA, Cathryn M GOULD, Lars J JENSEN, Toby J GIBSON and Francesca DIELLA, 2011. Phospho.ELM: a database of phosphorylation sites--update 2011. *Nucleic acids research* [online]. vol. 39, no. Database issue, s. D261–7, [vid. 27. February 2013]. ISSN 1362-4962. Dostupné z: doi:10.1093/nar/gkq1104
- EDDY, Sean R, 2004. PRIMER What is a hidden Markov model? vol. 22, no. 10, s. 1315–1317. ,
- ESMENJAUD-MAILHAT, Charlotte, Valérie LOBJOIS, Carine FROMENT, Roy M GOLSTEYN, Bernard MONSARRAT and Bernard DUCOMMUN, 2007. Phosphorylation of CDC25C at S263 controls its intracellular localisation. *FEBS letters* [online]. vol. 581, no. 21, s. 3979–85, [vid. 12. August 2013]. ISSN 0014-5793. Dostupné z: doi:10.1016/j.febslet.2007.07.024
- GNAD, Florian, Jeremy GUNAWARDENA and Matthias MANN, 2011. PHOSIDA 2011: the posttranslational modification database. *Nucleic acids research* [online]. vol. 39, no. Database issue, s. D253–60, [vid. 27. February 2013]. ISSN 1362-4962. Dostupné z: doi:10.1093/nar/gkq1159
- GNAD, Florian, Shubin REN, Juergen COX, Jesper V OLSEN, Boris MACEK, Mario OROSHI and Matthias MANN, 2007. PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome biology* [online]. vol. 8, no. 11, s. R250, [vid. 3. March 2013]. ISSN 1465-6914. Dostupné z: doi:10.1186/gb-2007-8-11-r250
- GONG, Y, T HIRANO, Y KATO, K YOSHIDA, Y SHOU, T OHIRA, N IKEDA, Y EBIHARA and H KATO, 2002. Phosphorylated tyrosine-containing proteins in primary lung cancer correlates with proliferation and prognosis. *British journal of cancer* [online]. vol. 86, no. 12, s. 1893–8, [vid. 19. August 2013]. ISSN 0007-0920. Dostupné z: doi:10.1038/sj.bjc.6600327

GUHA, Udayan, Raghothama CHAERKADY, Arivusudar MARIMUTHU, a Scott PATTERSON, Manoj K KASHYAP, H C HARSHA, Mitsuo SATO, Joel S BADER, Alex E LASH, John D MINNA, Akhilesh PANDEY and Harold E VARMUS, 2008. Comparisons of tyrosine phosphorylated proteins in cells expressing lung cancer-specific alleles of EGFR and KRAS. *Proceedings of the National Academy of Sciences of the United States of America* [online]. vol. 105, no. 37, s. 14112–7., ISSN 1091-6490. Dostupné z: doi:10.1073/pnas.0806158105

HANKS, SK and T HUNTER, 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *The FASEB Journal* [online]. [vid. 5. May 2013]. Dostupné z: <http://www.fasebj.org/content/9/8/576.short>

HORNBECK, Peter V, Indy CHABRA, Jon M KORNHAUSER, Elzbieta SKRZYPEK and Bin ZHANG, 2004. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics* [online]. vol. 4, no. 6, s. 1551–61, [vid. 16. November 2012]. ISSN 1615-9853. Dostupné z: doi:10.1002/pmic.200300772

HORNBECK, Peter V, Jon M KORNHAUSER, Sasha TKACHEV, Bin ZHANG, Elzbieta SKRZYPEK, Beth MURRAY, Vaughan LATHAM and Michael SULLIVAN, 2012. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic acids research* [online]. vol. 40, no. Database issue, s. D261–70, [vid. 3. March 2013]. ISSN 1362-4962. Dostupné z: doi:10.1093/nar/gkr1122

HUNTER, T and G D PLOWMAN, 1997. The protein kinases of budding yeast: six score and more. *Trends in biochemical sciences* [online]. vol. 22, no. 1, s. 18–22, [vid. 6. May 2013]. ISSN 0968-0004. Dostupné z: <http://www.ncbi.nlm.nih.gov/pubmed/9020587>

HUNTER, Tony, 2009. Tyrosine phosphorylation: thirty years and counting. *Current opinion in cell biology* [online]. vol. 21, no. 2, s. 140–6, [vid. 14. November 2012]. ISSN 1879-0410. Dostupné z: doi:10.1016/j.ceb.2009.01.028

HUNTER, Tony, 2012. Why nature chose phosphate to modify proteins. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* [online]. vol. 367, no. 1602, s. 2513–6, [vid. 7. March 2013]. ISSN 1471-2970. Dostupné z: doi:10.1098/rstb.2012.0013

IZARZUGAZA, Jose M G, Oliver C REDFERN, Christine a ORENGO and Alfonso VALENCIA, 2009. Cancer-associated mutations are preferentially distributed in protein kinase functional sites. *Proteins* [online]. vol. 77, no. 4, s. 892–903, [vid. 5. November 2012]. ISSN 1097-0134. Dostupné z: doi:10.1002/prot.22512

JAIN, A.K. and K.M. MOHIUDDIN, 1996. Artificial neural networks: a tutorial. *Computer* [online]. . S.l.: IEEE, vol. 29, no. 3, s. 31–44, [vid. 5. April 2013]. ISSN 00189162. Dostupné z: doi:10.1109/2.485891

JEMAL, Ahmedin, Freddie BRAY, Jacques FERLAY, Elizabeth WARD and David FORMAN, 2011. Global Cancer Statistics. [online]. vol. 61, no. 2, s. 69–90., Dostupné z: doi:10.3322/caac.20107.Available

JIMÉNEZ, JL, B HEGEMANN and JRA HUTCHINS, b.r. mtcPTM: a database for comparison of phosphorylation patterns from different experiments and structural analysis of phosphosites. *193.62.203.114* [online]. [vid. 22. August 2013]. Dostupné z: <ftp://193.62.203.114/pub3/mitocheck/mtcPTM-MS/text/revised-002-ref-GenomeBiology.rtf>

KESHAVA PRASAD, T S, Renu GOEL, Kumaran KANDASAMY, Shivakumar KEERTHIKUMAR, Sameer KUMAR, Suresh MATHIVANAN, Deepthi TELIKICHERLA, Rajesh RAJU, Beema SHAFREEN, Abhilash VENUGOPAL, Lavanya BALAKRISHNAN, Arivusudar MARIMUTHU, Sutopa BANERJEE, Devi S SOMANATHAN, Aimy SEBASTIAN, Sandhya RANI, Somak RAY, C J HARRYS KISHORE, Sashi KANTH, Mukhtar AHMED, Manoj K KASHYAP, Riaz MOHMOOD, Y L RAMACHANDRA, V KRISHNA, B Abdul RAHIMAN, Sujatha MOHAN, Prathibha RANGANATHAN, Subhashri RAMABADRAN, Raghothama CHAERKADY and Akhilesh PANDEY, 2009. Human Protein Reference Database--2009 update. *Nucleic acids research* [online]. vol. 37, no. Database issue, s. D767–72, [vid. 7. August 2013]. ISSN 1362-4962. Dostupné z: doi:10.1093/nar/gkn892

KLUMPP, Susanne and Josef KRIEGLSTEIN, 2002. Phosphorylation and dephosphorylation of histidine residues in proteins. *European journal of biochemistry / FEBS* [online]. vol. 269, no. 4, s. 1067–71., ISSN 0014-2956. Dostupné z: <http://www.ncbi.nlm.nih.gov/pubmed/11856347>

KOVAROVA, H, M HAJDUCH, M LIVINGSTONE, P DZUBAK and I LEFKOVITS, 2003. Analysis of state-specific phosphorylation of proteins by two-dimensional gel electrophoresis approach. *Journal of Chromatography B* [online]. vol. 787, no. 1, s. 53–61, [vid. 19. August 2013]. ISSN 15700232. Dostupné z: doi:10.1016/S1570-0232(02)00729-8

LEE, Tzong-Yi, Hsien-Da HUANG, Jui-Hung HUNG, Hsi-Yuan HUANG, Yuh-Shyong YANG and Tzu-Hao WANG, 2006. dbPTM: an information repository of protein post-translational modification. *Nucleic acids research* [online]. vol. 34, no. Database issue, s. D622–7, [vid. 1. May 2013]. ISSN 1362-4962. Dostupné z: doi:10.1093/nar/gkj083

LOWE, SW and AW LIN, 2000. Apoptosis in cancer. *Carcinogenesis* [online]. vol. 21, no. 3, s. 485–495, [vid. 19. August 2013]. ISSN 14602180. Dostupné z: doi:10.1093/carcin/21.3.485

MANNING, G, D B WHYTE, R MARTINEZ, T HUNTER and S SUDARSANAM, 2002. The protein kinase complement of the human genome. *Science (New York, N.Y.)* [online]. vol. 298, no. 5600, s. 1912–34, [vid. 6. August 2013]. ISSN 1095-9203. Dostupné z: doi:10.1126/science.1075762

MUSTELIN, Tomas, 2007. A brief introduction to the protein phosphatase families. *Methods in molecular biology (Clifton, N.J.)* [online]. vol. 365, s. 9–22., ISSN 1064-3745. Dostupné z: doi:10.1385/1-59745-267-X:9

NAEGLE, Kristen M, Melissa GYMREK, Brian A JOUGHIN, Joel P WAGNER, Roy E WELSCH, Michael B YAFFE, Douglas A LAUFFENBURGER and Forest M WHITE, 2010. PTMScout, a Web resource for analysis of high throughput post-translational proteomics studies. *Molecular & cellular proteomics : MCP* [online]. vol. 9, no. 11, s. 2558–70, [vid. 25. March 2013]. ISSN 1535-9484. Dostupné z: doi:10.1074/mcp.M110.001206

PINNA, L a and M RUZZENE, 1996. How do protein kinases recognize their substrates? *Biochimica et biophysica acta* [online]. vol. 1314, no. 3, s. 191–225., ISSN 0006-3002. Dostupné z: <http://www.ncbi.nlm.nih.gov/pubmed/8982275>

RADIOVJAC, Predrag, Peter H BAENZIGER, Maricel G KANN, Matthew E MORT, Matthew W HAHN and Sean D MOONEY, 2008. Gain and loss of phosphorylation sites in human cancer. *Bioinformatics (Oxford, England)* [online]. vol. 24, no. 16, s. i241–7, [vid. 2. November 2012]. ISSN 1367-4811. Dostupné z: doi:10.1093/bioinformatics/btn267

REIMAND, Jüri and Gary D BADER, 2013. Systematic analysis of somatic mutations in phosphorylation signaling predicts novel cancer drivers. *Molecular systems biology* [online]. vol. 9, no. 637, s. 637, [vid. 4. August 2013]. ISSN 1744-4292. Dostupné z: doi:10.1038/msb.2012.68

SACCO, Francesca, Livia PERFETTO, Luisa CASTAGNOLI and Gianni CESARENI, 2012. The human phosphatase interactome: An intricate family portrait. *FEBS letters* [online]. . S.l.: Federation of European Biochemical Societies, vol. 586, no. 17, s. 2732–9, [vid. 10. August 2013]. ISSN 1873-3468. Dostupné z: doi:10.1016/j.febslet.2012.05.008

SHAWVER, Laura K, Dennis SLAMON and Axel ULLRICH, 2002. Smart drugs: tyrosine kinase inhibitors in cancer therapy. *Cancer cell* [online]. vol. 1, no. 2, s. 117–23., ISSN 1535-6108. Dostupné z: <http://www.ncbi.nlm.nih.gov/pubmed/12086869>

SIMS, David, Borisa BURSTEINAS, Qiong GAO, Ekta JAIN, Alan MACKAY, Costas MITSOPOULOS and Marketa ZVELEBIL, 2010. ROCK: a breast cancer functional genomics resource. *Breast cancer research and treatment* [online]. vol. 124, no. 2, s. 567–72, [vid. 13. May 2013]. ISSN 1573-7217. Dostupné z: doi:10.1007/s10549-010-0945-5

SLEIJFER, Stefan, Isabelle RAY-COQUARD, Zsuzsa PAPAI, Axel LE CESNE, Michelle SCURR, Patrick SCHÖFFSKI, Françoise COLLIN, Lini PANDITE, Sandrine MARREAU, Annick DE BRAUWER, Martine VAN GLABBEKE, Jaap VERWEIJ and Jean-Yves BLAY, 2009. Pazopanib, a multikinase angiogenesis inhibitor, in patients with relapsed or refractory advanced soft tissue sarcoma: a phase II study from the European organisation for research

and treatment of cancer-soft tissue and bone sarcoma group (EORTC study 620. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* [online]. vol. 27, no. 19, s. 3126–32, [vid. 12. May 2013]. ISSN 1527-7755. Dostupné z: doi:10.1200/JCO.2008.21.3223

TALAVERA, David, Martin S TAYLOR and Janet M THORNTON, 2010. The (non)malignancy of cancerous amino acidic substitutions. *Proteins* [online]. vol. 78, no. 3, s. 518–29, [vid. 5. November 2012]. ISSN 1097-0134. Dostupné z: doi:10.1002/prot.22574

TAYLOR, Susan S, Malik M KESHWANI, Jon M STEICHEN and Alexandr P KORNEV, 2012. Evolution of the eukaryotic protein kinases as dynamic molecular switches. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* [online]. vol. 367, no. 1602, s. 2517–28, [vid. 5. January 2013]. ISSN 1471-2970. Dostupné z: doi:10.1098/rstb.2012.0054

THE UNIPROT CONSORTIUM, 2012. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic acids research* [online]. vol. 40, no. Database issue, s. D71–5, [vid. 28. February 2013]. ISSN 1362-4962. Dostupné z: doi:10.1093/nar/gkr981

TORKAMANI, Ali, Gennady VERKHIVKER and Nicholas J SCHORK, 2009. Cancer driver mutations in protein kinase genes. *Cancer letters* [online]. . S.l.: Elsevier Ireland Ltd, vol. 281, no. 2, s. 117–27, [vid. 20. August 2013]. ISSN 1872-7980. Dostupné z: doi:10.1016/j.canlet.2008.11.008

TYANOVA, Stefka, Jürgen COX, Jesper OLSEN, Matthias MANN and Dmitrij FRISHMAN, 2013. Phosphorylation variation during the cell cycle scales with structural propensities of proteins. *PLoS computational biology* [online]. vol. 9, no. 1, s. e1002842, [vid. 4. August 2013]. ISSN 1553-7358. Dostupné z: doi:10.1371/journal.pcbi.1002842

WISHART, David S, Craig KNOX, An Chi GUO, Dean CHENG, Savita SHRIVASTAVA, Dan TZUR, Bijaya GAUTAM and Murtaza HASSANALI, 2008. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic acids research* [online]. vol. 36, no. Database issue, s. D901–6, [vid. 27. February 2013]. ISSN 1362-4962. Dostupné z: doi:10.1093/nar/gkm958

WITZE, Eric S, William M OLD, Katheryn A RESING and Natalie G AHN, 2007. Mapping protein post-translational modifications with mass spectrometry. *Nature methods* [online]. vol. 4, no. 10, s. 798–806, [vid. 3. August 2013]. ISSN 1548-7091. Dostupné z: doi:10.1038/nmeth1100

WOLSTENCROFT, K J, R STEVENS, L TABERNERO and a BRASS, 2005. PhosphaBase: an ontology-driven database resource for protein phosphatases. *Proteins* [online]. vol. 58, no. 2, s. 290–4, [vid. 22. August 2013]. ISSN 1097-0134. Dostupné z: doi:10.1002/prot.20325

WU, Hsin-Yi, Vincent S TSENG, Lien-Chin CHEN, Hui-Yin CHANG, I-Chi CHUANG, Yeou-Guang TSAY and Pao-Chi LIAO, 2010. Identification of tyrosine-phosphorylated proteins associated with lung cancer metastasis using label-free quantitative analyses. *Journal of proteome research* [online]. vol. 9, no. 8, s. 4102–12., ISSN 1535-3907. Dostupné z: doi:10.1021/pr1006153

WU, Chang-Jiun, Tianxi CAI, Klarisa RIKOVA, David MERBERG, Simon KASIF and Martin STEFFEN, 2009. A predictive phosphorylation signature of lung cancer. *PloS one* [online]. vol. 4, no. 11, s. e7994, [vid. 19. August 2013]. ISSN 1932-6203. Dostupné z: doi:10.1371/journal.pone.0007994

ZANZONI, Andreas, Daniel CARBAJO, Francesca DIELLA, Pier Federico GHERARDINI, Anna TRAMONTANO, Manuela HELMER-CITTERICH and Allegra VIA, 2011. Phospho3D 2.0: an enhanced database of three-dimensional structures of phosphorylation sites. *Nucleic acids research* [online]. vol. 39, no. Database issue, s. D268–71, [vid. 22. August 2013]. ISSN 1362-4962. Dostupné z: doi:10.1093/nar/gkq936

ZHOU, H, J D WATTS and R AEBERSOLD, 2001. A systematic approach to the analysis of protein phosphorylation. *Nature biotechnology* [online]. vol. 19, no. 4, s. 375–8., ISSN 1087-0156. Dostupné z: doi:10.1038/86777