

# Moderní metody analýzy genomu: aplikace technologie NGS

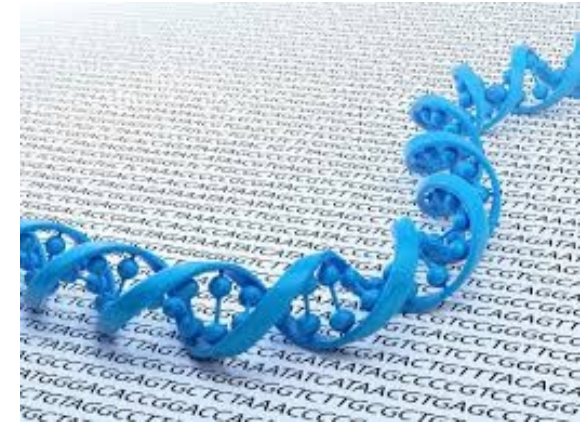
doc. Mgr. Martin Trbušek, Ph.D.

Interní hematologická a onkologická klinika, FN Brno

Lékařská fakulta MU

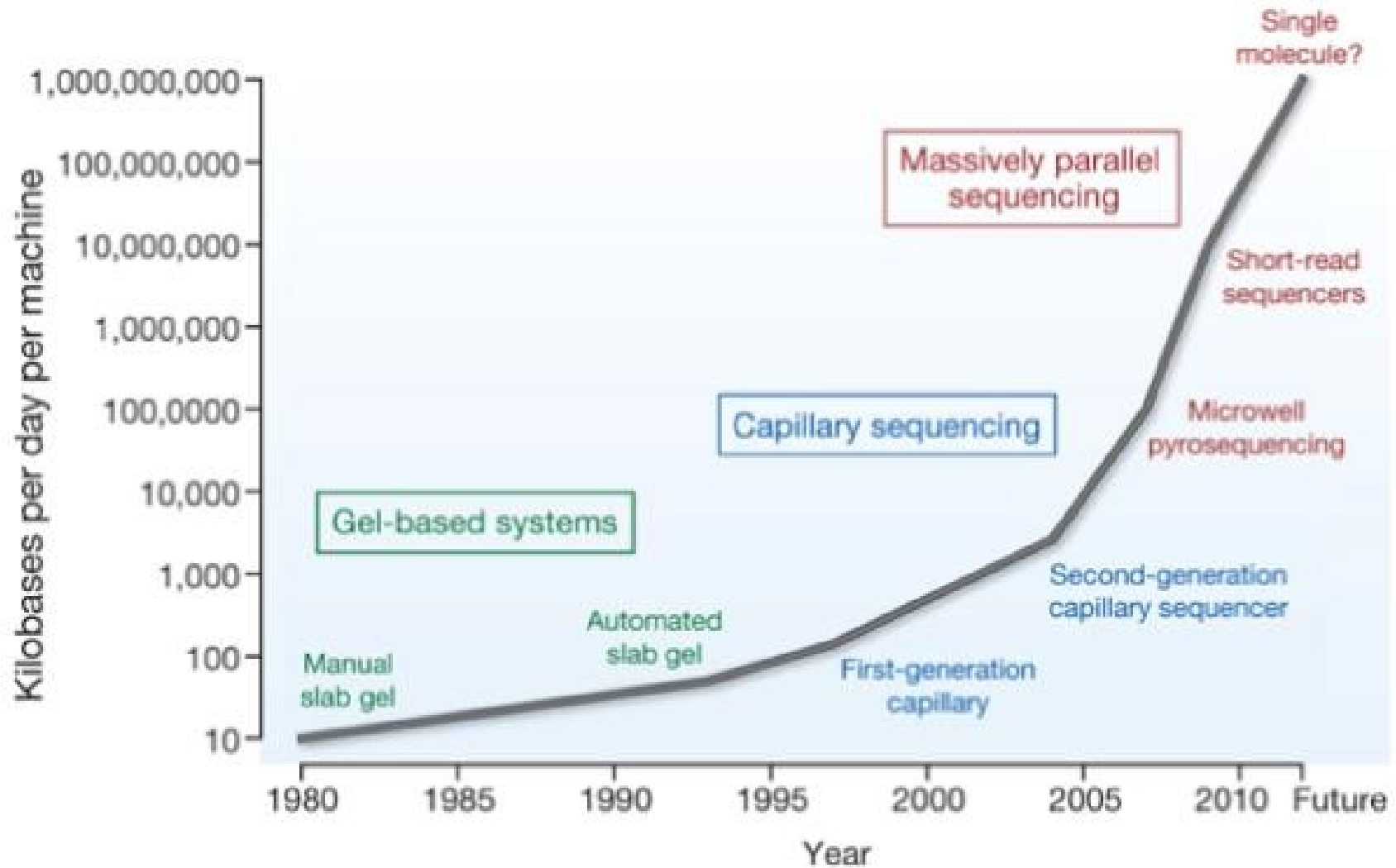


# Obsah

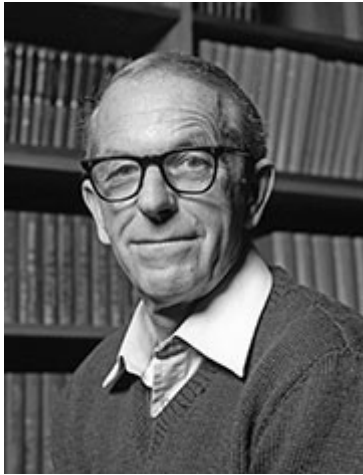


- Metodické přístupy NGS v onkologii
- *Somatická mutační teorie vzniku nádorů*
- Současné využití NGS v biomedicíně a onkologii
- Zkušenosti našeho pracoviště s NGS

# Vývoj sekvenování DNA



# Sangerovo sekvenování – klasický gel

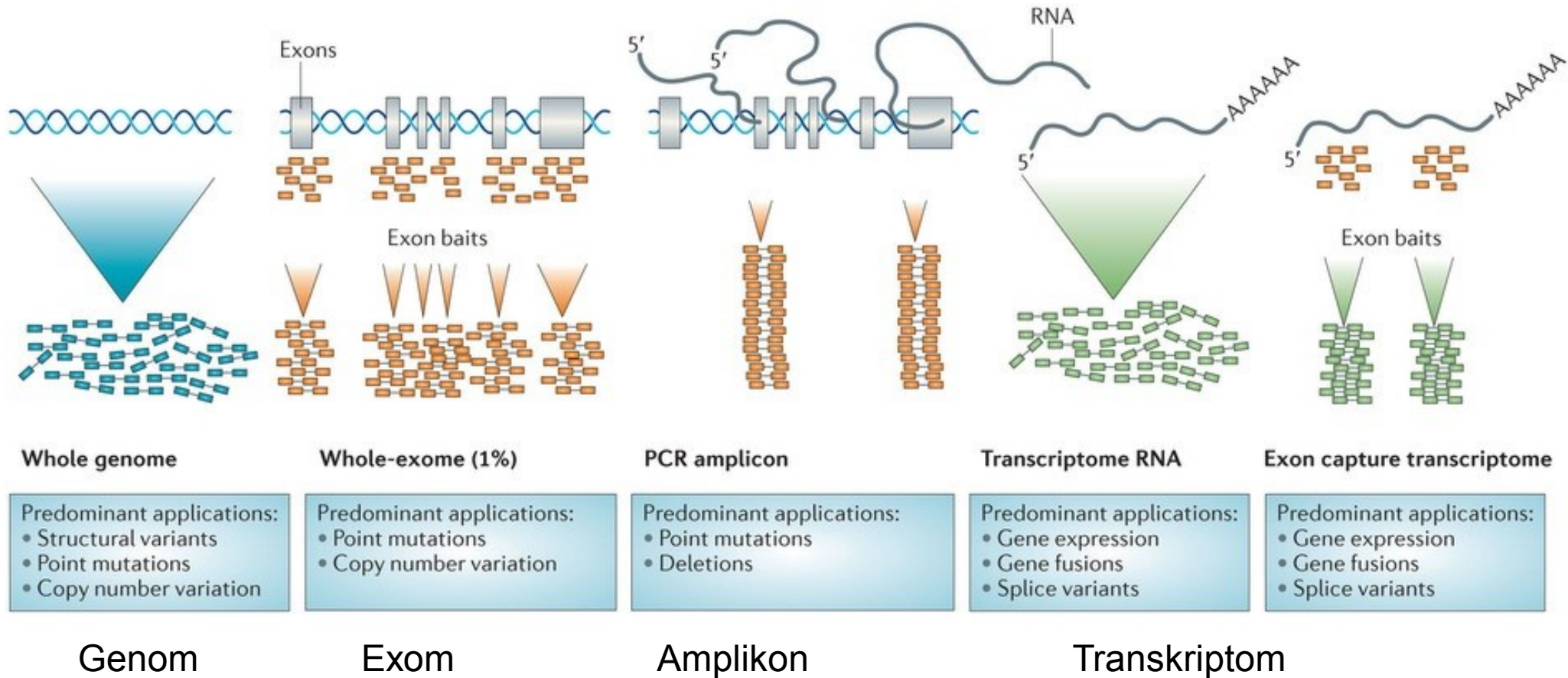


**Frederick Sanger**  
University of Cambridge



PAGE  
značení  $^{35}\text{S}$

# Varianty NGS: analyzovaná oblast



# Využití NGS vedle sekvenování

- SNV = single nucleotide variants (mutace/SNP)
- CNV = copy number variation (inzerce/delece)
- strukturní aberace (translokace/inverze)
- genová exprese (mRNA)
- epigenetika (metylované oblasti)
- interakce DNA-protein
- analýza technologie CRISPR



# Původ nádorů: koncepční teorie

## Somatic mutation theory (SMT)

VS.

## Tissue organization field theory (TOFT)

### **SMT:**

Základní nastavení buňky je klidové stádium a rakovina představuje „únik“ z tohoto stavu.

Maligní buňka vykazuje – díky svým genetickým změnám - selektivní výhodu růstu oproti zdravému protějšku.

### **TOFT:**

Základní nastavení buňky je nekonečná proliferace (blbost?....možná ne.....)

Jsou to naše tkáně, kdo drží buňky pod kontrolou (v klidu) a jejich dezorganizace vede k „obnově“ buněčného dělení

# Celogenomové sekvenování (WGS)

Sekvenace celé chromozomální DNA → úplná informace o genomu (pokryty i promotorové a regulační sekvence)

- 1. *De novo* assembly** - využívá překryvů sekvencí, předpokladem dostatečné pokrytí (>10x)
- 2. Resekvenování** - mapování na referenční sekvenci



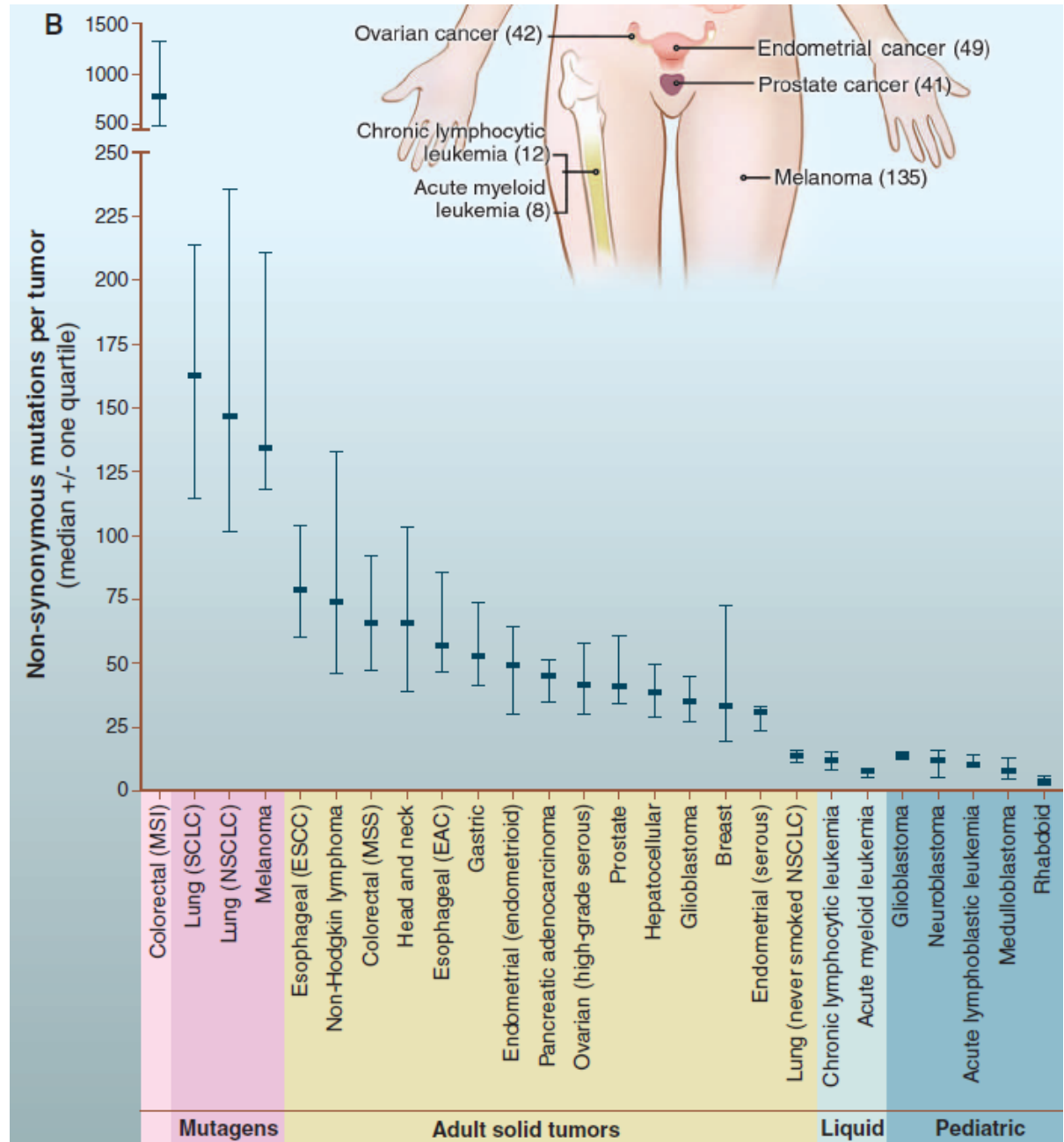
# WGS a analýza nádorů z různých tkání

Vysoce variabilní počet mutací

„Driver“ mutace  
vs.  
„Passenger“ mutace

Driver geny: ~125  
71 TS/54 ONC

PRINCIPY  
DARWINOVSKÉ SELEKCE

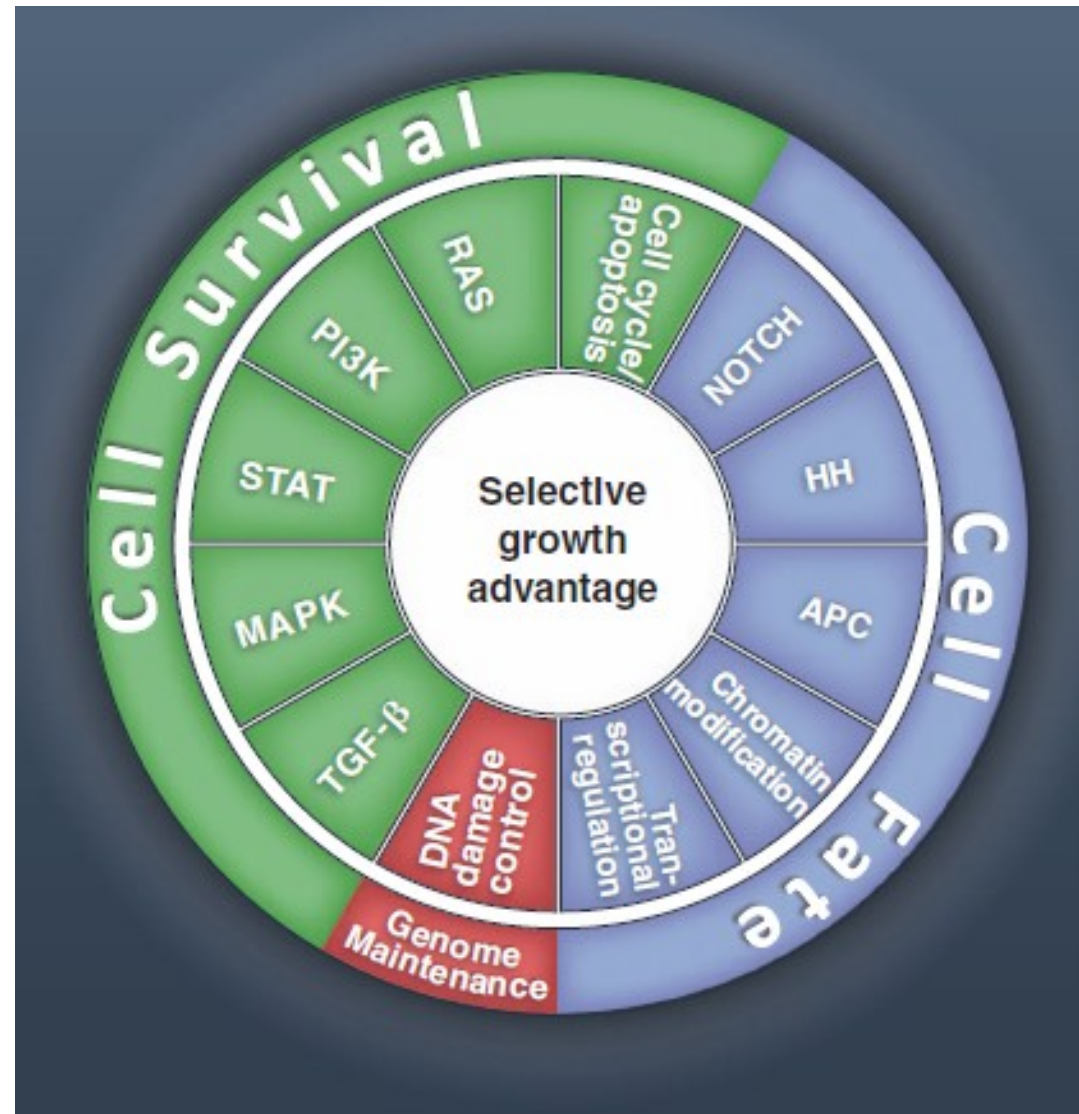


# 12 biochemických drah narušených u nádorových buněk

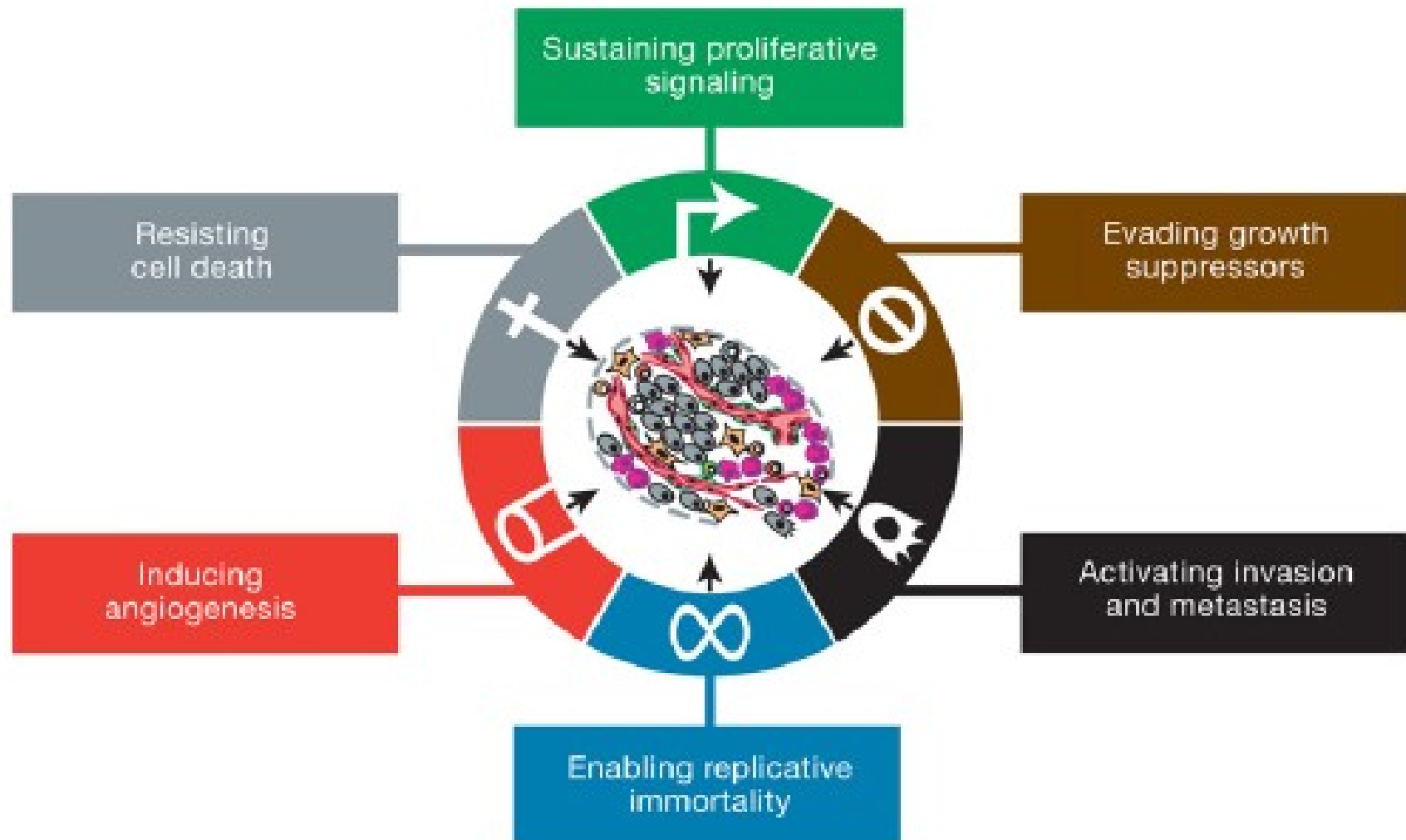
## Interpretace WGS

99.9% všech identifikovaných změn neposkytuje **žádnou selekční výhodu**

**Mutabilita** lidského genomu **je normální**. Nicméně normální je také to, že se tělo dokáže zbavit narušených buněk pomocí kontinuálně probíhající apoptózy....



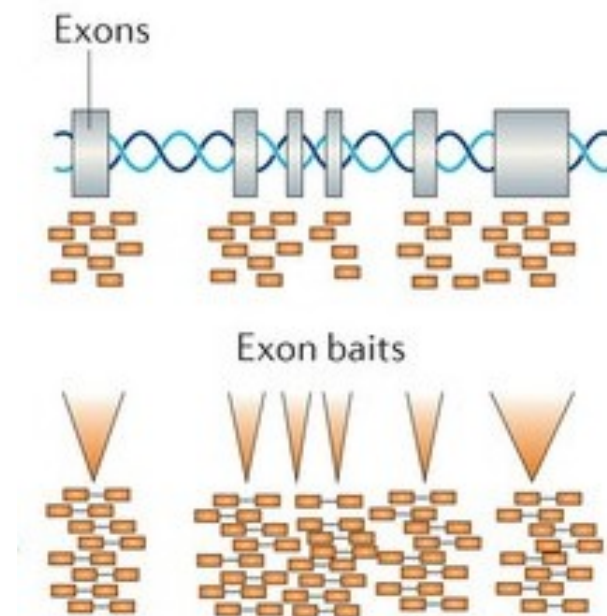
# Základní „schopnosti“ nádorových buněk



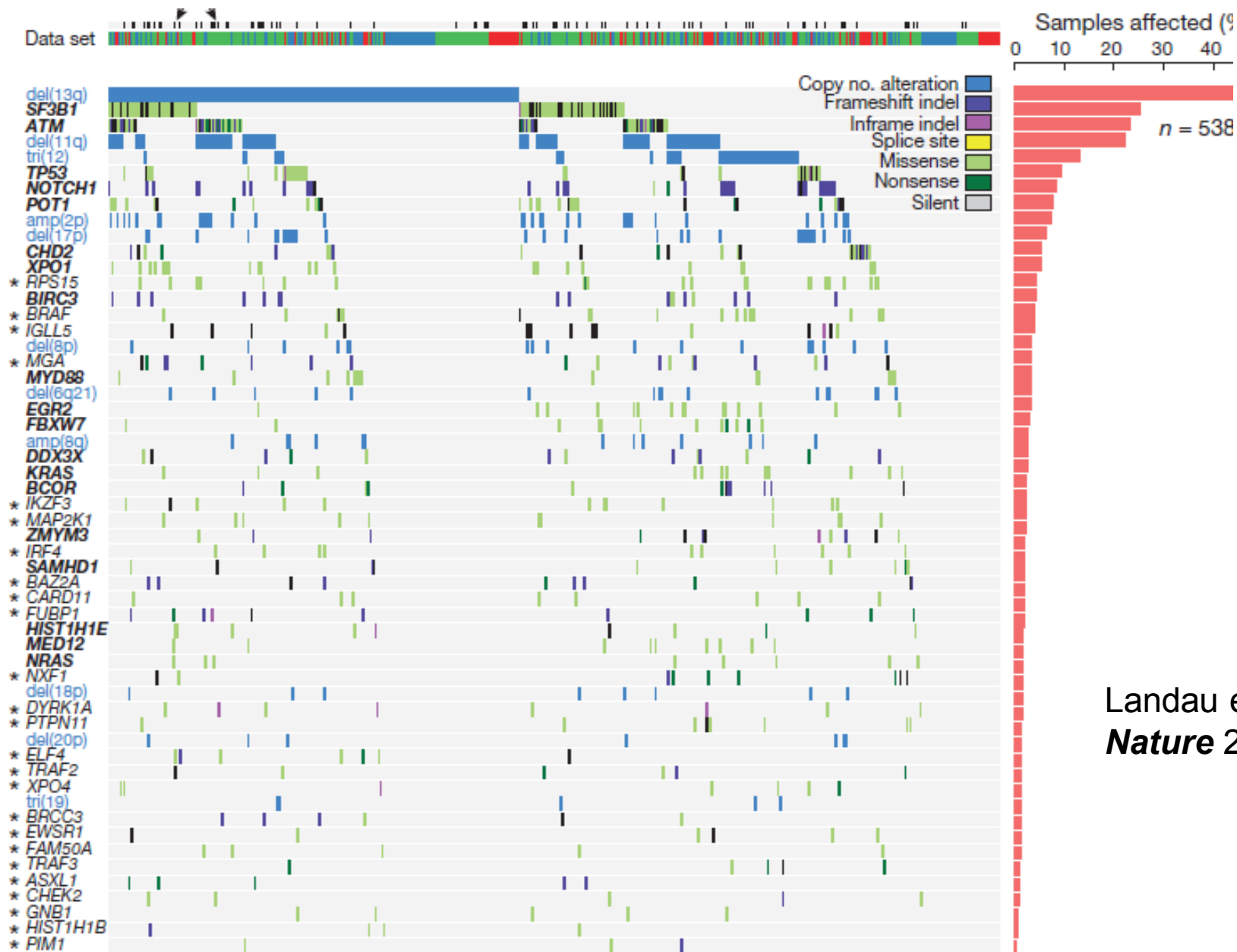


# Exomové sekvenování

- WES = whole exome sequencing
- Sekvenování kódujících oblastí = **exom** (asi 1 % genomu)
- **Efektivnější než WGS**: rychlost, cena, vyšší pokrytí



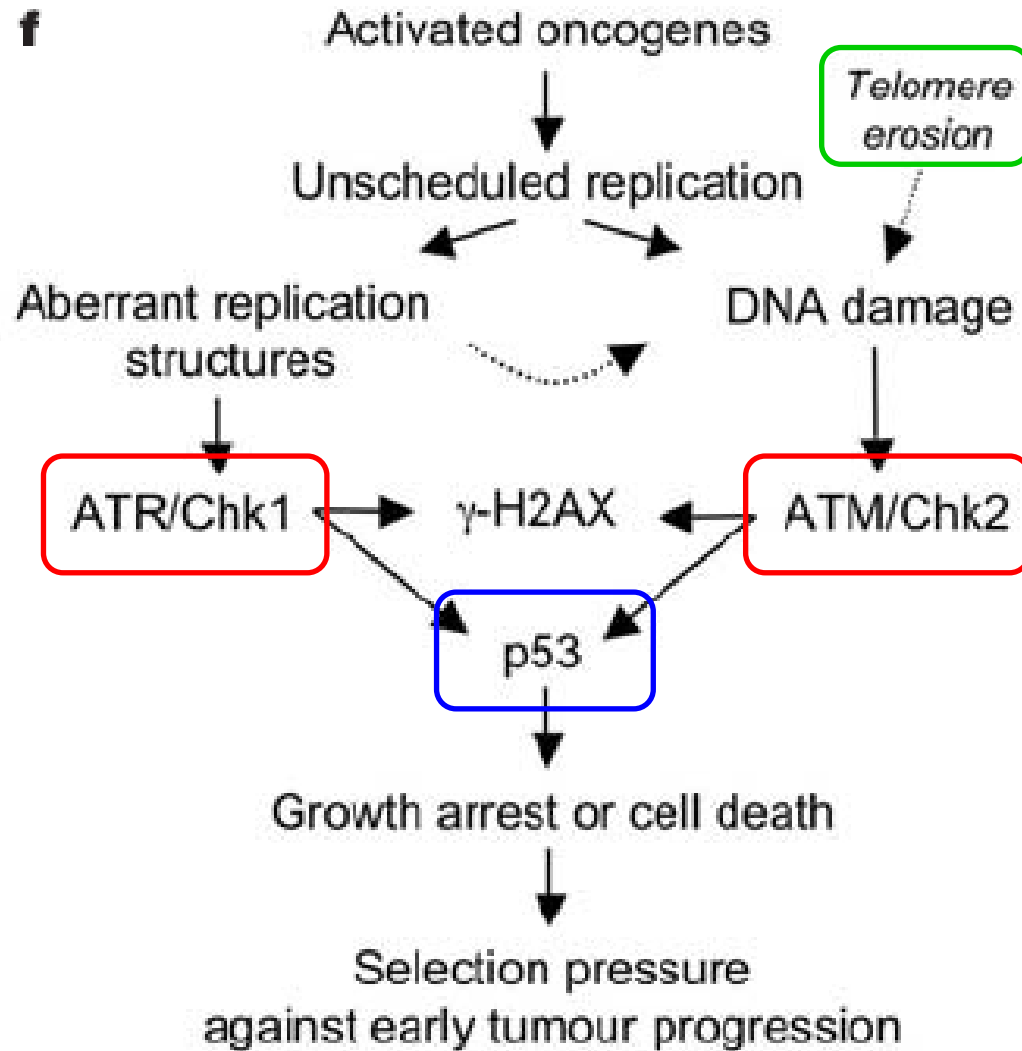
# WES a rekurentní mutace u CLL



Landau et al.,  
*Nature* 2015

Nejčastěji mutované geny: SF3B1, ATM, TP53

# Model vzniku a progresu nádorů



# Cílené sekvenování

- Targeted sequencing
- Princip: PCR amplifikace nebo využití cílených sond
- Identifikace mutačních variant pod detekčním limitem Sangerova sekv. (až 1% při vysokém pokrytí – tzv. *ultra-deep NGS*)
- Výhodné pro sledování **klonální evoluce nádorů**
- **Výhody oproti WGS / WES:** rychlost, cena, méně prostoru pro skladování dat
- Pro sekvenování velkého počtu vzorků (screening) nebo validaci genetických variant v populaci





# Tři způsoby přípravy knihovny

1. **multiplex PCR**: AmpliSeq (Life Tech.) až 6144 párů primerů
2. **single-plex PCR**: Microdroplet PCR (RainDance Tech.), Access Array System (Fluidigm)
3. **targeted capture** (cílený „záchyt“ sondami) s následnou multiplex PCR: TrueSeq Amplicon (Illumina), HaloPlex (Agilent Tech.), SeqCap EZ technology (Roche NimbleGen)

# Výstup cíleného NGS (gen ATM)

Count	Coverage	Frequency	Gene_function	RefGene	Exon_number	cDNA	Codon
1752	1752	100	exonic	ATM	exon40	c.5948A>G	p.N1983S
2261	2452	92,21	exonic	ATM	exon22	c.3161C>G	p.P1054R
690	2962	23,3	exonic	ATM	exon50	c.7311C>A	p.Y2437X
100	1203	8,31	exonic	ATM	exon24	c.3433_3435del	p.1145_1145del
74	1433	5,16	exonic	ATM	exon30	c.4578C>T	p.P1526P
46	1281	3,59	exonic	ATM	exon43	c.6258T>A	p.Y2086X
243	8231	2,95	splicing	ATM	exon19	c.2921+1G>A	p.P962Q
19	699	2,72	exonic	ATM	exon25	c.3705_3709del	p.P1235fs
25	1087	2,3	exonic	ATM	exon5	c.480delT	p.S160fs
24	1046	2,29	exonic	ATM	exon5	c.483G>C	p.Q161H
67	3357	2	exonic	ATM	exon26	c.3837G>A	p.W1279X
73	5626	1,3	exonic	ATM	exon26	c.3952_3960del	p.1318_1320del
64	5151	1,24	exonic	ATM	exon49	c.7181C>T	p.S2394L
11	904	1,22	exonic	ATM	exon63	c.9022C>T	p.R3008C
42	3514	1,2	exonic	ATM	exon10	c.1402_1403del	p.K468fs

# Klinické využití NGS



- Nejvhodnější „targeted capture“
- **Komerční kity:**
  - diagnostické kity (panely genů různých onemocnění)
  - nádorové panely (záchyt hereditárních nádorových onemocnění)
  - panely genů dle přání zákazníka (až příliš mnoho nabídek...)

# Původ rakoviny: role přímé dědičnosti

Dědičné nádory (včetně zárodečných nádorových syndromů)  
5-10% všech případů rakoviny

Např. syndrom *Li-Fraumeni* asociovaný s mutacemi v TP53  
nebo *xeroderma pigmentosum* zahrnující mutace v genech  
pro opravu DNA

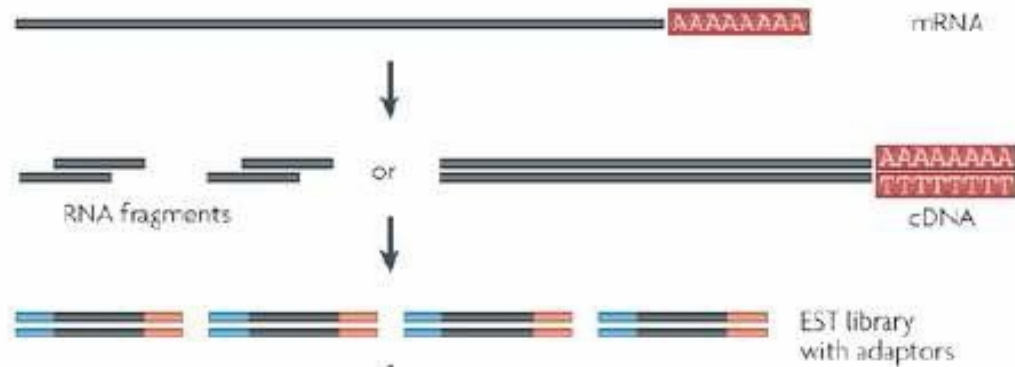
Sporadické nádory – zbytek, původ v somatické tkáni

Genetické defekty jsou příčinou v obou případech;

Navíc, 15-20% nádorových onemocnění zahrnuje infekční *agens*

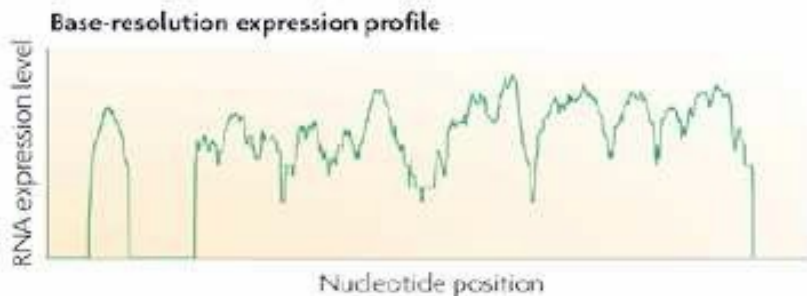
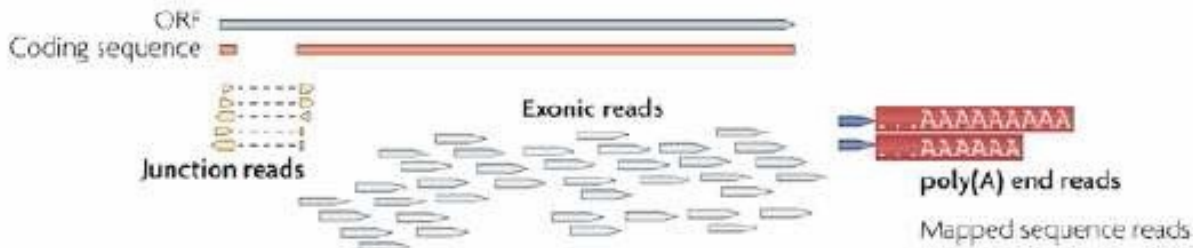
e.g. vysoce rizikové HPVs u karcinomu děložního čípku

# Sekvenování transkriptomu (RNA seq)



ATCACAGTGGGACTCCATAAATTTTCT  
CGAAGGACCAGCAGAAACGAGAGAAAA  
GGACAGAGTCCCAGCGGGCTGAAGGGG  
ATGAAACATTAAAGTCAAACAATATGAA  
.....

Short sequence reads

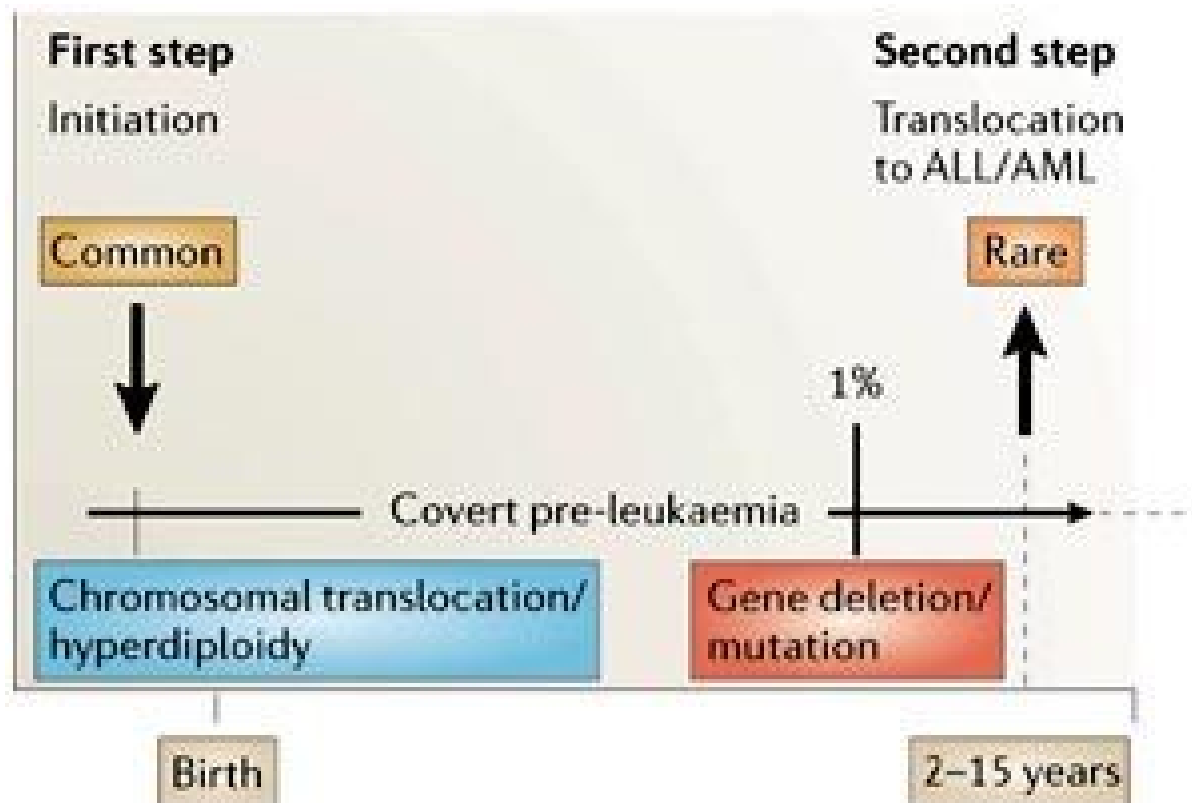


**Transkriptom =**  
soubor všech molekul  
RNA (mRNA, rRNA,  
tRNA a noncoding  
RNA)

Analýza fúzních (driver) genů

Wang et al. *Nat Biotechnol* 2009

# (Specifická) etiologie dětských leukémií

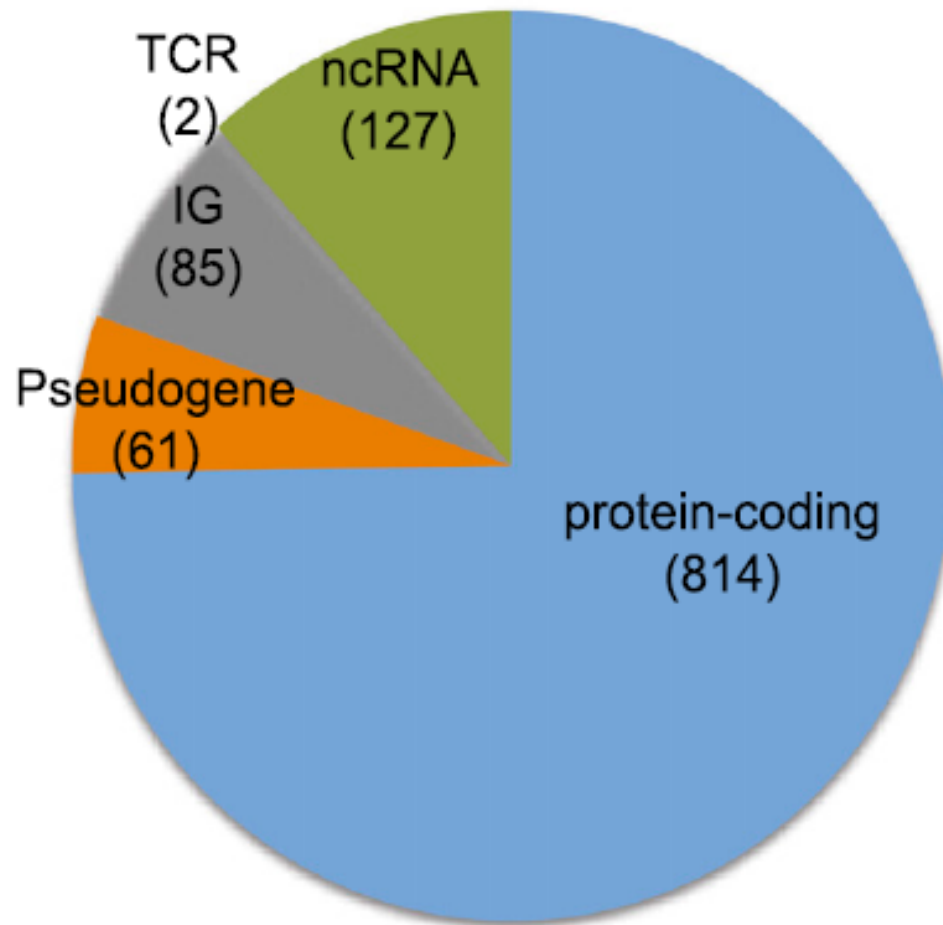


Analýza „Guthriho kartiček“  
nebo pupečnickové krve

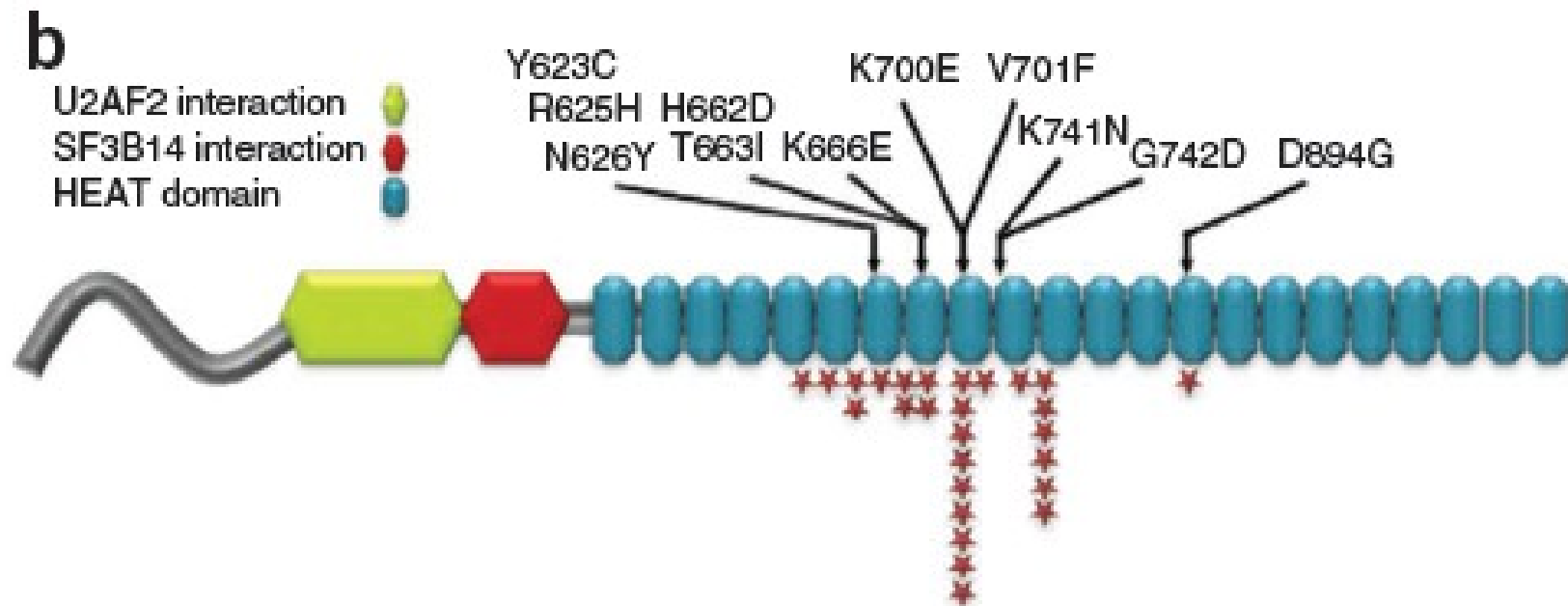
Copyright © 2006 Nature Publishing Group  
Nature Reviews | Cancer

...u jednovaječných dvojčat

# Aberantní exprese u CLL lymfocytů (RNA seq)



# Identifikace mutací SF3B1 u CLL: synergie různých přístupů NGS



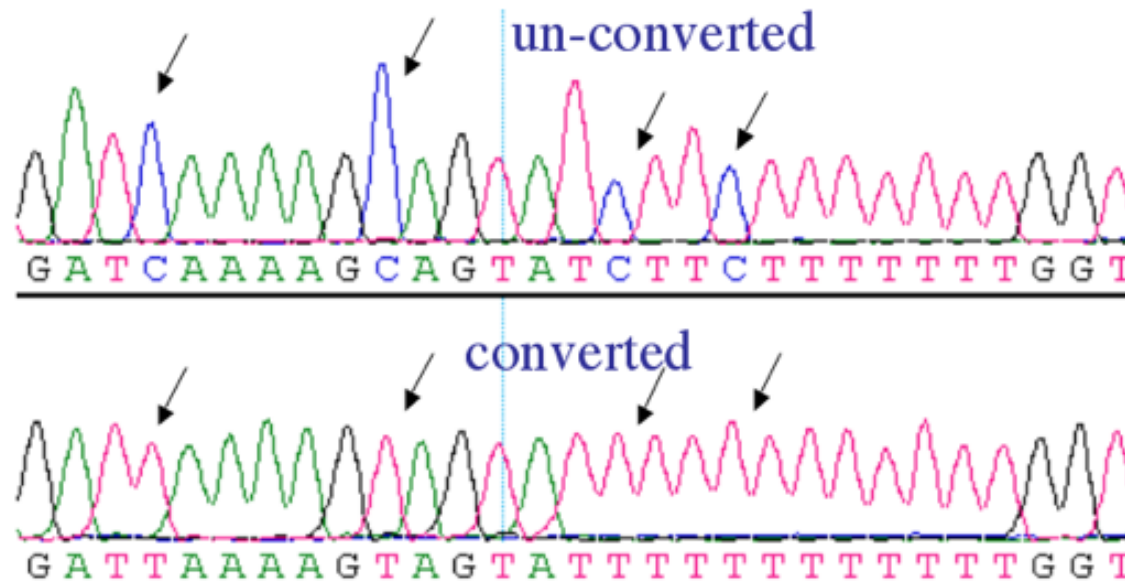


# NGS a epigenetika (metylace DNA)

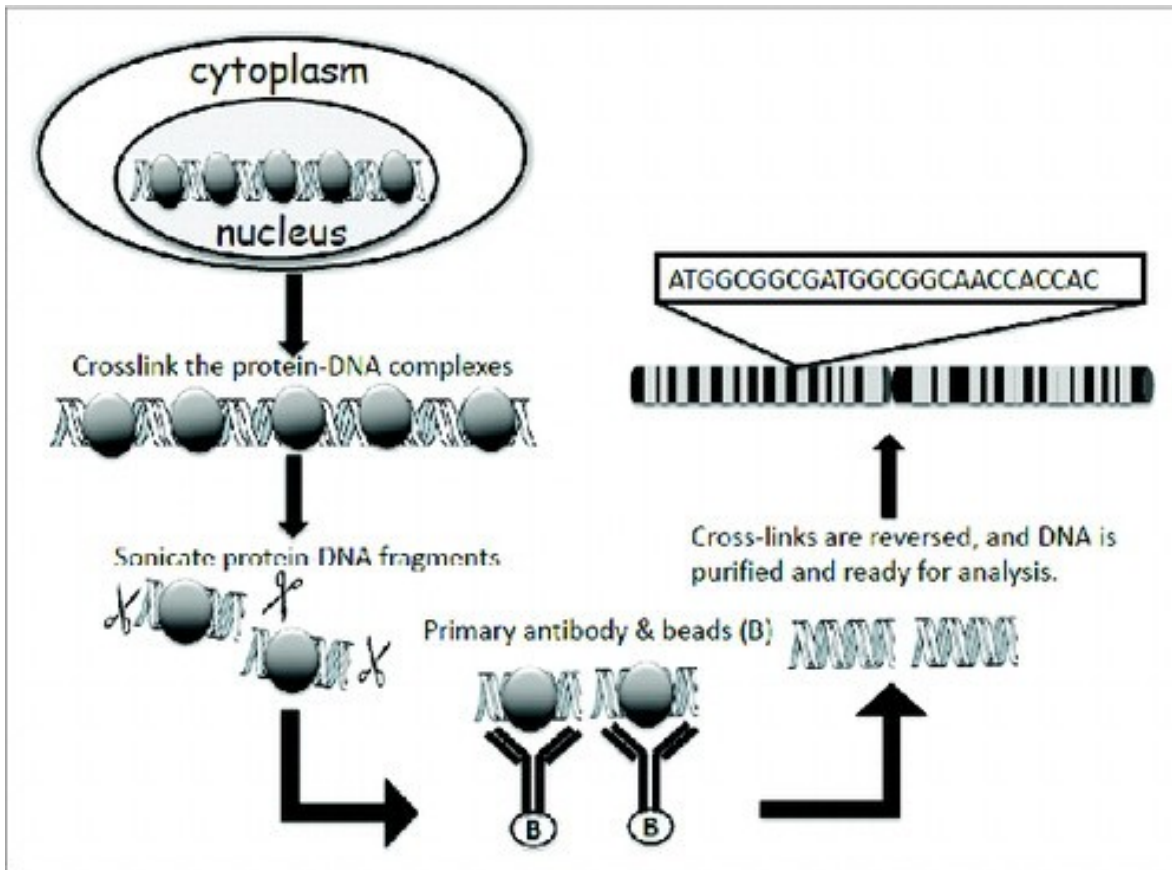
- Epigenetické umlčení tumor-supresorových genů (metylace promotorů)
- Globální (celogenomová) hypometylace

## Bisulfitová konverze + NGS:

- konverze C → U, Met-C se nemění
- přesná identifikace jednotlivých metylovaných bazí



# Chromatinová imuno-precipitace (ChIP-Seq)



Sledování **interakcí mezi proteiny, DNA a RNA** → vazebná místa pro TF, histony, a další proteiny

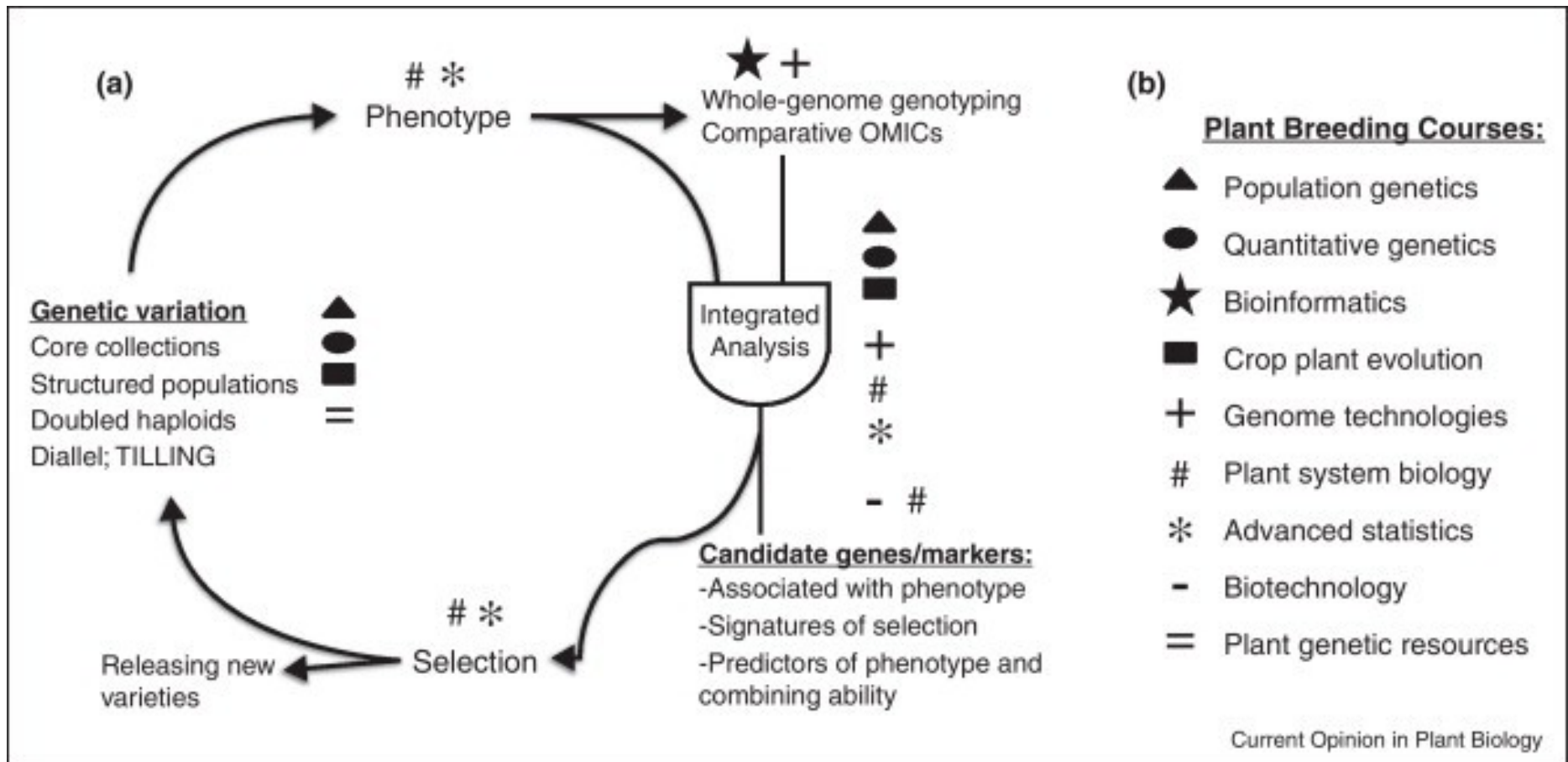
**Regulace genové exprese, epigenetické modifikace chromatinu**

# Reálné výstupy aplikace NGS

# Mezioborové využití NGS

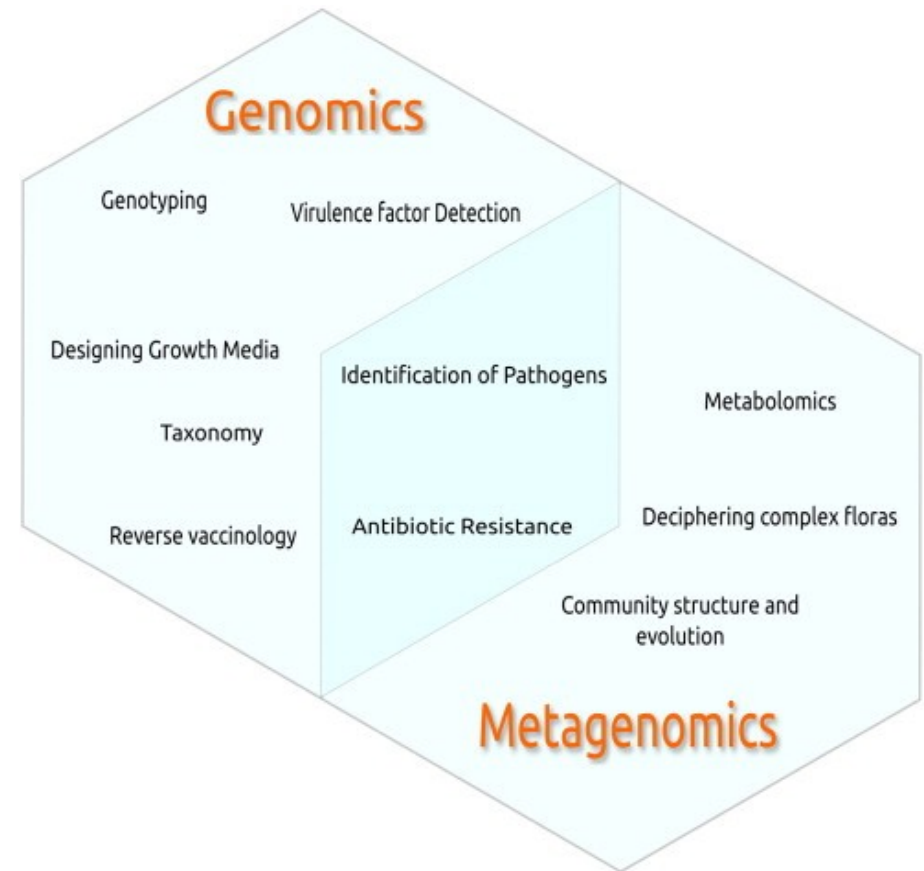
Studium **druhové diverzity** (genotypizace):

- fylogeneze živočišných druhů (Ellegren et al., 2012)
- identifikace nových virových variant (Kapgate et al., 2015)
- šlechtění plodin v zemědělství (Fridman et al., 2012)



**Metagenomika**: studium mikrobiálního složení v různých typech prostředí (střevní mikroflóra, zubní plak, půda, korálové útesy, mořské dno)

- bez potřeby kultivace
- identifikace patogenů, virulence, rezistence, atd.



# NGS a forenzní genetika

Analýza stop  
z místa činu

Polymorfní místa  
v genomu



Expand your forensics  
workflow with NGS:  
STRs, SNPs, and mtDNA

[Learn more >](#)

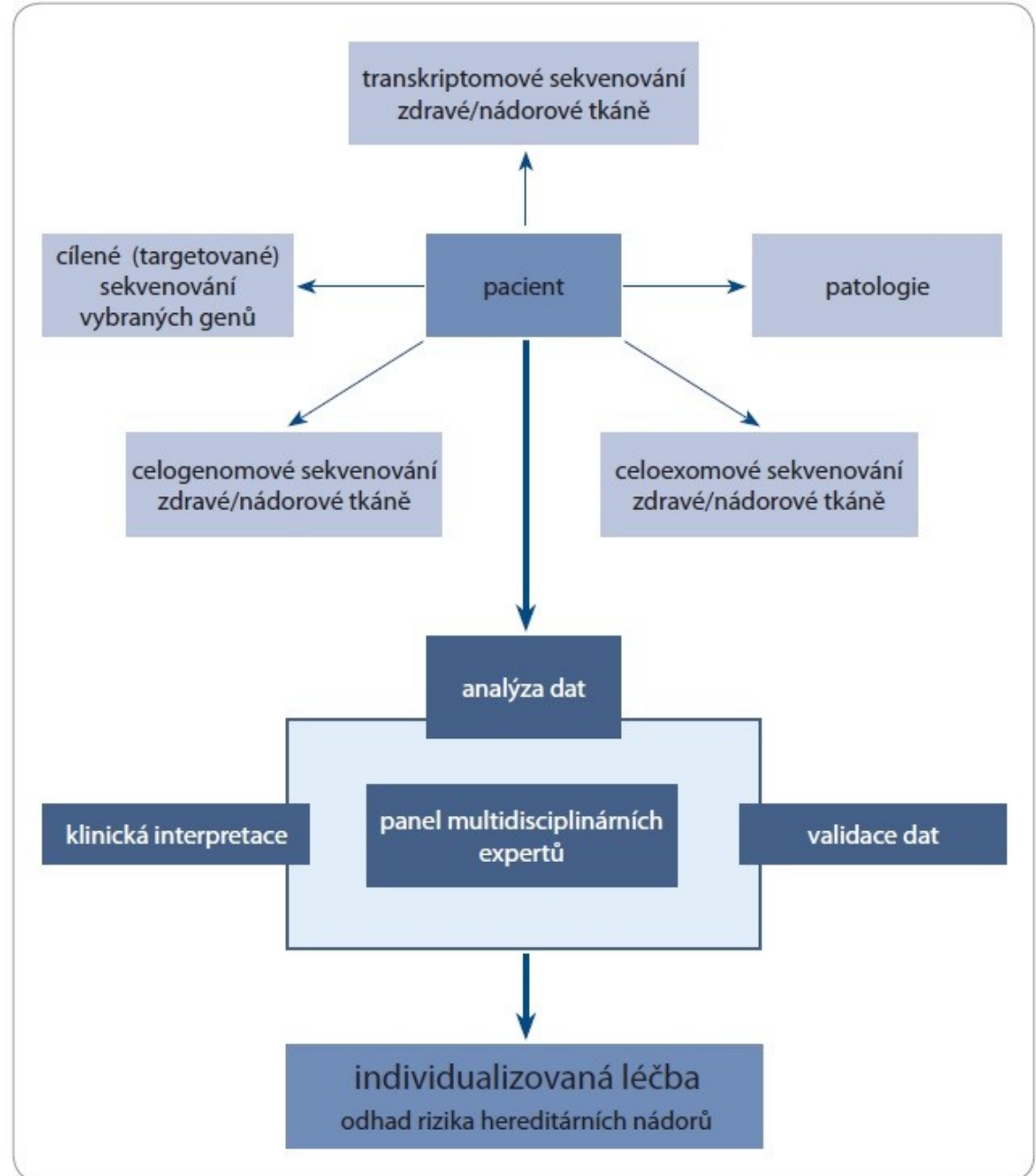
Borsting and Morling, **Next generation sequencing and its applications in forensic genetics.**  
Forensic Int Sci Genet, 2015

# Využití v medicíně



- molekulární diagnostika dědičných chorob a infekčních onemocnění
- prenatální diagnostika (neinvazivní, z fetální DNA v mateřské plazmě)
- farmakogenomika (identifikace nových terapeutických cílů, studium rezistence)
- Onkologie; bezbřehé možnosti!

# Optimální využití NGS: personalizovaná léčba





# První projekt WGS u nádorového onemocnění

Porovnání nádorových a zdravých buněk **pacienta s AML**

→ identifikace dvou známých a osmi nových somatických mutací

(Ley et al., *Nature* 2008)



# Využití NGS v experimentální a klinické onkologii

International Cancer Genome Consortium (ICGC):

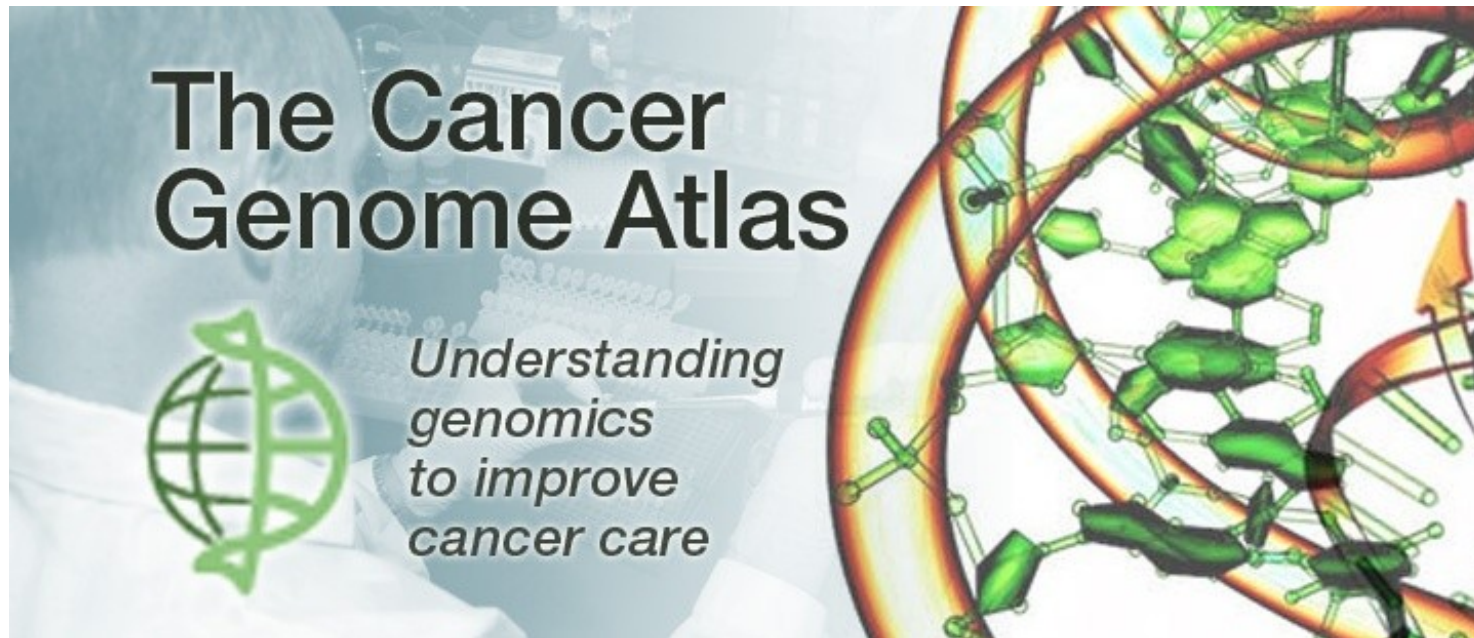
**Cancer Genome Project** – charakterizace genomických, transkriptomických a epigenomických změn u 50 nejčastějších nádorových onemocnění

14 zemí, sekretariát Toronto, Kanada



International  
Cancer Genome  
Consortium

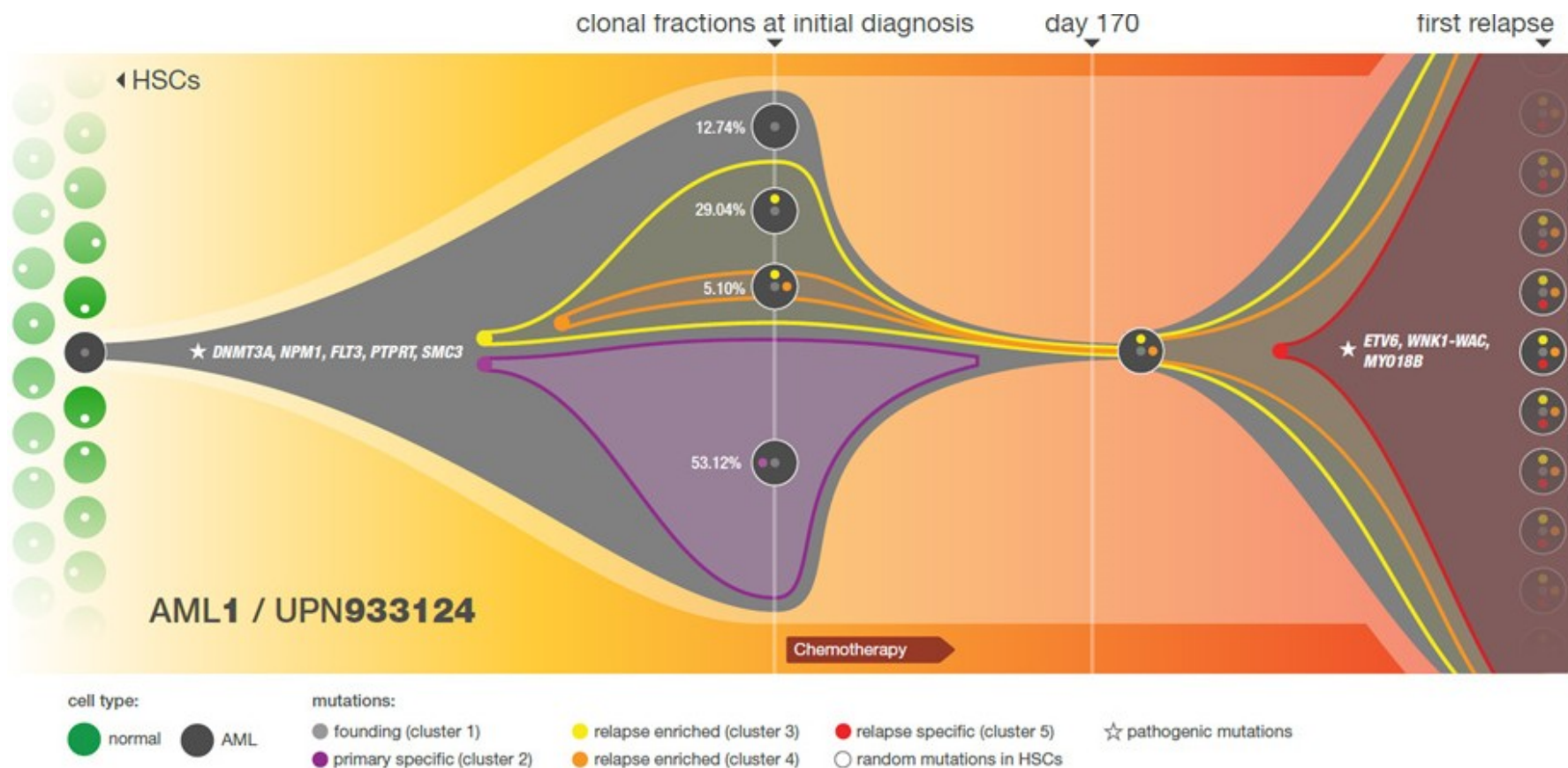
**National Cancer Institute (NCI):** projekt vytvoření atlasu nádorových genů – **The Cancer Genom Atlas (TCGA)** za účelem zlepšit nádorovou prevenci, včasnou detekci a léčbu



Skrínink nových léků na **panelu 60 buněčných linií**  
→ **syntetická letalita**

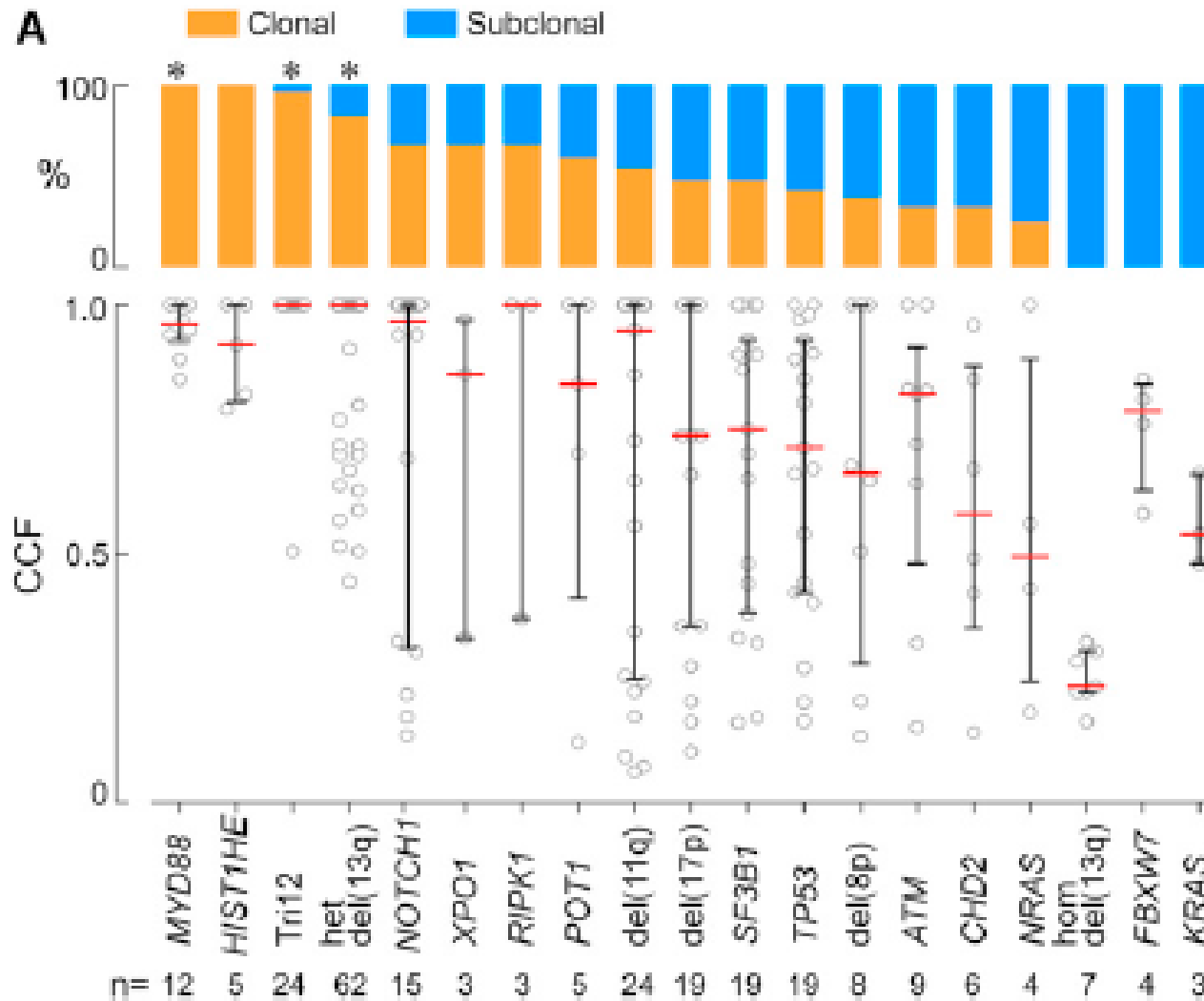
# Studium klonální evoluce AML pomocí WGS

Identifikace somatických mutací objevujících se při relapsu (selekce cytotoxickou terapií)



# Klonální architektura CLL

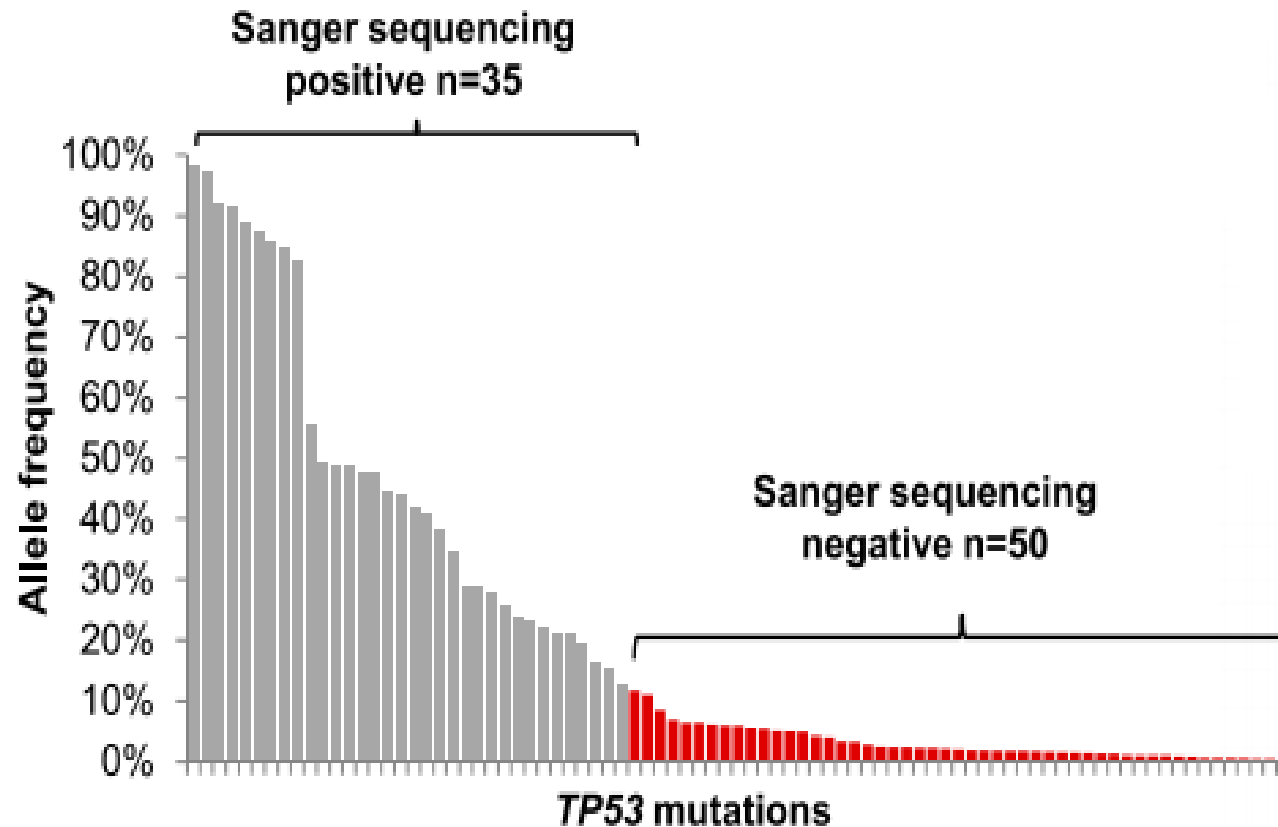
- časně klonální (del13q, tri12, MYD88) a pozdější sub-klonální (TP53, SF3B1) aberace → selekce léčbou a urychlení progresu



# Klinický dopad sub-klonů mut-TP53

– *ultra-deep sequencing* (median alel. frekvence 2%)

→ horší přežití pacientů, evoluce klonu při relapsu s následným vznikem chemorezistence

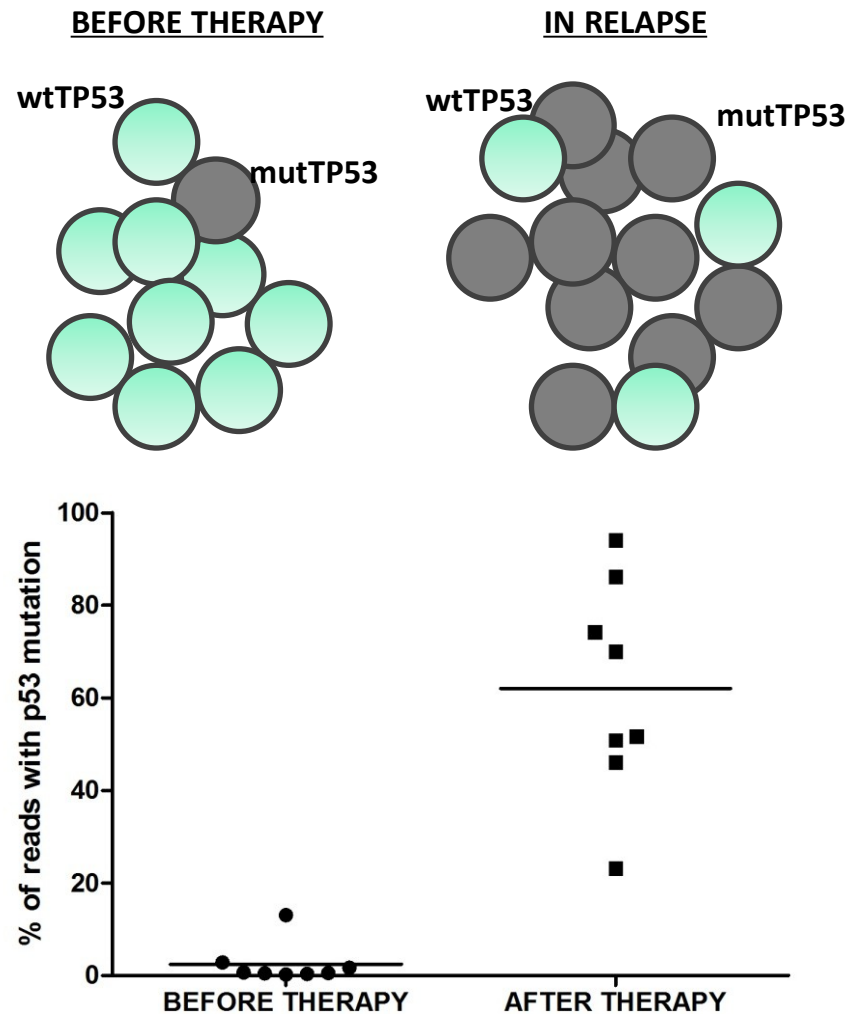


Zkušenosti našeho pracoviště s NGS

# Identifikace sub-klonů mut-TP53 před terapií

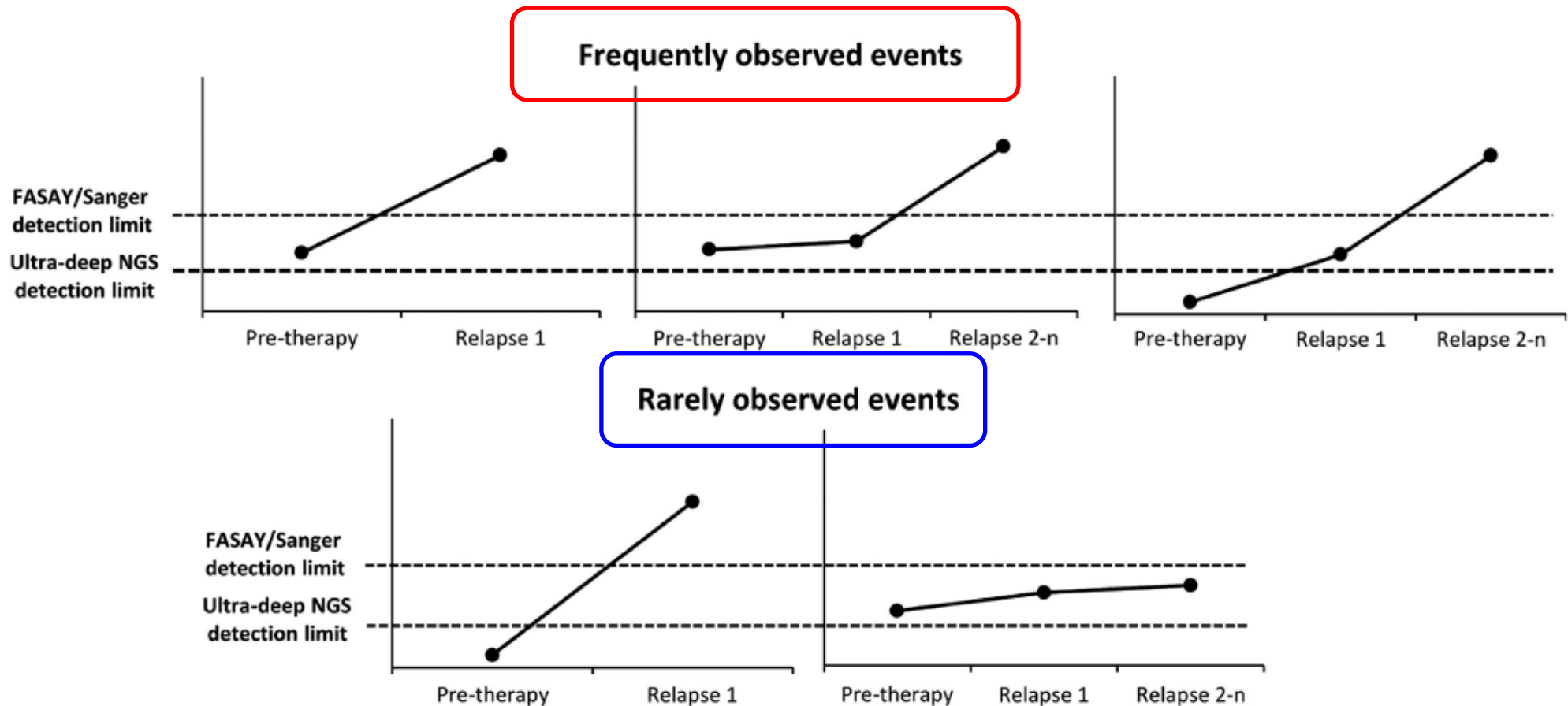
CLL buňky s mutacemi v genu TP53 jsou typicky součástí tzv. **minimální zbytkové choroby**. Frakce buněk rezistentních na terapii je příčinou pozdějšího relapsu onemocnění.

Cílem studie bylo identifikovat **pacienty s minoritními mutacemi v genu TP53** již před nasazením terapie

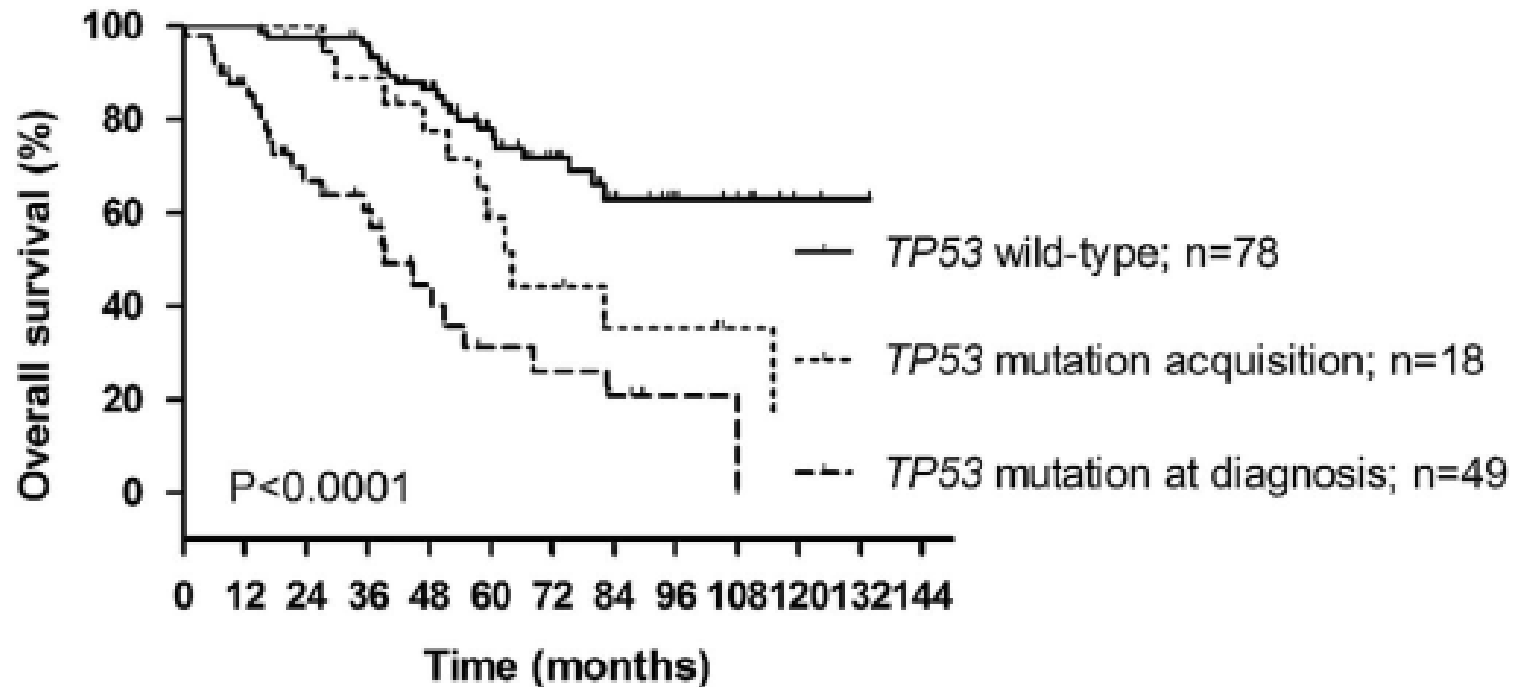




# Více a méně pravděpodobné scénáře selekce mutací v TP53



# Negativní dopad selektovaných mutací na přežití CLL pacientů

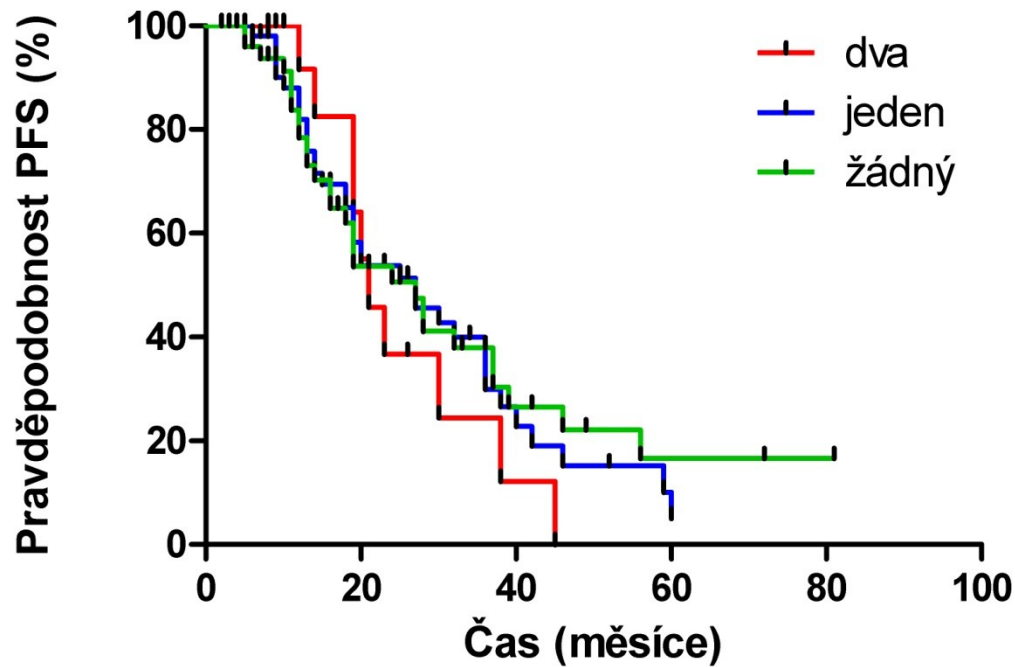


Přežití stanoveno od diagnózy

„Mutation acquisition“ = přítomnost mutace v relapsu

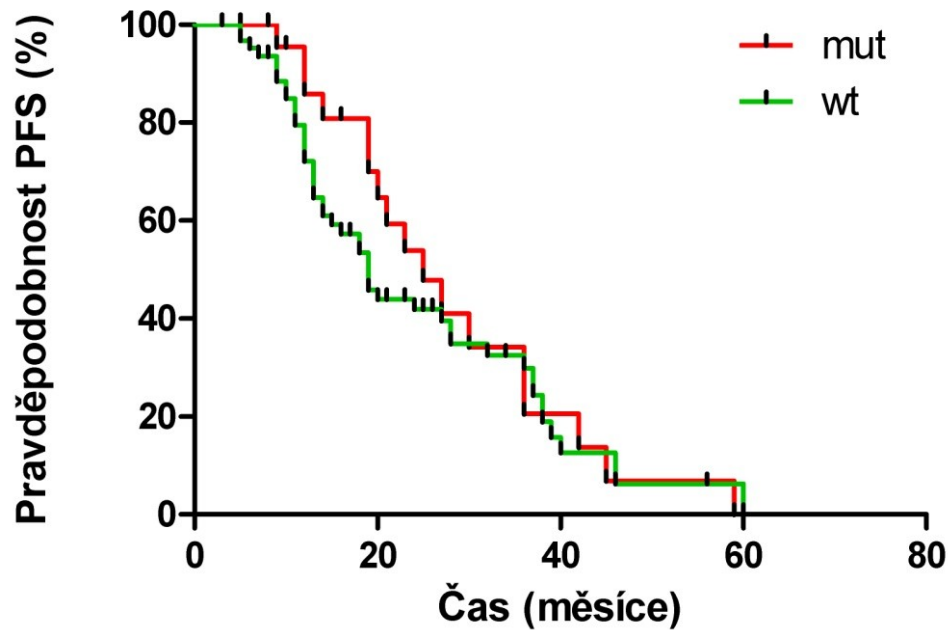
# Analýza genů ATM, SF3B1, NOTCH1, BIRC3 u terapie FCR, BR

PFS dle počtu mutovaných genů

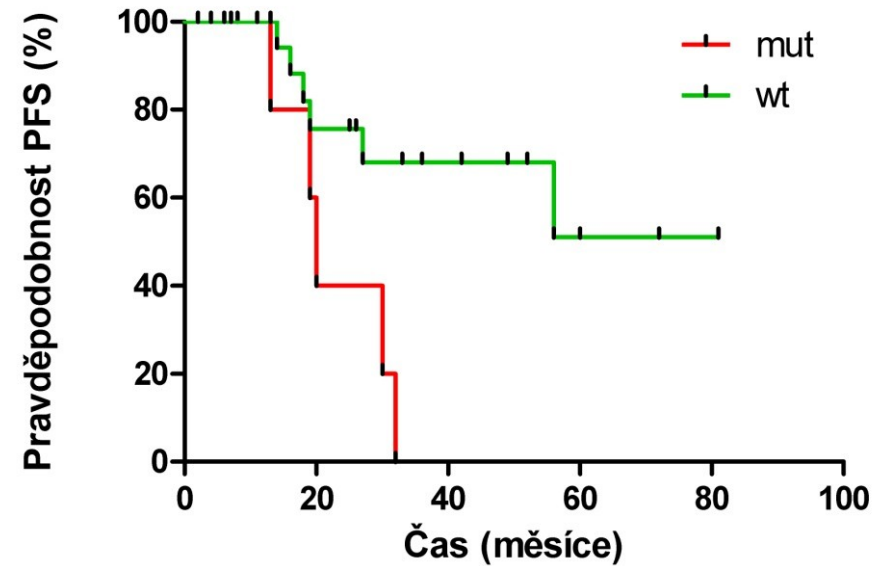


	Medián PFS (m)
2 mut. geny	21
1 mut. gen	27
Bez mutace	27
P = ns	

# PFS dle mutací v genu ATM

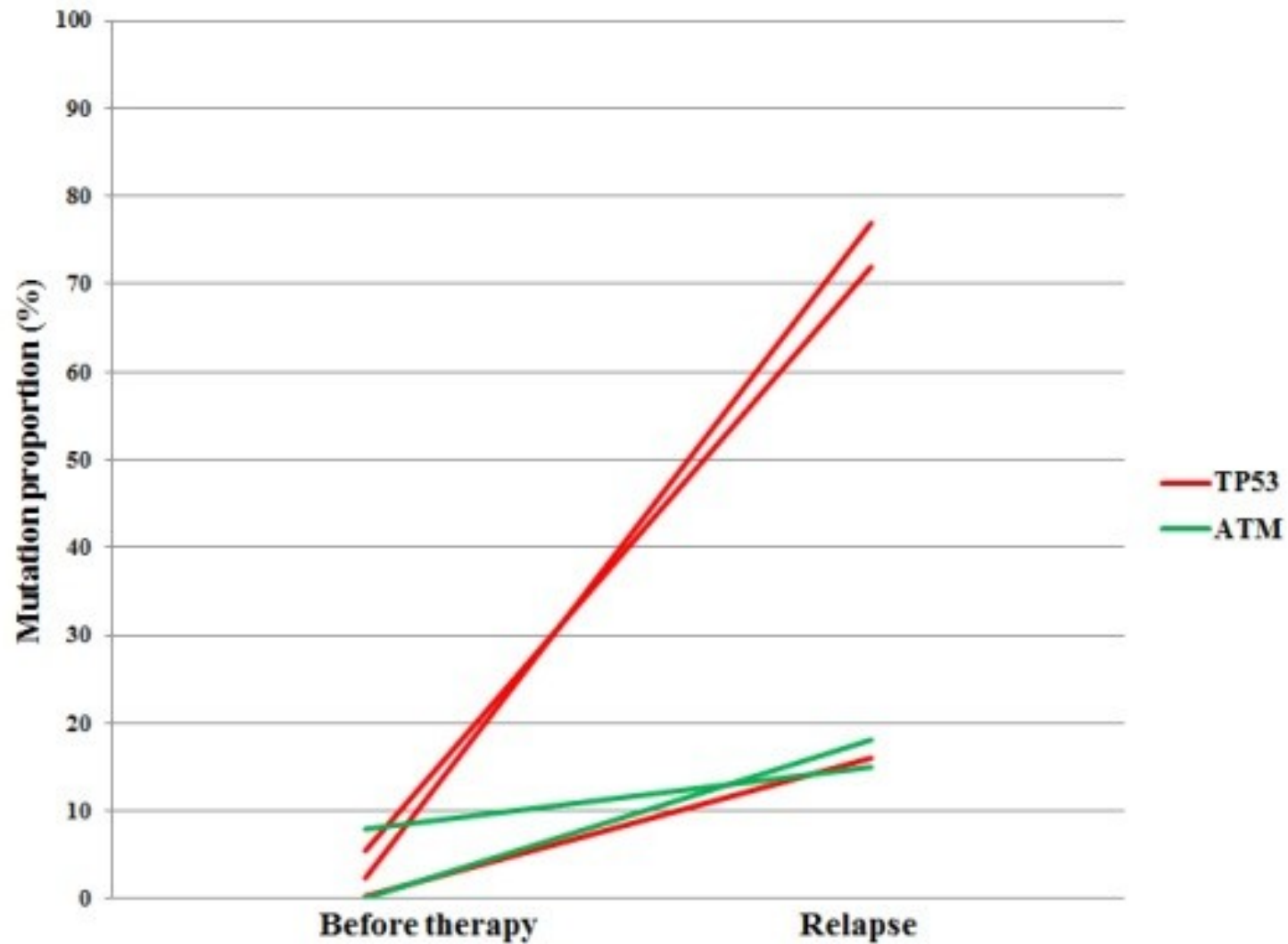


Nemut-IGHV	Medián PFS (m)
ATM-mut	25
ATM-wt	19
<b>P = ns</b>	

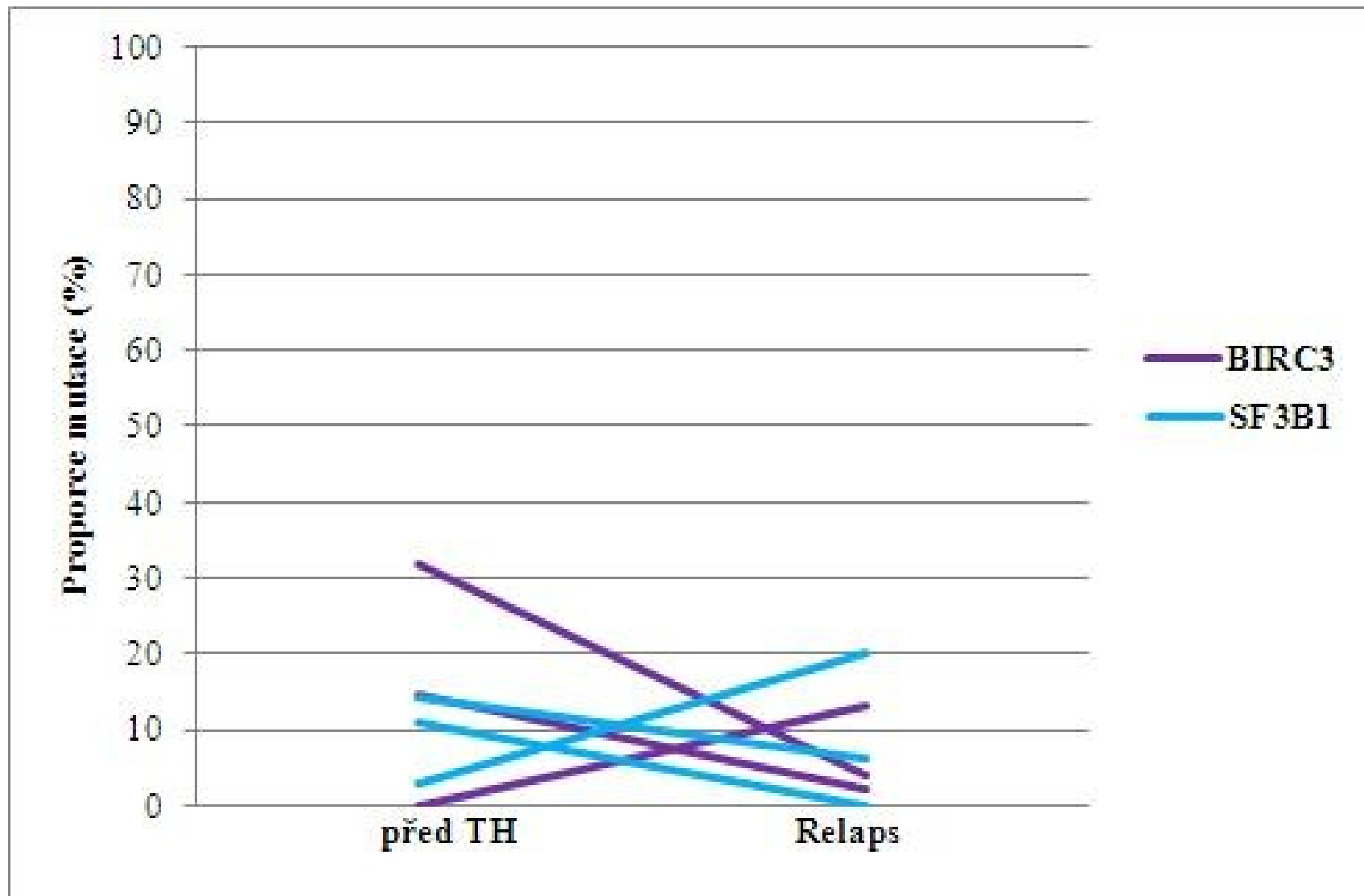


Mut-IGHV	Medián PFS (m)
ATM-mut	20
ATM-wt	nedosažen
<b>P = 0,014</b>	

# Klonální evoluce mutací TP53 a ATM



# Klonální evoluce mutací BIRC3 a SF3B1



# • Klinická a genetická heterogenita

## ➤ 840 typů NMD – 16 skupin

1. Muscular dystrophies

2. Congenital muscular dystrophies

3. Congenital myopathies

4. Distal myopathies

5. Other myopathies

6. Myotonic syndromes

7. Ion channel muscle diseases

8. Malignant hyperthermia

9. Metabolic myopathies

10. Hereditary cardiomyopathies

11. Congenital myasthenic syndromes

12. Motor neuron diseases

13. Hereditary ataxias

14. Hereditary motor and sensory neuropathies

15. Hereditary paraplegias

16. Other neuromuscular disorders

<http://www.musclegenetable.fr>

## ➤ 465 genů

[1. Muscular dystrophies](#)

[2. Congenital muscular dystrophies](#)

[3. Congenital myopathies](#)

[4. Distal myopathies](#)

[5. Other myopathies](#)

[6. Myotonic syndromes](#)

[7. Ion channel muscle diseases](#)

[8. Malignant hyperthermia](#)

[9. Metabolic myopathies](#)

[10. Hereditary cardiomyopathies](#)

[11. Congenital myasthenic syndromes](#)

[12. Motor neuron diseases](#)

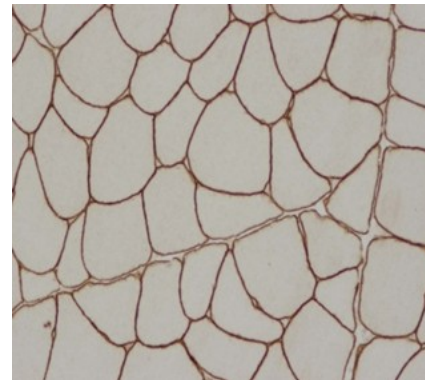
[13. Hereditary ataxias](#)

[14. Hereditary motor and sensory neuropathies](#)

[15. Hereditary paraplegias](#)

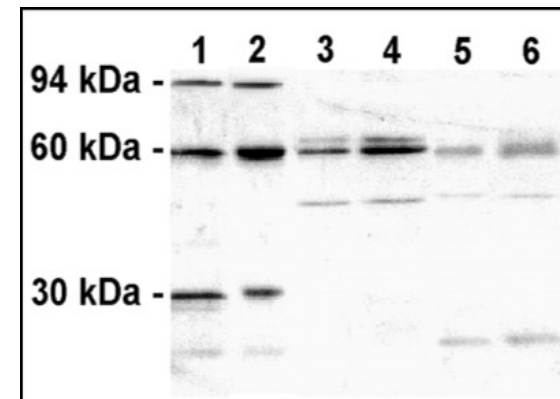
[16. Other neuromuscular disorders](#)

Výběr genu na základě analýzy svalové tkáně (problematické - rutinně se provádí imunohistochemická detekce jen některých proteinů ..... defekt proteinu nemusí vždy korelovat s výsledkem imunohistochemie)



41 + 29 + 25 + 15 + 22 ..... 132 genů ???

- 1) Analýza vytypovaných genů po sobě – „*gene by gene analysis*“, začíná se genem s nepravděpodobnějším výskytem mutace – ČASOVĚ A FINANČNĚ NÁROČNÉ
- 2) Analýza všech genů potenciálně asociovaných s onemocněním současně – masivní paralelní sekvenace (sekvenování nové generace)



Imunohistochemická detekce alfa-sarkoglykanu (kontrola, pacient s mutacemi v genu SGCA), Western blot kalpainu-3. (FNB, PAU, Hermanová M.)



*Genet Med.* 2015 May ; 17(5): 405–424. doi:10.1038/gim.2015.30.

## **Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology**

Sue Richards [Chair, ACMG],

1. Sekvenční varianty patogenní
2. Sekvenční varianty pravděpodobně patogenní
3. Sekvenční varianty nejasného významu
4. Sekvenční varianty pravděpodobně benigní
5. Sekvenční varianty benigní

# Použití NGS – analýza mikroorganismů

1. **Sekvenace rRNA genů** - detekce širokého spektra patogenů — použití sekvenování 16S u bakterií (**mikrobiom**) a 18S u mykotických patogenů (**mykobiom**) – srovnání spektra detekovaných druhů u zdravých a např. chronicky nemocných – cystická fibróza, Crohnova choroba
  - ? virové patogeny
2. **Přímá detekce nových patogenů** – doposud nepopsaných nebo obtížně kultivovatelných (pozor - chybějící sekvence v databázích – de novo sestavení genomu (??))

# Lidský mykobiom

- **Trávící trakt** - 335 druhů ze 158 rodů
  - 221 exkluzivně ve střevě
  - 88 exkluzivně v dutině ústní
  - z celkem 247 druhů nalezených ve střevě bylo vykultivováno jen 29 – zbytek pouze molekulárně biologická detekce
- **výplach úst** od 20 zdravých lidí - 101 druhů, 85 rodů
  - z toho 11 nekultivovatelných
  - **průměrně 15 různých druhů** – nejčastěji *Candida*, *Cladosporium*, následují *Aspergillus*, *Cryptococcus*, *Fusarium* a *Alternaria*, v jiné studii kromě výše uvedených nejčetnější *Malassezia* a *Epicoccum*

Doc. Mgr. Martina Lengerová, Ph.D

# Slabá místa NGS při využití pro klinickou diagnostiku

- **Vysoké „pozadí“ lidské DNA**
  - Deplece – odstranění pomocí speciálních technik
- **Bias v sekvencích?** – v důsledku izolace DNA – zatím neexistuje ideální metoda zpracování DNA pro detekci virů, bakterií a hub v jednom
- **Bioinformatické zpracování** – pro základní analýzy není komplikované, ale přece jen vyžaduje pokročilé znalosti

# Závěr: mikrobiální NGS

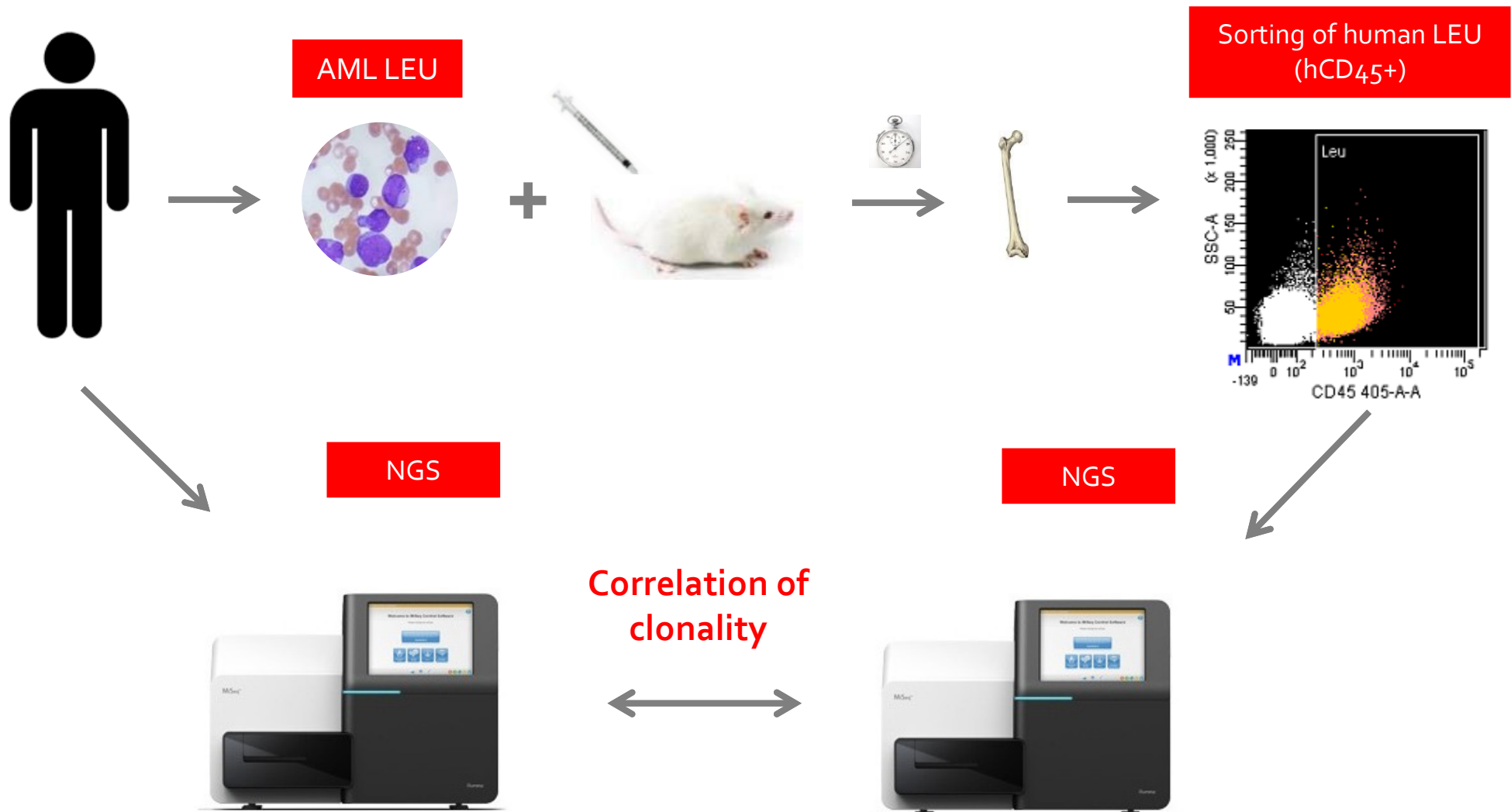
- Za současného stavu techniky může být NGS pouze doplňkem tradičních metod
  - Automatizace
  - Standardizace protokolů a bioinformatických nástrojů pro analýzu dat
  - Správa referenčních databází
  - Snížení nákladů a doby trvání vyšetření
- Velký potenciál má ve studiu interakcí hostitelského a mikrobiálního genomu, epidemiologie a mechanismů vzniku a šíření rezistence

Doc. Mgr. Martina Lengerová, Ph.D

# AZV project CELL AML (PharmDr. Martin Čulen, Ph.D.)

## Subproject A – Xenograft experiments (LF MU Brno)

### Scheme



# Naše zkušenosti s NGS: souhrn „onkologie“

## Pozitivní aspekty:

- Přesnost, citlivost, dobrá reprodukovatelnost
- Přijatelná rychlost

## Ne až tak pozitivní aspekty:

- Flexibilita

## K vážnému zamyšlení:

- Jak vysokou citlivost skutečně potřebujeme?
- Jsme schopni výsledky biologicky a klinicky interpretovat?
- Finanční aspekty (Illumina → monopol na trhu)
- Etické otázky (zejména WES vs. predispozice)

# Literatura

- Stratton 2009: The cancer genome.
- Simon 2013: Implementing personalized cancer genomics in clinical trials.
- Wang 2009: RNA-Seq: a revolutionary tool for transcriptomics
- Meaburn 2012: Next generation sequencing in epigenetics: insights and challenges.
- Mundade 2014: Role of ChIP-seq in the discovery of transcription factor binding sites, differential gene regulation mechanism, epigenetic marks and beyond.
- Padmanabhan 2013: Genomics and metagenomics in medical microbiology.
- Ellengren 2012: The genomic landscape of species divergence in Ficedula flycatchers.
- Kapgate 2015: Next generation sequencing technologies: Tool to study avian virus diversity.



- Fridman 2012: Next-generation education in crop genetics.
- Yang 2014: Application of next-generation sequencing technology in forensic science.
- Dong 2012: Exploring the cancer genome in the era of next-generation sequencing.
- Walsh 2010: Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing.
- Koubkova 2014: Sekvenování nové generace a možnosti jeho využití v onkologické praxi
- Guan 2012: Application of next-generation sequencing in clinical oncology to advance personalized treatment of cancer.
- Ley 2008: DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome.

- Ding 2012: Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing.
- Jardin 2014: Next generation sequencing and the management of diffuse large B-cell lymphoma: from whole exome analysis to targeted therapy.
- Wang 2011: *SF3B1* and Other Novel Cancer Genes in Chronic Lymphocytic Leukemia
- Landau 2013: Evolution and Impact of Subclonal Mutations in Chronic Lymphocytic Leukemia
- Rossi 2014: Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia.

DÍKY ZA POZORNOST!

[m.trbusek@volny.cz](mailto:m.trbusek@volny.cz)