



SeqOne platform

Návod k použití

V platnosti od SeqOne Platform v2.1

1. Úvod	7
2. Kontakt na výrobce a švýcarského zástupce	8
3. Dozor nad zdravotnických prostředků	8
4. Zamýšlené použití platformy	8
4.1. Zamýšlené použití	8
4.2. Indikace k použití	9
4.3. Popis	9
4.4. Populace pacientů	10
4.5. Zamýšlení uživatelé	10
4.6. Charakteristika	10
4.6.1. Analytická výkonnost	12
4.6.2. Bioinformatické moduly	13
4.6.3. Anotační moduly	21
4.6.4. Klinická výkonnost	23
5. Bezpečnostní opatření	24
5.1. Vhodné použití	24
5.2. Upozornění	24
5.3. Omezení	25
5.4. Poruchy	26
6. Minimální požadavky	26
6.1. Hardware	26
6.2. Síť	26
7. Ověřování	27
7.1. Přihlášení	27
7.2. Odhlášení	28
7.3. Domovská stránka	28
7.4. Správa uživatelských účtů	29
7.4.1. Správa "My Account"	30
7.4.2. Správa uživatelských nastavení - definice	30
Entity	30
Users	30
Worksets	30
Manifest	30
Panels - Transcript	30
Panels - Genes	31
Marker	31
Protocols	31
Reporting - Global settings	31
Reporting - Template list	32

Tags	32
API Keys	32
7.4.3. Správa administrátora - rozhraní	33
7.4.4. Přepínání mezi režimy Main a Preview	36
8. Popis operačních kroků	37
9. Vytvoření projektu	38
9.1. Vytvoření projektu	38
9.2. Řízení projektů	39
9.3. Delete project	39
9.4. Data DNA (sekvenování DNA, sekvenování ctDNA)	39
9.5. Data RNA (sekvenování RNA)	41
9.6. Nastavení projektu	42
9.6.1. Konfigurace výchozích metrik QC projektu	42
9.6.2. Nastavení výchozího profilu filtru projektu	42
10. Nahrání vzorku	43
10.1. Nahrávání vzorků	43
Z počítače	43
Přes SDS (SeqOne Data Sync)	43
Stav nahrávání vzorků	43
Kapacita nahrávání	44
10.2. Formát vstupního souboru	44
10.2.1. Soubory primární analýzy	44
10.2.2. Soubory sekundární analýzy	45
10.3. Metadata vzorku	47
10.4. Upravit vzorek	48
10.5. Odstranění vzorku	48
11. Analýza	49
11.1. Spuštění nové analýzy	49
11.1.1. Výběr pracovního setu pro analýzu	50
Worksets Analysis dostupné ze souborů FASTQ	50
GermlineVar	51
GermVar	51
GermlineVarLR	51
GermlineFamily	52
GermVar Family	52
CNVCapture	53
CNVExome	53
SomaVar	53
SomaCNVCapture	54
SomaRNA	54

SomaMSI	54
SomaHRD	55
SomaVar LF	55
SomaCGP	56
SomaHemato	56
SomaLBx	57
SomaMethyl	57
Soubory analýz dostupné ze souborů sekundárních analýz	58
GermVar Tertiary	58
GermVar Tertiary Family	58
SomaVar Tertiary	59
11.1.2. Volba režimu zpracování UMI	60
Prezentace režimů zpracování UMI	60
Proces ošetření UMI	61
Kompatibilita s analytickými worksety	61
11.2. Stav analýzy	62
11.3. Kvalita vzorku	62
12. Kontrola kvality dat	64
12.1. Metriky a zpráva o kontrole kvality	64
12.1.1. Metriky a kontroly kvality	64
12.1.2. Zpráva QC	64
12.1.3. Zpráva ID-Check	67
12.1.4. Technická zpráva	67
12.1.5. Zpráva UPD	69
Skóre UPD	69
Graf frekvence B-alely (BAF) ze zprávy UPD	70
13. Interpretace výsledků analýzy	71
13.1. Výsledky analýzy	71
13.1.1. Správa pracovního postupu interpretace analýzy	71
13.1.1.1. Zadání analýz pro interpretaci	71
13.1.1.2. Analýza uzamčení	72
13.1.1.3. Boční panel pro správu analýzy	73
Modální zobrazení Workflow	74
Modální zobrazení Patient	75
Modální zobrazení HPO	76
Modální zobrazení Panel	78
13.1.1.4. Boční panel pro správu hodnocení	79
13.1.2. Přehled výsledků analýzy na kartě Dashboard	80
13.1.3. Podrobné pokrytí podle exonů	81
Metodika	81

Zobrazení a výklad	82
13.1.4. Malé varianty	85
Metodika	85
Vizualizace a interpretace	89
Třídění variant	103
Konfigurace sloupců tabulky Variant	104
Úprava metadat vzorku z panelu analýzy	105
Varianty filtrování	107
Filtiry Genes & Variants	107
Sekce Samples & Quality	107
Sekce Frequencies	109
Sekce Genetics	110
Sekce Databases	111
Sekce Transmission & Phenotypes	112
Sekce Classification	115
Sekce Scores	116
Vyberte indikaci	118
Použití předdefinovaného profilu filtru	119
Úprava a ukládání profilů filtrů	119
Oblíbené profily filtrů	119
Kombinace profilů filtrů	120
Profily filtrů SeqOne	120
Označit jako zobrazenou funkci	121
Výběr variant	121
Hodnocení variant	121
Stránka variant	124
Vytvoření klinické zprávy	132
13.1.5. CNV	133
13.1.5.1. CNV v panelech genů	133
Metodika	133
Zobrazení a výklad	134
13.1.5.2. CNV v analýze SomaCGP a SomaLBx	139
Metodika	139
Vizualizace a interpretace	139
13.1.5.3. CNV v exomech a genomech	140
Metodika pro exome	140
Metodology pro genom	141
Zobrazení a výklad	141
Tabulka výsledků CNV	141
Podrobná stránka CNV	144

13.1.6. Prohlížeč velkých variant pro CNVs a SVs (LVV)	149
Tabulka výsledků CNVs & SVs	149
Třídění, filtrování, prohlížení a interpretace	151
Postranní panel CNV/SV (Details Drawer)	151
13.1.7. Znaký v analýze SomaCGP	153
Metodika výpočtu MSI skóre	153
Metodika výpočtu skóre TMB	153
Vizualizace a interpretace	154
13.1.8. Krátké tandemové repetice (STRs)	156
Metodika	156
Zobrazení a výklad	156
13.1.9. Fúzní geny	158
13.1.9.1. Fúze zjištěné z DNA	158
Metodika	158
Zobrazení a výklad	158
13.1.9.2. Fúze zjištěné z RNA	160
Metodika	160
Zobrazení a výklad	161
13.1.10. Klonálnost	164
13.1.11. Varianty spojování	165
Metodika	165
Vizualizace a interpretace	165
13.1.12. Stav MSI/MSS	167
Metodika	167
Zobrazení a výklad	167
13.1.13. Stav HRD/HRP	169
Metodika	169
Vizualizace a interpretace	171
13.1.14. Metylační data pro SomaMethyl	176
14. Zpráva	179
14.1.1. Výstupní soubory	179
15. Řešení problémů a chybové zprávy	180
16. Záruka	182
17. Právní upozornění	182
17.1. Podmínky použití	182
17.2. Omezení odpovědnosti	182
17.3. Autorská práva a ochranné známky	182

1. Úvod

Před použitím softwaru doporučujeme pečlivě si přečíst pokyny. Tištěnou verzi těchto pokynů vám na požádání poskytne tým podpory společnosti SeqOne (support@seqone.com). Abyste při používání platformy dosáhli co nejlepších výsledků, pečlivě si přečtěte všechna bezpečnostní

opatření označená symbolem . Různé regulační symboly a další symboly používané v tomto návodu a softwaru jsou následující:

Symbol	Význam
	Výrobce
	Označuje zplnomocněného zástupce usazeného ve Švýcarsku.
	Datum výroby (datum vydání softwaru)
	Diagnostický zdravotnický prostředek in vitro.
	Jedinečný identifikátor zařízení
	Odkaz na katalog výrobce
	Splňuje požadavky evropského nařízení (EU) 2017/746 Označení CE s identifikačním číslem oznámeného subjektu
	Přečtěte si návod k použití
	Varování Tento bezpečnostní symbol označuje, že při používání přístroje je třeba dbát zvláštní opatrnosti.

Štítek (včetně informací o verzi softwaru SeqOne a kontaktních údajů) a tento dokument popisující návod k použití (v různých jazycích) jsou k dispozici v sekci "About".

2. Kontakt na výrobce a švýcarského zástupce



SeqOne S.A.S

Web: <https://seqone.com/>

22 rue Durand
34000 Montpellier
FRANCIE

Podpora uživatelů :

support@seqone.com



Arazy Group Swiss GmbH

Bruderholzallee 53

4059 Basilej, Švýcarsko

swiss.ar@arazygroup.com

3. Dozor nad zdravotnických prostředků

Pokud při používání prostředku dojde nebo může dojít k závažnému incidentu, žádáme uživatele, aby incident bez zbytečného odkladu nahlásili výrobci a příslušnému orgánu v zemi, kde k incidentu došlo nebo může dojít.

Platné předpisy a seznam příslušných orgánů v EU jsou k dispozici na této adrese:

https://health.ec.europa.eu/medical-devices-sector/new-regulations/contacts_en



4. Zamýšlené použití platformy

4.1. Zamýšlené použití

SeqOne platform je softwarový diagnostický lékařský přístroj *in vitro* určený k podpoře zdravotnických pracovníků při analýze a interpretaci nezpracovaných genomických a sekvenačních dat od pacientů, kteří podstupují analýzu NGS nebo CGH array. Platforma poskytuje kvalitativní a prakticky využitelné poznatky a funguje jako systém pro podporu rozhodování, který pomáhá zdravotnickým pracovníkům při klinických rozhodnutích.

4.2. Indikace k použití

SeqOne Platform podporuje sekundární a terciární analýzu ze surových sekvenačních dat i samotnou terciární analýzu. Její analytické schopnosti nezávislé na patologii mají uplatnění v různých lékařských oborech, včetně onkologie a vzácných onemocnění.

4.3. Popis

SeqOne platform je softwarový diagnostický lékařský přístroj *in vitro*, který integruje sadu aplikací pracujících ve vzájemné součinnosti a poskytuje odborným uživatelům výkonnou sadu nástrojů pro interpretaci komplexních genomických a sekvenačních dat.

SeqOne platform integruje zavedené informace z celé řady **externích zdrojů**, včetně veřejných databází, vědecké literatury a vlastních znalostních databází. Tyto externí zdroje dat poskytují identifikovaným genomickým variantám kontext a význam.

Vstupní soubory pro různé aplikace v rámci platformy jsou zpracovávány pomocí přizpůsobené kombinace bioinformatických a anotací modulů, což zajišťuje, že každá analýza je optimalizována pro své specifické požadavky.

Bioinformatické moduly provádějí sekundární analýzu pomocí výpočetních metod, jako je zarovnávání a volání variant, k identifikaci genetických variant v datech sekvenování DNA a RNA porovnáním s referenčním genomem (hg19 nebo hg38). Data, která lze použít pro zárodečnou i somatickou analýzu, lze získat ze sekvenování celého genomu, sekvenování exomu nebo cílených genových panelů.

Anotační moduly provádějí terciární analýzu, která anotuje varianty s podrobnostmi, jako je gen a jeho známá funkce, umístění varianty v genu a publikace nebo klinické zprávy, které se o genu nebo variantě zmiňují. Tato anotace určí, která klasifikační pravidla ACMG a AMP/CAP/ASCO se vztahují na jednotlivé varianty, a na tomto základě lze dospět k verdiktu o patogenitě. Anotace je rovněž prováděna prostřednictvím DiagAI – sady nástrojů využívajících umělou inteligenci, které pomáhají analyzovat genetické varianty a rychle identifikovat potenciálně kauzální varianty.

Díky této modulární architektuře je SeqOne vysoce přizpůsobitelný různým typům analýz a výzkumných dotazů. Vstupní soubory pro různé aplikace v rámci platformy jsou zpracovávány pomocí na míru vytvořené kombinace těchto bioinformatických a anotačních modulů, což zajišťuje, že každá analýza je optimalizována pro své specifické požadavky. Integrací těchto různorodých zdrojů dat a analytických nástrojů umožňuje SeqOne odborným uživatelům získat hluboký vhled do komplexního prostředí genomických dat a jejich důsledků pro lidské zdraví a choroby.

Platforma funguje jako cloudové řešení typu SaaS (software jako služba) a všechny její funkce jsou dostupné prostřednictvím webového prohlížeče. Její součástí je SeqOne Flow, volitelný nástroj, který lze nasadit buď v cloudu AWS, nebo na serverech zákazníka, aby se usnadnil přenos dat a integrace pracovních postupů.

V rámci zabezpečení jsou všechna data v klidovém stavu šifrována pomocí šifrování AES-256 a provoz mezi uživateli a platformou je zabezpečen pomocí certifikátů SSL.

4.4. Populace pacientů

Řešení je vhodné pro každého pacienta vyžadujícího genomickou analýzu. Cílovými pacienty mohou být osoby s rakovinou, vzácnými chorobami, dědičnými nemocemi, genetickou predispozicí k rakovině nebo zdravé osoby.

4.5. Zamýšlení uživatele

SeqOne Platform je určena pro zdravotnické pracovníky odpovědné za rutinní klinickou genomickou analýzu. To zahrnuje analýzu a interpretaci surových souborů generovaných sekvenátorem ze sekvenovaných biologických vzorků. Cílovými uživateli platformy jsou genetici, bioinformatici, biologové, inženýři a technici platformy NGS.

4.6. Charakteristika

Každá aplikace (workset) se skládá z několika **bioinformatických a anotací modulů**, z nichž každý je zodpovědný za konkrétní analytické úkoly. V této kapitole jsou výkony hodnoceny na úrovni modulů, aby bylo zajištěno podrobné posouzení přesnosti a spolehlivosti každé složky.

Analytické látky detekované SeqOne Platform spadají do dvou kategorií genomových změn:

- Genetické varianty, jako jsou SNV, indely, MNV, CNV, UPD, genové fúze, strukturální varianty, splicing varianty, Alu inserce a FLT3-ITD (Internal Tandem Duplication).
- Signatures nebo komplexní biomarkery, jako je stav IGHV, nestabilita mikrosatelitových míst (MSI), Tumor Mutational Burden (TMB), metylace nebo defekty opravy dvouřetězcových zlomů prostřednictvím homologní rekombinace (HRD, Homologous Recombination Deficiency).

Bioinformatický modul	Workset	Typ analyzované látky
Detekce CNV Exom Germline	CNVExome, GermVar, GermVar Family	Genetická varianta
Detekce CNV Panel Germline	CNVCapture	Genetická varianta
SNV/Indel Germline	GermlineVar, GermlineFamily	Genetická varianta
SV Germline	GermlineVar, GermlineFamily	Genetická varianta
UPD	GermVar, GermVar Family, GermVar Tertiary, GermVar Tertiary Family	Genetická varianta
Alu inserce	GermlineVar, GermlineFamily, SomaVar, SomaVar LF, SomaCGP, SomaHemato, GermVar, GermVar Family	Genetická varianta
MNV	GermlineVar, GermlineFamily, GermVar, GermVar Family, GermVar Tertiary, GermVar Family Tertiary, SomaVar, SomaVar LF, SomaCGP, SomaHemato, SomaLBx, SomaVar Tertiary	Genetická varianta
SNV/Indel Somatická	SomaVar, SomaHemato, SomaRNA, SomaCGP	Genetická varianta
SNV/Indel somatická LF	SomaVar LF	Genetická varianta
CNV Somatické velké panely	SomaCGP	Genetická varianta
Index metylace	SomaMethyl	Signatura
Fúze somatické DNA	SomaCGP, SomaVar, SomaVar LF, SomaHemato	Genetická varianta
Somatic LBx DNA fúze	SomaLBx	Genetická varianta
TMB	SomaCGP	Signatura
MSI	SomaCGP, SomaMSI	Signature
GI pravděpodobnost (GiP)	SomaHRD	Signature
Sentieon SNV/Indel Germline	GermVar, GermVar Family	Genetická varianta
Mosaické varianty Germline	GermVar, GermVar Family	Genetická varianta
CNV WGS	GermVar, GermVar Family	Genetická varianta
Sentieon SNV/Indel Somatická	SomaLBX	Genetická varianta
Fúzní RNA	SomaRNA	Genetická varianta
Splicingová RNA	SomaRNA	Genetická varianta
CNV Malý somatický panel	SomaCNVCapture	Genetická varianta
Stav IGHV	SomaHemato	Signatura
Podskupiny IGHV (subsets)	SomaHemato	Signatura
CFH	GermlineVarLR	Genetická varianta
FLT3	SomaVar, SomaHemato	Genetická varianta

4.6.1. Analytická výkonnost

Výkonnost uvedená v tomto návodu k použití odpovídá nejnovějším verzím worksetů pro diagnostiku in vitro, které jsou uvedeny v následující tabulce:

Název worksetu	Verze
CNVCapture	2.6
CNVExome	2.2
GermlineVar	2.6
GermlineFamily	2.6
GermlineVar LR	1.2
GermVar	3.12
GermVar Family	3.12
GermVar Tertiary	1.16
GermVar Tertiary Family	1.14
SomaCGP	1.6
SomaCNVCapture	2.4
SomaHemato	0.6
SomaHRD	1.5
SomaLBx	0.7
SomaMethyl	0.3
SomaMSI	1.1
SomaRNA	3.3
SomaVar	2.9
SomaVar LF	2.1
SomaVar Tertiary	2.4

4.6.2. Bioinformatické moduly

- **Detekce CNV Exome Germline a Panel Germline**

Typ varianty	Verze genomu	Pozitivní kontroly	Skutečné pozitivní výsledky	Falešně negativní	Kritéria přijatelnosti	Citlivost
Detekce CNV Exom Germline (GermlineVar)	GRCh37	8	8	0	100%	100%
Detekce CNV Exom Germline (GermVar)	GRCh37	8	8	0	100%	100%
Detekce CNV Panel Germline	GRCh37	15	15	0	100%	100%
Detekce CNV Panel Germline	GRCh38	11	11	0	100%	100%

- **SNV / Indel Germline – údaje z panelu**

Verze genomu	Typ varianty	Kritéria přijatelnosti	Citlivost před filtrováním	Citlivost po filtrování
GRCh37	SNV	100 %	98 %	100 %
GRCh38	SNV	100 %	98 %	100 %

Verze genomu	Typ varianty	Kritéria přijatelnosti	Přesnost před filtrem	Přesnost po filtrování
GRCh37	SNV	100 %	90 %	100 %
GRCh38	SNV	100 %	90 %	100 %

- **SNV / Indel Germline – data exomu**

Verze genomu	Type of variant	Kritéria přijetí*	Citlivost před filtrováním	Citlivost po filtrování
GRCh37	SNV	> 98 %	98,7	98,5 %
	Indel	> 80 %	88,0 %	87,5 %
GRCh38	SNV	> 98 %	98,5 %	98,3 %
	Indel	> 80 %	84,9 %	84,3

* pro detekci variant nacházejících se v oblastech s vysokou spolehlivostí u NA12878 a NA24385.

Verze genomu	Type of variant	Kritéria přijatelnosti*	Přesnost před filtrem	Přesnost po filtrování
GRCh37	SNV	> 90 %	92,1 %	94,4 %
	Indel	> 80 %	84,7 %	89,0 %
GRCh38	SNV	> 90 %	94,2 %	95,4 %
	Indel	> 80 %	82,2 %	87,6 %

* pro detekci variant nacházejících se v oblastech s vysokou spolehlivostí NA12878 a NA24385.

- **SNV / Indel Germline - Opakovatelnost a Reprodukovatelnost**

- **Opakovatelnost**

Typ varianty	Citlivost	Přesnost
SNV	98,59 % ± 0 %	95,92 % ± 0 %
indel	87,75 % ± 0 %	81,63 % ± 0 %
"All"	98,38 % ± 0 %	95,63 ± 0 %

- **Reprodukovatelnost**

Typ varianty	Citlivost	Přesnost
SNV	98,57 % ± 0,03 %	95,93 % ± 0,08 %
indel	86,58 % ± 0,42 %	78,97 % ± 0,96 %
celkem	98,34 % ± 0,03 %	95,58 ± 0,09 %

- **Mozaikové varianty Germline**

- **Citlivost SNV**

Verze genomu	VAF biny	Počet skutečných variant	Kritéria přijatelnosti	Citlivost 100x
GRCh38	[0,05; 0,1[	3704	40 %	47,6 %
GRCh38	[0,1; 0,2[	747	70 %	73,1 %
GRCh38	[0,2; 0,3]	685	90 %	90,7 %
GRCh38	[0,3; 1,0[	76	90 %	93,4 %

- **Citlivost Indelů**

Verze genomu	VAF biny	Počet skutečných variant	Kritéria přijatelnosti	Citlivost 100x
GRCh38	[0,05; 0,1[	87	40 %	43,7 %
GRCh38	[0,1; 0,2]	14	70 %	78,6 %
GRCh38	[0,2; 0,3]	15	90 %	100 %
GRCh38	[0,3; 1,0[	0	'N/A'	'N/A'

- **SV Germline**

Typ varianty	Verze genomu	Pozitivní kontroly	Skutečné pozitivní výsledky	Falešně negativní	Kritéria přijatelnosti	Citlivost
SV Germline	GRCh37	2	2	0	100 %	100 %

- **Alu inserce**

Verze genomu	Pozitivní kontroly	Skutečné pozitivní výsledky	Falešně negativní	Kritéria přijatelnosti	Citlivost
GRCh37	1	1	0	100 %	100 %
GRCh38	1	1	0	100 %	100 %

- **MNV**

Verze genomu	Typ varianty	Kritéria přijetí	Přesnost exomu	Přesnost genomu
GRCh37	MNV	> 85 %	87,0 %	86,9 %
GRCh38	MNV	> 85 %	93,0 %	91,5 %

- **SNV / Indel somatický**

- **Citlivost**

Verze genomu	Typ varianty	Pozitivní kontroly	Skutečné pozitivní výsledky	Falešně negativní	Kritéria přijatelnosti*	Citlivost
GRCh37	SNV (≥ 5 %)	121	121	0	100 %	100 %
	Indel (≥ 5 %)	21	21	0	100 %	100 %
GRCh38	SNV (≥ 5 %)	121	121	0	100 %	100 %
	Indel (≥ 5 %)	21	21	0	100 %	100 %

*Kritérium přijatelnosti pro detekci variant je definováno jako 100% citlivost pro SNV a indely s očekávanou VAF nad 5 %.

○ **Správnost**

			Přesnost			
			GRCh37		GRCh38	
Gene	Variant	Očekávaná frekvence	Průměr frekvence*	Odchylka	Průměrná frekvence*	Odchylka
KIT	p.D816V	10 %	9,07 % +/- 0,79 %	0,97 %	9,07 % +/- 0,79 %	0,97 %
PIK3CA	p.E545K	8,80 %	7,88 % +/- 0,84 %	1,04 %	7,87 % +/- 0,83 %	1,05 %
EGFR	p.E746_A750del	1,90 %	1,53 % +/- 0,5 %	0,40 %	1,53 % +/- 0,5 %	0,40 %
KRAS	p.G12D	6,30 %	6,27 % +/- 0,76 %	0,57 %	6,27 % +/- 0,76 %	0,57 %
KRAS	p.G13D	15 %	14,59 % +/- 0,86 %	0,69 %	14,59 % +/- 0,86 %	0,69 %
EGFR	p.G719S	24,50 %	24,5 % +/- 1,39 %	1,15 %	24,5 % +/- 1,39 %	1,15 %
PIK3CA	p.H1047R	17,50 %	17,47 % +/- 1,04 %	0,77 %	17,47 % +/- 1,04 %	0,77 %
EGFR	p.L858R	2,80 %	3,58 % +/- 0,51 %	0,83 %	3,58 % +/- 0,51 %	0,83 %
NRAS	p.Q61K	12,50 %	11,27 % +/- 1,19 %	1,27 %	11,27 % +/- 1,19 %	1,27 %
EGFR	p.T790M	0,90 %	1,33 % +/- 0,15 %	0,43 %	1,33 % +/- 0,15 %	0,58 %
BRAF	p.V600E	10,70 %	11,68 % +/- 0,62 %	0,98 %	11,68 % +/- 0,62 %	0,98 %

*Průměr frekvencí je získán z 12 opakovaných měření vzorku HD200.

● **SNV/Indel somatická LF**

Genom verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Kritéria přijatelnosti*	Citlivost
GRCh37	42	42	0	100 %	100 %
GRCh38	42	42	0	100 %	100 %

*Kritérium přijatelnosti pro detekci variant je definováno jako 100% citlivost u SNV a indelů s očekávanou hodnotou VAF vyšší než 0,1 %.

● **CNV somatické velké panely**

Verze genomu	Typ varianty	Pozitivní kontroly	Skutečné pozitivní	Falešně negativní	Kritéria přijatelnosti	Citlivost
GRCh37	Zisk	27	26	1	90 %	96,3 %
	Loss	20	19	1	90 %	95,0 %
	LOH	11	11	0	90 %	100,0 %
GRCh38	Zisk	27	26	1	90 %	96,3 %
	Loss	20	19	1	90 %	95,0 %
	LOH	11	10	1	90 %	90,9 %

- Index metylace

Verze genomu	Kritéria přijatelnosti	Vážená shoda
GRCh37	>99 %	99,88

- Fúze somatické DNA

Genom verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Přijetí Kritéria	Citlivost
GRCh37	9	9	0	100 %	100 %
GRCh38	9	9	0	100 %	100 %

- Somatic LBx DNA fúze

Verze genomu	Frakce nádoru	Pozitivní kontroly	Skutečně pozitivní	Falešně negativní	Kritéria přijatelnosti	Citlivost
GRCh38	1 %	22	21	1	85 %	95,5 %
	0,5 %	22	19	3	85 %	86,4 %
	0,1 %	22	21	1	85 %	95,5 %

- TMB

Verze genomu	Kritéria přijatelnosti	Citlivost	Přesnost
GRCh37	>80 %	83 %	85 %

- MSI

Verze genomu	Pozitivní kontroly	Negativní kontroly	Skutečně pozitivní	Falešně negativní	Kritéria přijatelnosti	Citlivost	Přesnost
GRCh37	22	24	22	0	100 %	100 %	100 %
GRCh38	6	17	6	0	100 %	100 %	100 %

- GiP

Genom verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Přijetí Kritéria	Citlivost
GRCh37	171	165	6	90 %	96,5 %

Genom verze	Pozitivní kontroly	Skutečně falešně	Falešně pozitivní	Přijetí Kritéria	Přesnost
GRCh37	171	165	7	90 %	95,9 %

- Sentieon SNV / Indel Germline

Verze genomu	Type of variant	Kritéria přijatelnosti*	Citlivost Exome	Citlivost Genom
GRCh37	SNV	> 90 %	98,9 %	98,5 %
	Indel	> 80 %	96,5 %	98,5 %
GRCh38	SNV	> 90 %	98,5 %	99,3 %
	Indel	> 80 %	95,5 %	99,1 %

* pro detekci variant nacházejících se v oblastech s vysokou spolehlivostí u NA12878 a NA24385

Verze genomu	Type of variant	Kritéria přijatelnosti*	Přesnost Exome	Přesnost genomu
GRCh37	SNV	> 90 %	98,7 %	98,5 %
	Indel	> 80 %	96,5 %	98,6 %
GRCh38	SNV	> 90 %	99,6 %	99,6 %
	Indel	> 80 %	97,2 %	99,5 %

* pro detekci variant nacházejících se v oblastech s vysokou spolehlivostí u NA12878 a NA24385.

- CNV WGS

Verze genomu	Pozitivní kontroly	Skutečně pozitivní	Falešně pozitivní	Kritéria přijatelnosti	Citlivost
GRCh38	Duplikace	1 kb – 5 kb	8	10 %	37,5 %
	Duplikace	≥ 5 kb	3	10 %	100 %
	Delece	1 kb – 5 kb	321	80 %	61,7 %
	Delece	≥ 5 kb	132	80 %	92,4 %

Ačkoli citlivost na malé delece (1–5 kb) nedosahuje akceptačních kritérií, je toto snížení přijatelným kompromisem za účelem maximalizace detekce významnějších velkých událostí a duplikací.

Verze genomu	Pozitivní kontroly	Skutečné pozitivy	Falešně pozitivní	Kritéria přijatelnosti	Přesnost
GRCh38	Duplikace	1 kb – 5 kb	8	"NA"	30
	Duplikace	≥ 5 kb	3	"NA"	75 %
	Delece	1 kb – 5 kb	321	"NA"	94,3 %
	Delece	≥ 5 kb	132	"NA"	92,4 %

- **Sentieon SNV/Indel somatický**

Verze genomu	Typ varianty	Pozitivní kontroly	Skutečně pozitivní	Falešně pozitivní	Kritéria přijatelnosti	Citlivost
GRCh38	SNV	39	37	2	85 %	95 %
	Indel	21	16	5	75 %	76 %

- **Fúzní RNA**

Genom Verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Přijetí Kritéria	Citlivost
GRCh37	20	20	0	100 %	100 %
GRCh38	20	19	1	95 %	95 %

- **Splicing RNA**

Genom Verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Přijetí Kritéria	Citlivost
GRCh37	3	3	0	100 %	100 %
GRCh38	3	3	0	100 %	100 %

- **CNV Malý panel somatických**

Genom Verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Přijetí Kritéria	Citlivost
GRCh37	4	4	0	100 %	100 %
GRCh38	4	4	0	100 %	100 %

- **Status IGHV IMGT**

Genom verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Přijetí Kritéria	Citlivost
GRCh37	122	116	6	95 %	95,9 %

- **Status IGHV igblast**

Verze genomu	Pozitivní kontroly	Skutečně pozitivní	Falešně negativní	Kritéria přijatelnosti	Citlivost
GRCh37	122	117	5	95 %	95,1 %

- **Podskupiny IGHV**

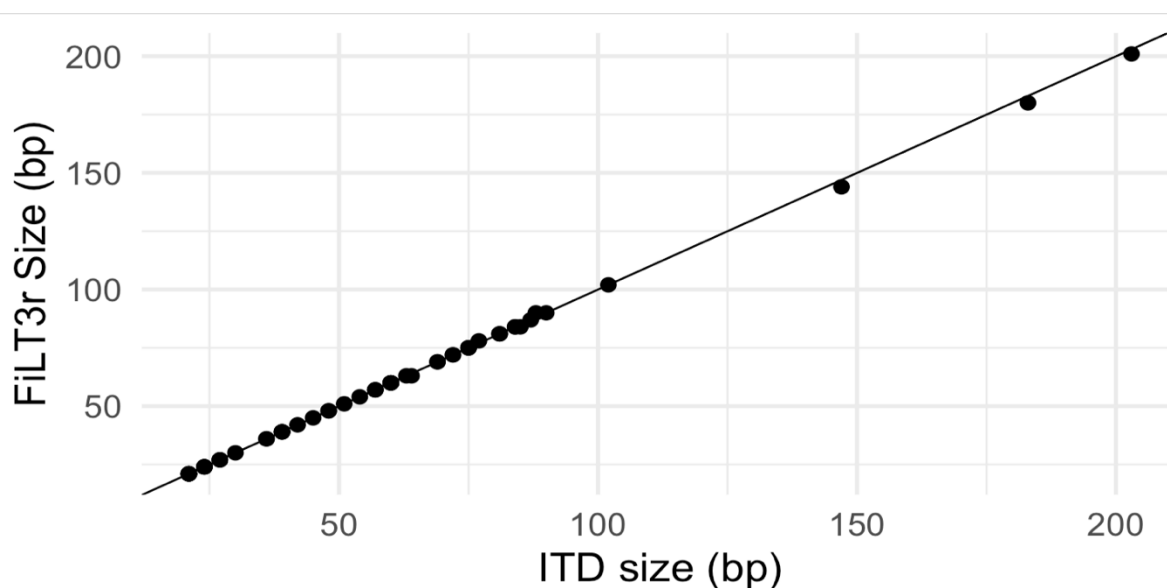
Genom verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Přijetí Kritéria	Citlivost
GRCh37	61	10	0	95 %	100 %

Genom verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Přijetí Kritéria	Přesnost
GRCh37	61	10	0	95 %	100 %

- CFH

Genom verze	Pozitivní kontroly	Skutečně falešně	Falešně negativy	Přijetí Kritéria	Citlivost
GRCh38	4	4	0	100 %	100 %

- FLT3



Pozorovaná vs. teoretická velikost události FLT3-ITD. Pearsonův koeficient $R^2 > 0,99$

Výsledky ukazují 100% citlivost při detekci událostí FLT3-ITD a téměř dokonalou korelaci mezi očekávanými a pozorovanými velikostmi událostí z analýzy.

- UPD

- Citlivost

Type of analysis	Verze genomu	Pozitivní kontroly	Negativní kontroly	Skutečné pozitivní výsledky	Falešně negativní	Kritéria přijatelnosti	Citlivost
Singleton	GRCh37/38	7	23	7	0	80 %	100 %
Trio	GRCh37/38	7	2	6	1	80 %	86 %

- Přesnost

Type of analysis	Verze genomu	Pozitivní kontroly	Negativní kontroly	Skutečné pozitivní výsledky	Falešně pozitivní	Kritéria přijatelnosti	Přesnost
Singleton	GRCh37/38	7	23	7	0	80 %	100 %
Trio	GRCh37/38	7	2	6	0	80 %	100 %

4.6.3. Anotační moduly

- "DiagAI Score"

Datový soubor	Analýza typu	Metrické vyhodnoceno	Kritéria Kritéria	Metrika hodnota
Exome pozitivní na vzácné onemocnění	Zárodečná linie	Střední pořadí	= 1	1
Exome pozitivní na vzácná onemocnění	Zárodečná linie	top_10	> 90 %	95 %
Exome CNV vzácných onemocnění	Zárodečná linie	Střední pořadí	= 1	1
Exome s pozitivním nálezem CNV u vzácných onemocnění	Zárodečná linie	top_1	> 80 %	83,3 %

- Somatické skóre DiagAI

Datový soubor	Analýza typu	Metrika vyhodnoceno	Kritéria přijatelnosti	Hodnota metriky
Pozitivní na rakovinu TSO500	"Somatic"	top_10	> 90 %	100 %

- **DiagAI užší výběr**

Datový soubor	Analýza typu	Metrika hodnoceno	Přijetí Kritéria	Hodnota metriky
Exome pozitivní na vzácné onemocnění	"Germline"	citlivost	> 90	95
Exome pozitivní na vzácné onemocnění	Zárodečná linie	Průměrná délka	< 25	18,475
Exome pozitivní na vzácné onemocnění	Zárodečná linie	Střední délka	< 25	17

- **DiagAI smartpicks**

Datový soubor	Analýza typu	Metrika hodnoceno	Přijetí Kritéria	Hodnota metriky
Exome pozitivní na vzácné onemocnění	"Germline"	citlivost	> 40 %	45
Exome pozitivní na vzácné onemocnění	"Germline"	přesnost	> 80 %	78 %

- **ACMG CNV**

Datový soubor	Typ analýzy	Metrika vyhodnoceno	Kritéria přijatelnosti	Hodnota metriky
Datový soubor CNV s poznámkami odborníků ACMG	"Germline"	citlivost	> 70 %	76

- **ACMG SNV**

Datový soubor	Typ analýza	Metrika vyhodnoceno	Kritéria přijatelnosti	Hodnota metriky
Datový soubor s poznámkami odborníků ACMG	"Germline"	přesnost	> 50 %	56 %

- **DiagAI UP2**

Datový soubor	Analýza typu	Metrika vyhodnoceno	Kritéria přijatelnosti	Metrika hodnota
Varianty způsobující onemocnění	Zárodečná linie	citlivost	90 %	92 %

- **DiagAI UP2 SV**

Datový soubor	Analýza typu	Metrika hodnoceno	Kritéria přijatelnosti	Metrika hodnota
Varianty způsobující onemocnění	'Germline'	citlivost	90 %	98 %

- **DiagAI Phenogenius**

Datový soubor	Analýza typu	Metrika hodnoceno	Kritéria přijatelnosti	Metrika hodnota
Spojivosti genů a symptomů	Zárodečná linie	top_100	50 %	49,8 %
Spojivosti mezi geny a symptomy	Zárodečná linie	top_200	60 %	60,7 %
Souvztažnosti genů a symptomů	Zárodečná linie	top_500	75 %	77,0 %
Spojivosti mezi geny a symptomy	Zárodečná linie	Medián v top 100	<20	15
Souvztažnosti genů a symptomů	Zárodečná linie	Průměr v top 100	<30	25

- **Compermed**

Datový soubor	Analýza typu	Metrika vyhodnoceno	Kritéria přijatelnosti	Metrika hodnota
Pozitivní na rakovinu TSO500	"Somatic"	citlivost	90 %	91 %

4.6.4. Klinická výkonnost

Platforma SeqOne není určena k posuzování konkrétního klinického stavu nebo fyziologického stavu. Ačkoli poskytuje informace o fyziologickém stavu pacienta, jako je identifikace určitých variant, není určena k posuzování konkrétních klinických stavů nebo fyziologických stavů. Platforma SeqOne proto sama o sobě nevyžaduje prokázání klinické výkonnosti.

5. Bezpečnostní opatření

5.1. Vhodné použití

(Viz část **Určení uživatelé**)

Jakékoli nevhodné použití platformy je zakázáno.

Uživatelé jsou odpovědní za kontrolu, zda je platforma vždy funkční, a za zajištění nezbytných podmínek pro její správný výkon.

5.2. Upozornění

 Poskytnuté výsledky by neměly být jediným základem pro změnu péče o pacienta. Uživatelé proto musí při používání softwaru a jeho výsledků postupovat s náležitou opatrností. Diagnózy by měly být potvrzeny různými klinickými a lékařskými vyšetřeními.

Vyšetření onemocnění pomocí NGS může odhalit sekundární nebo náhodné nálezy, jako je predispozice k rakovině, z genů nesouvisejících s primárním onemocněním (Kuhlen a Borkhardt 2015). Navzdory politickým prohlášením agentur jako AMP a ACMG (Green et al. 2013; Allyse a Michie 2013; představenstvo ACMG 2015; Hegde et al. 2015) se národní směrnice týkající se hlášení náhodných nálezů liší. Proto je zásadní zajistit dokument „informovaného souhlasu“ (Anderson et al. 2016). Ošetřující lékař je odpovědný za to, aby pacient nebo jeho zákonný zástupce porozuměl potenciálním výsledkům a rizikům předtím, než objedná jakýkoli test NGS.

Tento produkt nezaručuje detekci všech genetických variant přítomných ve vzorku. Pravděpodobnost detekce hledané varianty je ze své podstaty závislá na kvalitě a vlastnostech vstupního vzorku (např. integrita vzorku, příprava knihovny, hloubka a rovnoměrnost sekvenování, bioinformatické filtrování a typ či kontext varianty). Negativní výsledek definitivně nevylučuje přítomnost patogenní varianty, pokud je její frekvence (VAF) pod detekčními prahy použitého analytického modulu. Spolehlivá detekce rozsáhlých genomových přestaveb (strukturální varianty, včetně CNV a delecí celých chromozomů) není zaručena. To je způsobeno inherentním omezením dat sekvenování nové generace (NGS) s krátkými čteními, což ztěžuje identifikaci těchto událostí. Jelikož čtení NGS jsou krátká (přibližně 150–200 bp), je toto omezení charakteristické pro samotnou sekvenační technologii a nejedná se o omezení SeqOne Platform. (Viz Tan R. et al., Hum Mutat. 2014 Jul;35(7):899–907. doi: 10.1002/humu.22537 a odkazy v něm uvedené). Uživatelé musí vyhodnotit vhodnost vybraného pracovního souboru a kvalitu sekvenování ve vztahu k očekávaným variantám, aby zajistili, že byl pokryt příslušný klinický kontext.

Klinická rozhodnutí založená na výsledcích screeningu novorozenců nebo Carrier screeningu musí být přijímána s mimořádnou opatrností a nikdy se nesmí spoléhat výhradně na výstupy tohoto softwaru. U screeningu novorozenců vyžaduje pozitivní nález okamžitou klinickou

korelaci se symptomy dítěte a urgentní potvrzující testy před zahájením léčby měnící život nebo dietních omezení. U carrier screeningu negativní výsledek nevylučuje riziko, že se jedná o nosiče, ale spíše jej snižuje na úroveň zbytkového rizika, protože určité genomové varianty mohou klesnout pod detekční limit softwaru. Vzhledem k tomu, že výsledky z tohoto zařízení přímo ovlivňují nevratná reprodukční rozhodnutí a kritickou neonatální péči, musí být všechny nálezy vyžadující zásah ověřeny pomocí alternativní validované metody a interpretovány kvalifikovaným odborníkem v kontextu úplné klinické anamnézy pacienta a relevantních biochemických nebo fenotypových údajů.

Nově objevené varianty, které dosud nejsou důkladně zdokumentovány ve vědecké literatuře, mohou být nesprávně klasifikovány.

Výsledky jsou přesné v době analýzy a odrážejí současné vědecké poznatky. Opakování analýzy na novější verzi platformy však může přinést odlišné výsledky.

Výsledky klasifikace z SeqOne Platform se snaží být v souladu s obecně přijímaným konsensem. Uvědomte si, že jakékoli výsledky generované po změně nastavení klasifikace jsou mimo rozsah CE-IVD a nejsou proto validovány.

Tento software obsahuje komponenty umělé inteligence (AI). Uživatelé si musí být vědomi následujících omezení, která jsou vlastní těmto nástrojům založeným na AI, a zohlednit je při interpretaci výsledků:

- Trénovací data pocházejí převážně z evropských/kavkazských populací předků vzhledem ke složení současných klinických referenčních databází;
- Výkonnost se může lišit u vzácných variant s omezeným zastoupením v trénovacích datech;
- Skóre spolehlivosti odráží nejistotu modelu, nikoli klinickou jistotu; musí být interpretována ve spojení s klinickým kontextem;
- Výkonnost se může zhoršit v důsledku datového posunu způsobeného novými sekvenčními platformami nebo protokoly přípravy knihoven, které nejsou zastoupeny v trénovacích datech;
- Výkonnost prioritizace genů nástroje DiagAI Phenogenius je hodnocena na základě asociací genů a symptomů a závisí na kvalitě a úplnosti anotace termínů HPO ze strany ošetřujícího lékaře.

5.3. Omezení

"Carrier screening" podléhá specifickým vnitrostátním právním předpisům a etickým pravidlům. Je odpovědností laboratoře a poskytovatele zdravotní péče zajistit, aby byly v rámci jejich konkrétní jurisdikce splněny všechny zákonné požadavky, včetně povinného genetického poradenství, objednávky pod vedením lékaře a informovaného souhlasu.

5.4. Poruchy

V případě poruchy:

- Okamžitě přestaňte software používat
- Zkuste identifikovat nebo odstranit příčiny podle postupu řešení problémů popsaného v této příručce (viz část **Řešení problémů a chybové zprávy**)
- Pokud se vám nepodaří identifikovat nebo odstranit příčinu pomocí této příručky, zašlete e-mail týmu podpory na adresu support@seqone.com.

6. Minimální požadavky

6.1. Hardware

Softwarové moduly platformy SeqOne jsou instalovány v režimu klient-server. Poskytovatelem služby musí být linuxový server, aby byla zajištěna kompatibilita s většinou nástrojů běžně používaných v bioinformatice. Každá položka systému je zapouzdřena do běhového prostředí, jako je kontejner Docker® , který lze nasadit na libovolnou vyhovující platformu Docker® bez ohledu na základní hardwarovou architekturu. Uživatelé musí mít k dispozici počítač s připojením k internetu a webový prohlížeč (viz níže).

Minimální systémové požadavky na klientský počítač jsou následující:

- 4 GB RAM
- Dvujádrový procesor
- Přístup k internetu
- Prohlížeč Chrome (od verze 36.0.1985) nebo Firefox (od verze 31) nebo Microsoft Edge (všechny verze založené na Chromiu).

6.2. Síť

Minimální doporučená rychlost nahrávání/stahování pro používání platformy přes internet je 100 Mb/s. Pro připojení k platformě se používá protokol HTTPS. Pro správný přístup k platformě doporučujeme požádat oddělení IT o autorizaci následujících adres :

*.eu.seqone.com

*.common.seqone.com

7. Ověřování

7.1. Přihlášení

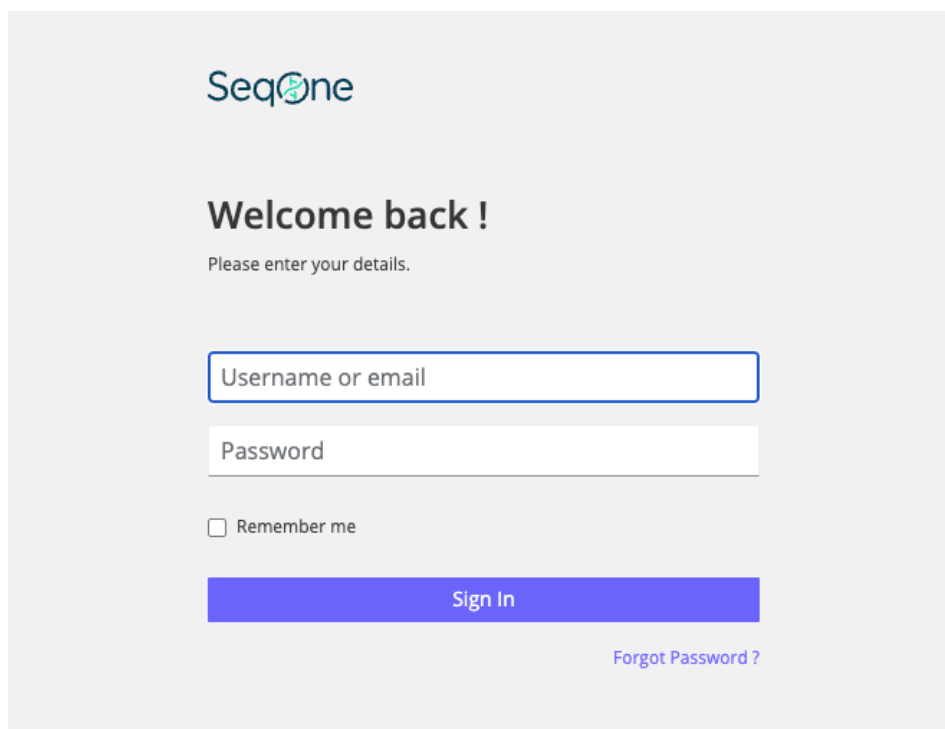
Z bezpečnostních důvodů mohou uživatelé přistupovat k platformě pouze po přihlášení na domovské obrazovce.

Domovská obrazovka je k dispozici na následující adrese URL:

- Pro Evropu a Afriku: <https://auth.eu.seqone.com/>
- Pro Švýcarsko: <https://auth.ch.seqone.com/>
- Pro Asii a Oceánii: <https://auth.au.seqone.com/>
- Pro Ameriku: <https://auth.us.seqone.com/>

Při přihlašování na domovské obrazovce musíte zadat své:

- Přihlašovací jméno nebo e-mailová adresa v poli "Username" nebo "Email".
- Heslo v poli "Password"



Obrázek 1: Přihlašovací stránka

Pokud je pro váš účet aktivováno dvoufaktorové ověřování, budete muset zadat šestimístný ověřovací kód zaslaný e-mailem na adresu přidruženou k vašemu účtu SeqOne. Tento kód je platný po dobu 30 minut.

Pokud máte přístup k několika pracovním prostorům nebo entitám, budete vyzváni k výběru entity, ke které se chcete připojit, prostřednictvím rozevírací nabídky v části "Choose a workspace".

Heslo si můžete změnit kliknutím na "Forgot Password". Podle pokynů zadejte e-mailovou adresu spojenou s vaším účtem SeqOne a obdržíte e-mail, který vám umožní obnovit heslo.

Zásady používání hesel se řídí následujícími pravidly:

- Minimální délka: 12 znaků.
- Složitost: hesla musí obsahovat alespoň jedno velké písmeno, jedno malé písmeno, jednu číslici a jeden speciální znak. Čím větší je soubor znaků, tím silnější je heslo.
- Historie hesel: žádné opakované použití předchozích hesel.
- Žádná maximální délka: aby bylo možné používat tajné fráze.

Uživatelé nemohou měnit své vlastní ID uživatele.

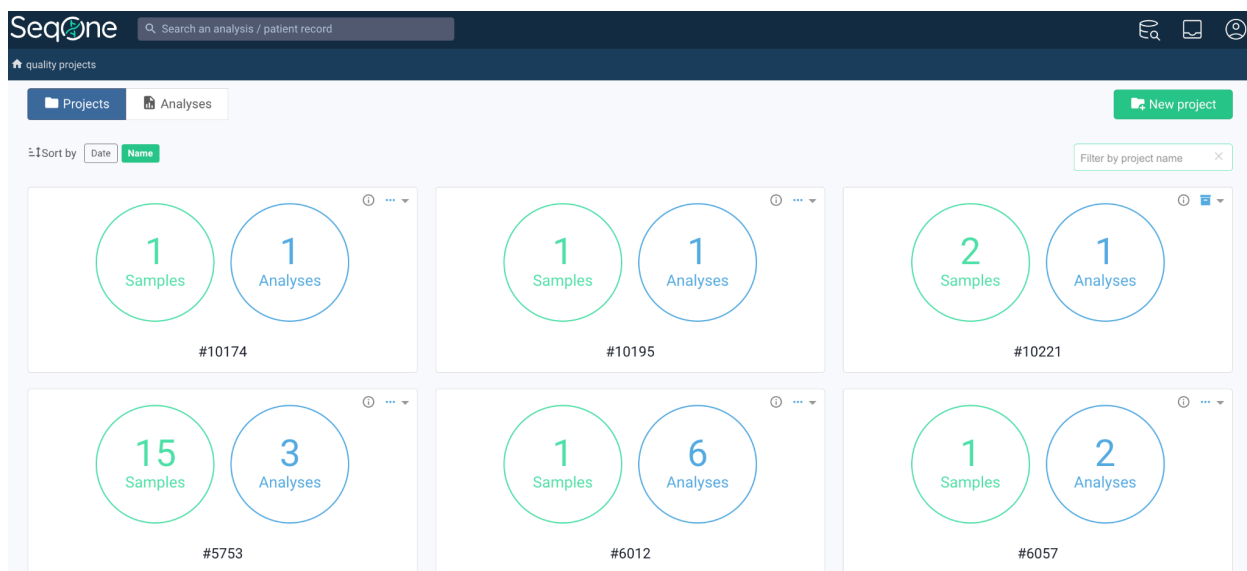
7.2. Odhlášení

Uživatelé se mohou odhlásit kliknutím na logo uživatelského účtu v pravém horním rohu stránky a poté na "Logout".

7.3. Domovská stránka

Po ověření přistupuje uživatel na stránku ovládacího panelu, která je rozdělena na dvě části:

- vyhledávací a akční banner, který je přístupný v celé navigaci zleva doprava:
 - ikonu SeqOne, abyste se mohli kdykoli vrátit na ovládací panel,
 - vyhledávací panel pro přímý přístup k dané analýze,
- VKB search icon  pro vyhledávání variant vyhodnocených v databázi VKB (viz část **Hodnocení variant**),
- ikona  , která otevře seznam úkolů uživatele ("To do list"), z něhož může přistupovat k analýzám, které mu byly přiděleny (viz oddíl **Zadání analýz pro interpretaci**),
- lištu nabídek, dostupnou prostřednictvím ikony  , která umožňuje přístup k uživatelským parametrům, panelu správce, uživatelské příručce, podmínkám používání, e-mailové adrese podpory a která umožňuje odhlášení uživatele (viz část **Správa uživatelských účtů**).
- hlavní obrazovka pro správu projektů a analýz.



Obrázek 2: Uživatelský panel

Do uživatelského panelu můžete kdykoli vstoupit kliknutím na logo .

7.4. Správa uživatelských účtů

Svůj uživatelský účet můžete spravovat kliknutím na logo v pravém horním rohu hlavního panelu.

V této části má uživatel možnost:

- přístup k nastavení účtu kliknutím na "My account"
- přístup k uživatelským parametrům kliknutím na "Entity Settings"
- změna entity, ke které je uživatel připojen, kliknutím na "Entity" a výběrem požadované entity v části "Choose your workspace",
- přístup k panelu administrátora kliknutím na "Admin", který umožňuje přístup k seznamu uživatelů s přístupem k entitě a vytváření, úpravu nebo mazání uživatelských účtů.
- kliknutím na "About" získáte přístup k uživatelské příručce v různých jazycích, informacím o verzi softwaru SeqOne a kontaktním údajům,
- přístup k podmínkám vaší licence SeqOne kliknutím na "Terms of Use",
- kliknutím na "Support" získáte přístup k e-mailové adrese podpory pro jakékoli dotazy nebo problémy zjištěné na platformě,
- přístup ke zkratce kliknutím na "Keyboard Shortcuts"
- přepínání mezi standardní verzí a verzí preview aplikace kliknutím na "SeqOne Mode" a výběrem buď "SeqOne" (standard mode) nebo "SeqOne Preview" (preview mode). Aktuálně aktivní režim je označen zaškrtnutím.
- odhlášení kliknutím na "Logout".

7.4.1. Správa "My Account"

V sekci "My account" může uživatel:

- nahrát svůj podpis a zahrnout jej do klinické zprávy,
- změnit své heslo kliknutím na "Change your password" (otevře se v nové kartě) v sekci "Manage your credentials".

7.4.2. Správa uživatelských nastavení - definice

V uživatelském nastavení může uživatel spravovat nastavení, soubory a konfigurace popsané níže.

Entity

Entitou se rozumí zařízení nebo tým, který si pořídil licenci platformy SeqOne. Platforma je navržena tak, aby ji mohlo používat několik členů v rámci jedné skupiny. V uživatelských nastaveních slouží karta *Účet* ke správě informací týkajících se subjektu.

Users

U této karty se prosím podívejte na **Správa administrátora - rozhraní**.

Worksets

Na této kartě jsou uvedeny různé worksety analýzy, které jsou k dispozici u entity, k níž je uživatel připojen. Protože na platformě SeqOne může vedle sebe existovat několik verzí určitých pracovních sad, může uživatel v této nabídce určit, která verze má být při spuštění analýzy předem vybrána.

Změny provedené v tomto seznamu jsou sledovány, takže se automaticky zobrazuje uživatel, který definoval verzi, jež se má spustit jako výchozí workset, a datum změny.

Manifest

Manifest zobrazuje genomické oblasti zájmu (ROI) jako soubor BED, který může být doplněn souvisejícími anotacemi. Nový manifest můžete nahrát prostřednictvím karty *Manifest* v uživatelských nastaveních kliknutím na "Add manifest" (Přidat manifest) a výběrem souboru manifestu, který chcete nahrát. Po nahrání bude tento manifest k dispozici všem uživatelům při vytváření nového projektu a při použití panelu *in silico* v analýze.

Panels - Transcript

Uživatelé mohou nahrát nový seznam referenčních transkriptů (RefSeq) prostřednictvím karty *Transcript*.

Panels - Genes

Tato karta umožňuje importovat nebo generovat panely genů na základě NCBI GeneID kliknutím na "Add gene list".

Očekávaným formátem souboru je tabulkový soubor (např. soubor .tsv) bez záhlaví sloupců, který je podrobně popsán na platformě. První sloupec očekávaného souboru je věnován identifikátorům genů (NCBI Entrez gene ID), druhý sloupec (nepovinný) obsahuje transkript zájmu (RefSeq ID) a třetí sloupec (nepovinný) obsahuje symbol HGNC genu. Tento soubor lze také vygenerovat vložením požadovaného seznamu genů oddělených čárkou do pole "Gene symbols" (Symboly genů).

Marker

Tato karta slouží k importu seznamu markerů genetické identity kliknutím na "Add marker config". Očekávaným formátem souboru je tabulkový soubor (.tsv nebo .txt soubor) se záhlavím obsahujícím 5 sloupců s podrobnostmi: chromozom (CHR), identifikátor varianty nebo rsID (ID), genomická pozice (POS), referenční alela (REF) a alternativní alela (ALT). Soubor nesmí obsahovat více než 100 řádků včetně záhlaví.

Protocols

Analytický protokol je předdefinovaná kombinace úkolů navržená ke standardizaci interpretace analýzy. Příslušná karta umožňuje uživatelům zobrazit seznam protokolů dostupných v entitě, jejich verze a vytvářet je.

Chcete-li vytvořit protokol, klikněte na "New protocol", poté přidejte a seřadte úkoly, které odpovídají SOP vaší laboratoře. Chcete-li upravit existující protokol, klikněte na ikonu tužky vedle něj a uložte novou verzi protokolu.

Vyberte protokol při spuštění analýzy nebo jej připojte později z karty Workflow v panelu analýzy. Označte úkoly jako dokončené přímo v panelu analýzy. Každý dokončený úkol automaticky zaznamená uživatele a datum pro plnou sledovatelnost.

Reporting - Global settings

Sekce Global settings na kartě Reporting umožňuje uživatelům povolit nebo zakázat automatické uzamčení analýzy, když se stav změní na "Reporting Validating", a přizpůsobit sekce klasifikace a reportování. Chcete-li tak učinit, v části Germline / Somatic klikněte na "New section", zadejte název sekce a vyberte tag, který chcete přidružit k požadované klasifikaci. Varianty, které odpovídají jednomu z vybraných tagů nebo klasifikací, jsou do této sekce přidány automaticky. Chcete-li tuto funkci používat, musíte nejprve vytvořit tagy na příslušné kartě.

Reporting - Template list

Karta Reporting pod "Template list" umožňuje uživatelům konfigurovat šablony používané při exportu výsledků analýzy. Na této kartě můžete přizpůsobit informace obsažené v generovaných reportech.

Tags

Karta "Tags" umožňuje uživatelům spravovat vlastní štítky pro organizaci a filtrování variant v rámci VKB a reportu. Uživatelé mohou vytvářet tagy kliknutím na "Add tag", nebo přejmenovávat a mazat tagy, které jsou dostupné všem uživatelům v rámci entity, pomocí odpovídajících ikon pera a koše.

API Keys

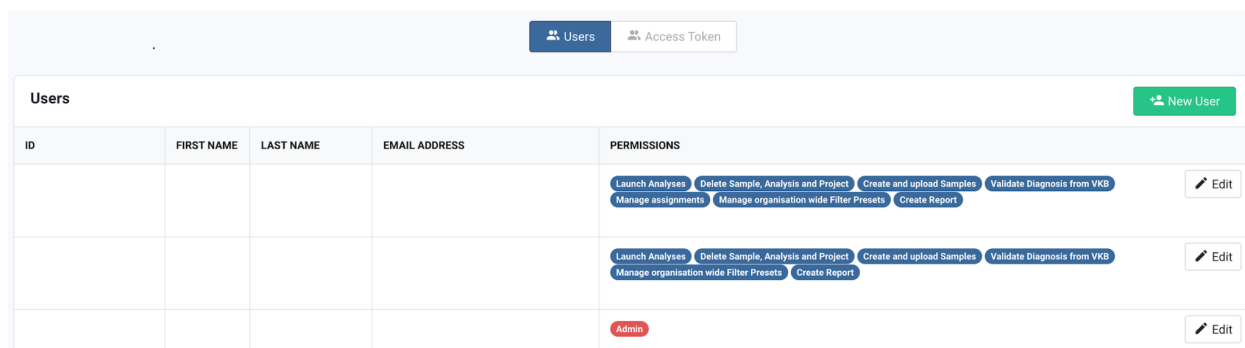
Klíč API je jedinečný autentizační token, který umožňuje externím aplikacím nebo skriptům programově komunikovat s platformou SeqOne. Na kartě API Keys si můžete prohlédnout všechny existující klíče spolu s přidruženými informacemi.

Nový klíč vytvoříte kliknutím na "+ New API Key". Prošlé klíče lze zobrazit nebo skrýt pomocí přepínače "Show expired key". Každý klíč lze odstranit jednotlivě. Klíče API musí být uchovávány v tajnosti, protože udělují přístup k platformě s příslušnými oprávněními.

7.4.3. Správa administrátora - rozhraní

Toto rozhraní je určeno správcům účtů a umožňuje jim :

- přístup k seznamu uživatelů v entitě,
- vytvářet nové uživatele,
- upravovat uživatelská práva: vytvářet projekty a nahrávat vzorky, spouštět analýzy, mazat vzorky, analýzy nebo projekty, vyhodnocovat varianty ve VKB, vytvářet zprávy, spravovat profily filtrů, spravovat přiřazení.
- aktivovat duální ověřování,
- deaktivovat uživatelský účet.



ID	FIRST NAME	LAST NAME	EMAIL ADDRESS	PERMISSIONS	
				Launch Analyses Delete Sample, Analysis and Project Create and upload Samples Validate Diagnosis from VKB Manage assignments Manage organisation wide Filter Presets Create Report	Edit
				Launch Analyses Delete Sample, Analysis and Project Create and upload Samples Validate Diagnosis from VKB Manage organisation wide Filter Presets Create Report	Edit
				Admin	Edit

Obrázek 3: Panel správce

ID odpovídá uživatelskému jménu uživatele. Musí být psáno malými písmeny a nesmí obsahovat žádné diakritické znaménka ani mezery; speciální znaky, jako je tečka a podtržítka, jsou přípustné. Při vytváření nového uživatele je vhodné dodržovat nomenklaturu jméno.příjmení.

Pokud jde o e-mailovou adresu uživatele, z bezpečnostních důvodů důrazně nedoporučujeme používat obecnou e-mailovou adresu.

Protože jsou oprávnění ve výchozím nastavení deaktivována, je na správci, aby je aktivoval ručně:

Permissions

Admin privileges

Grant full access to all system features, users management and settings



Create and upload Samples



Create Report



Delete Sample, Analysis and Project



Edit restricted panels



Launch Analyses



Manage assignments



Manage organisation wide Filter Presets



Manage VKB Knowledge Base



Protocols Management



Report template



Unlock Analyses



Update project settings (lock filters for now)



Validate Diagnosis from VKB



Advanced

Activate 2FA

Enable email two-factor authentication for enhanced security



Temporarily suspending the access. The data is retained, but the user cannot log in or use the system until re-enabled.

 Disable User

Obrázek 4: Oprávnění uživatele

Zabezpečení můžete zvýšit také aktivací duálního ověřování. Tato možnost odešle uživateli jedinečný kód při každém připojení k platformě.

Role "Manage organization wide Filter Presets" umožňuje uživateli upravovat veřejné profily filtrů. Pokud tuto roli nemáte, můžete upravovat pouze soukromé profily filtrů.

Oprávnění spravuje administrátor entity z panelu Admin. Při vytvoření uživatele jsou všechna oprávnění ve výchozím nastavení zakázána a administrátor musí ručně povolit práva vyžadovaná pro daného uživatele. Práva přidělená uživateli se vztahují na všechny entity (pracovní prostory), ke kterým má uživatel přístup na platformě SeqOne (například pokud má uživatel přístup k více entitám, platí pro ně stejná sada oprávnění).

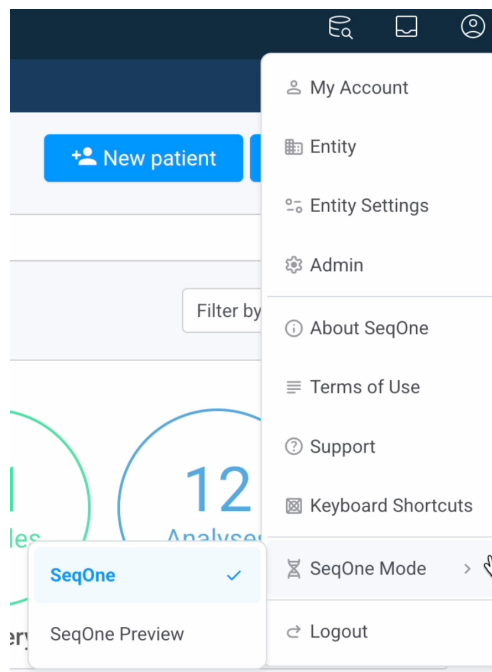
Pokud se zobrazí chybová zpráva, kontaktujte zákaznickou podporu na adrese support@seqone.com.

7.4.4. Přepínání mezi režimy Main a Preview

Připojení k SeqOne přes výchozí přihlašovací URL přeměří uživatele do režimu "Main" jejich subjektu. Tento výchozí režim hostí pracovní sady označené značkou CE a další místní certifikace.

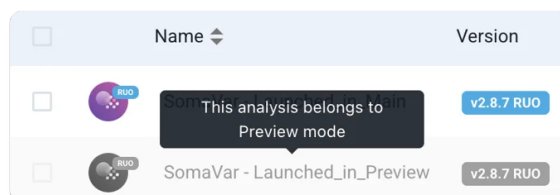
Pro včasný přístup k novým funkcím a verzím pracovních sad přepněte do režimu "Preview" (RUO: Research Use Only) tak, že přejdete do dolní části nabídky nastavení a v části SeqOne One vyberete "Preview".

Režimy "Main" a "Preview" sdílejí stejná základní data (projekty, vzorky, analýzy), oprávnění, prostředky (soubory, šablony, manifesty, seznamy genů), znalosti (VKB) a frekvence (data o frekvenci vzorků, výskyty).



Při navigaci v režimu "Preview" se zobrazí logo "SeqOne Preview RUO" jako připomínka, že tento režim obsahuje funkce a pracovní sady RUO.

V režimu "Main" se analýzy spuštěné v "Preview" zobrazují šedě. Pokud kliknete na jednu z těchto zašedlých analýz, platforma vás vyzve k přepnutí do režimu Preview, abyste si ji mohli prohlédnout.



8. Popis operačních kroků

Krok	Popis
1	Vytvoření projektu Analýzy prováděné na platformě SeqOne jsou organizovány v rámci projektů, které definují charakteristiky analyzovaných vzorků prostřednictvím metadat zadaných uživatelem (typ sekvenování, metoda přípravy, geny zájmu).
2	Nahrání vzorku Samples se nahrávají jako surové sekvenační soubory (FASTQ), které je třeba umístit do určeného projektu. Po nahrání souborů se před jakoukoli analýzou provede ověřovací krok, který zajistí integritu zaslaných souborů.
3	Spuštění analýzy Po nahrání dat na servery SeqOne si uživatel může vybrat z nabídky aplikací nebo "workssets," v závislosti na typu analyzovaného vzorku a typu hledaných genomických změn (SNV, CNV, genové fúze, splicing varianty atd.). Tyto aplikace provádějí bioinformatickou analýzu.
4	Kontrola kvality dat Aplikace dostupné na platformě SeqOne zahrnují několik modulů věnovaných analýze kvality dat pomocí různých sekvenačních metrik, zarovnání a hloubky pokrytí.
5	Interpretace výsledků analýzy Genomové odchylky zjištěné ve vzorku se zobrazují v různých záložkách nebo "viewers", kde může uživatel vyhledat diagnostickou variantu odpovědnou za patologii pacienta.
6	Získání terapeutických doporučení a klinických studií souvisejících s patologií pacienta.
7	Vydání zprávy Výsledky analýzy lze exportovat do zprávy ve formátu PDF, kterou lze následně stáhnout.

9. Vytvoření projektu

Analýzy prováděné na platformě SeqOne jsou organizovány v rámci projektů, které definují charakteristiky analyzovaných vzorků prostřednictvím metadat zadaných uživatelem (typ sekvenování, metoda přípravy, geny zájmu).

9.1. Vytvoření projektu

Po kliknutí na tlačítko "New project" jsou uživatelé vyzváni k zadání hodnot několika typů informací nebo metadat, které se používají k popisu projektu a souvisejících vzorků.

- **Project description**
- **Sekvenování:**
 - Assay type: typ analýzy (Germline nebo Somatic),
 - Projekty Germline (typ testu *Germline*) mají zaškrťovací políčko UMI Trimming, které umožňuje zpracovávat vzorky obsahující UMI. Zaškrtnutím této možnosti budou UMI ořezány ze sekvencí FASTQ, stejně jako v režimu zakázaného UMI pro somatické pracovní sady (viz část **Volba režimu zpracování UMI**).
 - Sequencing platform: sekvenační platforma
 - Library type: typ provedeného read sekvenování (*Paired-end* nebo *Single-end*),
 - Data type: DNA, RNA, ctDNA nebo Methylation Sequencing,
 - Pokud je jako typ dat vybrána DNA, je třeba zadat metodu sekvenování pomocí rozevírací nabídky "*DNA Sequencing method*" a metodu amplifikace (Capture nebo Amplicon) prostřednictvím "*Amplification method*".
 - File input type: *Soubory Fastq, VCF*
 - Genome: verze referenčního genomu (GRCh37 nebo GRCh38),
 - Low-Quality Mapping Region: volání variant v oblastech s nízkou kvalitou zarovnání pomocí čtení MAPQ0,
 - Sequencing method: Gene Panel, Whole exome sequencing, Comprehensive Genomic Profiling, Shallow WGS, Whole genome sequencing, CGH MicroArrays.
 - Secondary analysis provider: AmoyDX, Pillar, Thermofisher,
 - Manifest: Manifest v případě panelu genů,
 - ID marker (nepovinné): soubor s označením identity.
- **Vzorku:**
 - (Nepovinné) Nová vzorová metadata lze přidat kliknutím na možnost "Můžete si také vytvořit vlastní metadata". Po výzvě zadejte název a hodnotu metadat.

Po zadání různých metadat projektu jsou uživatelé vyzváni k vytvoření projektu kliknutím na tlačítko "Create project".

9.2. Řízení projektů

Prostřednictvím "Dashboardu" mohou uživatelé spravovat různé projekty vytvořené na platformě. Mohou zejména:

- Přístup k projektům získáte na kartě "Projects" a ke vzorkům a analýzám každého projektu kliknutím na loga "Samples" a "Analysis"
- Přístup k analýzám je možný přímo přes kartu Analysis.
- Nový projekt vytvoříte kliknutím na "New project" (viz část **Vytvoření projektu**).

9.3. Delete project

Chcete-li projekt odstranit, musí uživatelé:

- Vyberte projekt
- Klikněte na kartu "Settings" a poté na "Delete project".
- Potvrďte zaškrtnutím "Yes" v zobrazené zprávě. Tato akce je nevratná.

9.4. Data DNA (sekvenování DNA, sekvenování ctDNA)

Low-Quality Mapping Region. Při sekvenování genových panelů, exomů nebo genomů může uživatel při volání variant zahrnout oblasti s nízkou kvalitou zarovnání. Tyto oblasti odpovídají tzv. čtením MAPQ0, jejichž kvalita zarovnání je rovna 0, jako jsou pseudogeny a homologní oblasti genomu.

Tato možnost se vztahuje na konkrétní cílové geny nebo oblasti. Pro použití této možnosti je uživatel vyzván k načtení manifestu zaměřeného na požadované oblasti nebo k jeho výběru prostřednictvím rozbalovací nabídky.

Tato možnost může snížit kvalitu a přesnost výsledků. Může vnést šum do frekvence alel na pozadí a mít negativní dopad na prioritizaci variant. Může také prodloužit dobu analýzy v důsledku vyššího výpočetního zatížení.

Exome. V případě sekvenování exomu nemusí uživatelé specifikovat manifest, který má být použit; jsou vyzváni, aby si vybrali výrobce použitého kitu kliknutím na jednu z ikon v kategorii Exome kit provider. Rozbalovací nabídka pak uživateli umožní vybrat požadovaný manifest. Pokud požadovaný manifest není na platformě k dispozici, je uživatel vyzván, aby kontaktoval uživatelskou podporu support@segone.com.

Gene panel. V případě panelu genů jsou uživatelé vyzváni, aby v rozbalovací nabídce vybrali "Amplification method" a také manifest (ve formátu BED) odpovídající genům, které je zajímají. Pokud požadovaný manifest není na platformě k dispozici, lze jej nahrát výběrem možnosti "Upload new manifest" (Nahrát nový manifest) a následováním pokynů na obrazovce.

Shallow WGS, CGH MicroArrays, Whole genome sequencing. V případě sekvenování genomu nemusí uživatel určovat, který manifest má být použit.

Capture. Při výběru metody amplifikace Capture je k dispozici ikona UMI (Unique Molecular Identifiers). Jedinečné molekulární identifikátory jsou indexy, které se přidávají před amplifikací, aby pomohly eliminovat duplicitu PCR a snížit šum pozadí po analýze.

Amplicon. Při analýze ampliconů je třeba věnovat zvláštní pozornost amplifikačním primerům. Při vytváření projektu pro data ampliconů se uživatelům doporučuje vybrat manifest, ve kterém oblasti odpovídající cílům neobsahují sekvence primerů. Pokud takový BED není k dispozici, lze zvolit možnost ořezat pevnou délku na konci sekvencí.

Identity monitoring. Pro metody sekvenování odpovídající panelu genů, Exome, Genomovému profilování (CGP) nebo Genomovému (mělké WGS) ze souborů FASTQ je uživatel vyzván, aby z rozbalovací nabídky vybral soubor markerů pro monitorování identity. Pokud požadovaný soubor není na platformě k dispozici, lze jej importovat výběrem možnosti "Upload new markers" (Nahrát nové markery) a následováním pokynů na obrazovce (viz také odstavec **Marker** v části **Správa uživatelských nastavení**)).

9.5. Data RNA (sekvenování RNA)

Pokud je jako typ dat vybrána RNA, musí uživatelé prostřednictvím rozevírací nabídky *Stranded protocol* určit, zda jejich protokol zachovává informace o původním vlákně. Pokud je protokol specifický pro dané vlákno, je třeba zadat pořadí syntézy vláken (*dopředu-dozadu* nebo *dozadu-dopředu*) pomocí možnosti *Read orientation* (*Orientace čtení*). Nakonec je třeba zadat metodu sekvenování (Gene panel).

Read orientation. Pro zadání orientace sekvence jsou k dispozici dvě možnosti:

- **Forward-Reverse** : první read (R1) pro každý pár odpovídá sekvenci přepsané messengerové RNA (sense strand).
- **Reverse-Forward**: první read (R1) pro každý pár odpovídá komplementární sekvenci přepisované messengerové RNA (antisense vlákno).

Obecně je orientace reads u níže uvedených sad následující:

- Agilent: Reverse-Forward
- Illumina: Unstranded
- Roche Kapa Hypercap: Reverse-Forward
- Twist: Reverse-Forward

V případě pochybností se obraťte na podporu (support@seqone.com).

Po zadání typu dat a správné konfiguraci projektu jsou uživatelé vyzváni k ověření kliknutím na tlačítko "Next step".

9.6. Nastavení projektu

Na kartě "Settings" projektu můžete :

- upravit název a popis projektu v *polích* "Project name" a "Project description",
- upravit profil kontroly kvality (QC),
- změnit výchozí profil filtru v tabulce variant,
- trvale odstranit projekt.

9.6.1. Konfigurace výchozích metrik QC projektu

Uživatel si může vybrat, zda chce použít výchozí metriky kontroly kvality, nebo definovat vlastní prahové hodnoty na kartě Nastavení projektu. Uživatel je vyzván, aby se obrátil na uživatelskou podporu (support@seqone.com) a definoval požadované hodnoty.

Budou použity vlastní metriky QC :

- *QC checks* v zobrazení Analyses projektu,
- *Kontroly kontroly kvality* v části *automatické validace kontroly kvality* na kartě kontroly kvality analýzy,
- na kartu *GeneCov* analýzy.

9.6.2. Nastavení výchozího profilu filtru projektu

Uživatel může na kartě Nastavení projektu nastavit předdefinovaný profil filtru, který se použije na výchozí tabulku variant (viz také část **Použití předdefinovaného profilu filtru**).

Profil filtru lze vybrat pro každou workset pomocí rozevírací nabídky, pokud je veřejný. Profil filtru se pak použije ve výchozím nastavení při zobrazení tabulky variant pro všechny analýzy projektů pro všechny uživatele v organizaci.

Pokud uživatel tento parametr při vytváření projektu nezmění, bude výchozí hodnotou profil definovaný společností SeqOne.

Výchozí profily filtrů lze uzamknout kliknutím na přepínač "*Lock Variants Filters*".

10. Nahrání vzorku

10.1. Nahrávání vzorků

Uživatelé mohou importovat své vzorky pomocí ikony "Add samples" v každém projektu. Vzorky lze do platformy SeqOne importovat dvěma způsoby:

Z počítače

Uživatelé mohou nahrát své soubory FastQ přímo ze složky v počítači kliknutím na kartu "From your computer" na ikoně "Add files". Uživatelé mohou také přetáhnout soubory ze svého počítače do vyhrazené oblasti v rozhraní.

Přes SDS (SeqOne Data Sync)

SDS je volitelný prémiový modul pro automatické vytváření projektů, přenos dat, spouštění analýz a obnovu strukturovaných dat (TSV) pro integraci do systému správy laboratorních informací (LIMS).

Tento modul kompatibilní s operačními systémy Linux, MacOS a Windows je spustitelný, který umožňuje uživatelům ověřovat se a komunikovat s rozhraním SeqOne API v předem definovaném a vysoce zabezpečeném prostředí.

Minimální požadavky:

- Linux minimálně 1.19
- MacOS 10.15 s minimálním procesorem Intel
- MacOS s minimálně procesorem Apple Silicon (čip M1)
- Minimálně 64bitový systém Windows 10
- 4 GB RAM (8 GB pro Windows)
- Čtyřjádrový procesor
- Přístup k internetu a platformě SeqOne (app.eu.seqone.com)
- Přístup k systému ukládání dat (přímo nebo přes síť).

Pokud chcete toto řešení použít nebo získat další informace, kontaktujte náš tým podpory (support@seqone.com).

Stav nahrávání vzorků

Platforma SeqOne používá různé stavy, které uživatelům umožňují sledovat průběh nahrávání vzorků:

- *Upload pending*: probíhá import,
- *Verifying sample...:* ověřuje se shoda vzorku,
- *Mapping...:* krok zarovnání vzorku,
- *Verified sample*: ověřené vzorky
- *Upload failed*: import se nezdařil.

Kapacita nahrávání

Kapacita nahrávání souborů (zejména souborů fastq.gz) závisí přímo na internetovém připojení uživatele. Při nahrávání vzorků je proto třeba brát v úvahu rychlost internetového připojení. Pro jasnější představu o odhadovaných časech nahrávání viz následující tabulka:

Estimation du temps d'attente				
	ADSL		Fibre optique (Mbits/s)	
	Ethernet : 10 Mbits/s	Wifi : 7,5 Mbits/s	Ethernet : 100 Mbits/s	Wifi : 90 Mbits/s
Poids du fichier fastq				
100Mo	1min 20s	1min 45s	8s	9s
1Go	13min 20s	17min 45s	1min 20s	1 min 30s
10Go	2h 13min 20s	2h 57mn	13min 20s	15 min
100Go	22h 22min	30h	2h 13min 20s	2h 30min

Obrázek 5: Odhadovaná doba nahrávání v závislosti na velikosti souboru a rychlosti internetového připojení.

10.2. Formát vstupního souboru

10.2.1. Soubory primární analýzy

Formát souboru, který se má nahrát pro každý vzorek, závisí na použité sekvenační platformě. Názvy souborů musí odpovídat konvenci pro pojmenování souborů v každém sekvenátoru.

Před importem do platformy je třeba soubory komprimovat pomocí přípony .gz (gzip). V případě párového sekvenování se očekávají dva soubory odpovídající reads z obou konců každé molekuly.

Pokud byl vzorek v době importu platformou správně rozpoznán, zobrazí se celý název souboru. Poté je možné soubor přejmenovat kliknutím na něj. Párové soubory se automaticky spojí do jednoho vzorku.

Platforma podporuje následující sekvencery:

- [Illumina](#) formát fastq.gz, pouze párové zakončení,
- [Element Biosciences](#) : formát fastq.gz, pouze párové koncovky,
- [MGI](#) : formát fastq.gz, pouze párové řetězce. Očekávané názvy souborů jsou: sample_1.fq.gz a sample_2.fq.gz nebo flowcellID_L01_sample_1.fq.gz a flowcellID_L01_sample_2.fq.gz.

10.2.2. Soubory sekundární analýzy

Formát souboru, který se má nahrát pro každý vzorek, závisí na sekvenační platformě a aplikaci použité pro sekundární analýzu. Soubory musí být před importem z platformy komprimovány s příponou .gz (gzip), pokud není uvedeno jinak.

Formát importu souborů sekundárních analýz vytvořených pomocí aplikací Illumina, Oxford Nanopore Technologies, Agilent, AmoyDx, Pillar, Thermofisher závisí na použité pracovní sadě (viz část **Soubory analýz dostupné ze souborů sekundárních analýz**).

Pokud platforma při importu správně rozpozná vzorek, zobrazí celý název souboru. Kliknutím na soubor jej můžete přejmenovat. Soubory, které mají stejný základní název, se automaticky spojí do jednoho vzorku.

Soubory DRAGEN. Soubory a formáty očekávané pro jednotlivé worksety jsou následující, pokud jsou při vytváření projektu vybrány soubory VCF - Secondary Analysis data a metadata Illumina:

GermVar Tertiary, GermVar Tertiary Family sample.hard-filtered.vcf.gz
sample.cnv.vcf.gz
sample.sv.vcf.gz
sample.repeats.vcf.gz

Soubory Oxford Nanopore Technologies. Očekávané soubory a formáty jsou při výběru souborů VCF následující - data sekundární analýzy a metadata Oxford Nanopore Technologies při vytváření projektu. Uživatelé mohou importovat buď samotný soubor Variant (SNP), nebo jej kombinovat s doplňkovými soubory pro analýzu CNV, STR nebo SV:

GermVar Tertiary, GermVar Tertiary Family sample.wf_snp.vcf.gz
sample.wf_cnv.vcf.gz
sample.wf_str.vcf.gz
sample.wf_sv.vcf.gz

Thermofisher/Affymetrix files. Formát importu pro soubory sekundární analýzy CGH je formát .txt generovaný aplikací Affymetrix společnosti Thermofisher. Tento formát je podporován pouze v případě, že jsou při vytváření projektu vybrány metadata Thermofisher/Affymetrix i CGH MicroArrays. Soubory a formáty očekávané pro každou pracovní sadu jsou následující:

GermVar Tertiary ARRAY_SampleID.txt

Soubory CGH MicroArrays. Formátem pro import souborů sekundární analýzy CGH je formát .xls generovaný aplikací Agilent Cytogenomics. Tento formát je podporován pouze v případě, že jsou

při vytváření projektu vybrána metadata Agilent i CGH MicroArrays. Soubory a formáty očekávané pro jednotlivé worksety jsou následující:

GermVar Tertiary IntervalBasedReport_sample.xls

Soubory Alissa Reporter. Importní formát pro soubory sekundární analýzy Alissa Reporter je formát .xls generovaný aplikací Alissa Reporter společnosti Agilent. Tento formát je podporován pouze v případě, že jsou při vytváření projektu vybrána metadata Agilent i CGH MicroArrays. Soubory a formáty očekávané pro jednotlivé worksety jsou následující:

GermVar Tertiary IntervalBasedReport_sample.xls
ABERRATION_&_LOH_INTERVAL_sample.xls

PacBio-HiFi files. Pokud jsou při vytváření projektu vybrány soubory VCF (Secondary Analysis data) a metadata PacBio-HiFi, očekávané soubory a formáty pro každý workset jsou následující:

GermVar Tertiary, GermVar Tertiary Family sample.hard-filtered.vcf.gz
sample.cnv_sv.vcf.gz
sample.repeats.vcf.gz

Soubory VCF pro panely somatických genů. Formátem pro import souborů sekundární analýzy SNV, CNV a fúzí je formát .vcf.gz generovaný aplikacemi Illumina, AmoyDx, Pillar a Thermofisher. Původní názvy souborů je třeba převést tak, aby odpovídaly požadovanému formátu pojmenování popsanému níže.

Při výběru souborů VCF se očekávají následující soubory a formáty - data sekundární analýzy a metadata Gene Panel při vytváření somatického projektu. Uživatelé mohou importovat buď samotný soubor Variant (SNV), nebo jej kombinovat s doplňkovými soubory pro analýzu CNV nebo fúzí:

SomaVar Tertiary sample.snv.vcf.gz
sample.cnv.vcf.gz
sample.fus.vcf.gz

Při výběru metadat „VCF files - Secondary Analysis data“ a „Comprehensive Genomic Profiling“ během vytváření somatického projektu se očekávají následující soubory a formáty. Uživatelé mohou importovat multi-variant VCF soubor, například soubor VCF „All Variants“ z Oncomine™ Comprehensive Assay V3/Plus.

SomaVar Tertiary Oncomine_<LibPrepID>_<analysisID>.vcf

10.3. Metadata vzorku

Pro každý nahraný vzorek mohou uživatelé přiřadit hodnoty metadat, které vybrali při vytváření projektu, nebo definovat nové hodnoty kliknutím na ikonu v pravém horním rohu seznamu vzorků . 

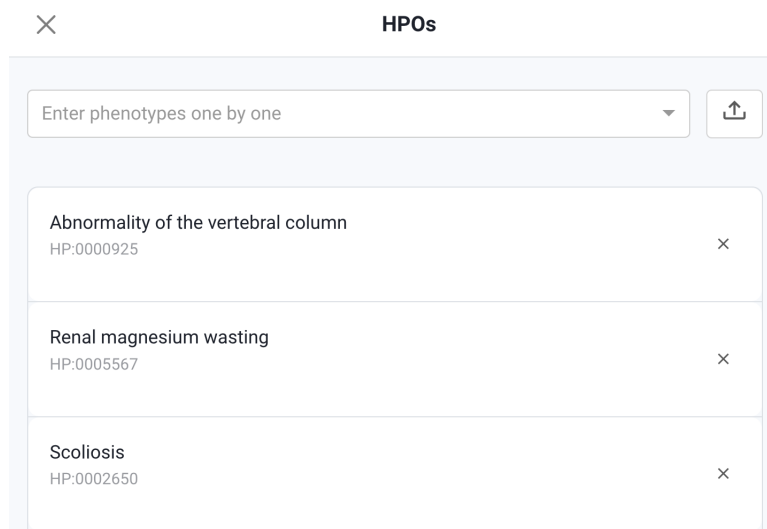
Některá vzorová metadata lze nahrát přímo jako soubor s hodnotami oddělenými tabulátory výběrem ikony . 

Tato kritéria hrají důležitou roli při zajišťování sledovatelnosti údajů a při vzájemném porovnávání různých analýz prováděných na platformě. Do procesu identifikace variant jsou zapojena některá metadata, jako je pohlaví, ID nemoci a HPO.

Pokud je pole „Assay Type“ vyplněno hodnotou „Germline“ , jsou metadata spojená se vzorkem ve výchozím nastavení „HPO“, „Sex“ a skupina „GnomAD“. Pokud je pole „Assay Type“ vyplněno hodnotou „Somatic“, jsou metadata spojená se vzorkem ve výchozím nastavení „Disease“ a „Sex“.

Chcete-li přidat nebo upravit HPO ve vzorku, můžete po výběru vzorku kliknout na "+ Add" ve sloupci HPO. Tím se otevře vyskakovací okno umožňující postupné přidávání HPO vyhledáním fenotypu nebo kódu HPO ve vyhledávacím řádku připojeném k rozhraní API hpo.jax.org. Je také

možné importovat textový soubor obsahující HPO kódy kliknutím na ikonu  . Poté je k dispozici přehled zadáných termínů s možností případného odstranění termínů.



Obrázek 6: Vyskakovací okno pro import HPO

10.4. Upravit vzorek

Název i metadata spojená se vzorkem může uživatel kdykoli upravit, a to výběrem vzorku v seznamu "Samples" a kliknutím na ikonu . Aby byly změny účinné, musí být následně potvrzeny.

10.5. Odstranění vzorku

Uživatelé mohou soubory FastQ vzorku z platformy odstranit kliknutím na ikonu  po výběru požadovaného vzorku. Tyto soubory uvolní místo v úložišti a výsledky budou stále k dispozici na platformě SeqOne.

Chcete-li vzorek odstranit, klikněte po výběru požadovaného vzorku na ikonu  na kartě "Samples". "Sample" lze trvale odstranit až po odstranění všech analýz, které jsou s ním spojeny.

11. Analýza

11.1. Spuštění nové analýzy

Uživatelé mohou spustit novou analýzu kliknutím na "New analysis" na obrazovce Overview pro každý projekt.

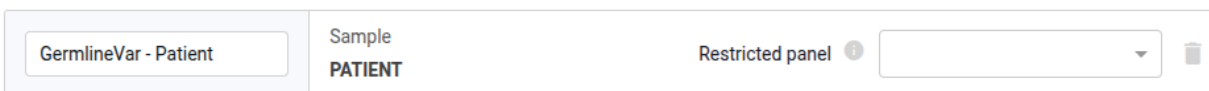
Spuštění analýzy se provádí ve třech různých krocích, které jsou popsány níže.

Krok "1. Workset" umožňuje vybrat vhodnou pracovní sadu analýzy (viz část **Výběr pracovního setu pro analýzu**).

Krok "2. Select Samples" je rozdělen na dvě části:

- 2.1. Nastavení možnosti:
 - popisuje vybranou sadu a referenční genom
 - umožňuje přidání *výplně* prostřednictvím funkce "Extra calling padding" (viz část **Výběr pracovního setu pro analýzu**).
 - Umožňuje vybrat režim zpracování UMI (viz část **Volba režimu zpracování UMI**), pokud je to vhodné.
 - umožňuje zvolit CNV baseline, MSI baseline nebo upravit minimální práh VAF, je-li to možné.
 - Consente di scegliere il protocollo.
- 2.2. Select sample(s): umožňuje vybrat vzorky pro analýzu. Uživatelé mohou konkrétně vybrat nové vzorky, které jsou označeny jako "New" (Nové), pomocí funkce "select all new" (Vybrat všechny nové).

V kroku "3. Confirm Analysis" se provede shrnutí vybraných vzorků. Pro každou analýzu, která má být spuštěna, je možné použít omezený panel výběrem jednoho nebo více genových panelů prostřednictvím možnosti "Restricted panel" (viz část **Modální zobrazení Panel**).



The screenshot shows a user interface for selecting a panel. On the left, there is a box containing the text "GermlineVar - Patient". To the right, there is a section labeled "Sample" with the word "PATIENT" below it. Further right, there is a label "Restricted panel" followed by a dropdown menu. The dropdown menu is currently open, showing "PATIENT" as the selected option. There is also a small trash icon to the right of the dropdown menu.

Možnost "Restricted panel" použije na analýzu filtr v podobě jednoho nebo více panelů genů. Tato možnost nemá vliv na volání variant, které je určeno manifestem projektu a přidáním *paddingem*. Volba uzamčených panelů je vratná; uživatel může tyto panely upravit na kartě Panel postranního panelu pro správu analýzy po provedení analýzy (viz část **Modální zobrazení Panel**).

11.1.1. Výběr pracovního setu pro analýzu

K dispozici je několik analytických worksetů pro zodpovězení různých biologických otázek. "Worksets" navržené při vytváření nové analýzy závisí na metadatech, která byla definována při vytváření projektu, a mohou se proto lišit v závislosti na datech.

Worksets Analysis dostupné ze souborů FASTQ

Pokud jsou při vytváření projektu vybrána metadata *Fastq files - Primary Analysis data*, jsou k dispozici následující worksety.

S výjimkou projektů, kde je metoda amplifikace ampikon, *umožňují* některé worksety (tj. GermlineVar, GermlineFamily, SomaVar, SomaVar LF (verze ≥ 2.0), SomaCGP (verze ≥ 0.6), GermVar, GermVar Family, SomaLBx, SomaMethyl) *rozšířit* oblast vyhledávání pro volání variant a přidat k původnímu manifestu 0 až 100 bp padding.

Extra calling padding (between 0 and 100bp)  Apply extra padding on the manifest for calling, this won't affect the computation of QC metrics.

Tato možnost je k dispozici v kroku 2 (Select Samples) spuštění analýzy prostřednictvím funkce "Extra calling padding". Ve výchozím nastavení je hodnota nastavena na 0 a musí se jednat o celé přirozené číslo.

Některé pracovní sady (např. SomaVar (verze ≥ 2.6) a SomaHemato (verze ≥ 0.2)) nabízejí možnost nastavit minimální prahovou hodnotu VAF pro detekci variant. Tato prahová hodnota může být nastavena v rozmezí 0,01 % až 1 %, přičemž výchozí hodnota je 1 %.

Minimal VAF threshold (between 0.01% and 1%)  Apply a custom minimal VAF threshold for short variant detection. Sensitivity and precision cannot be guaranteed when deviating from the default pipeline value (1% for SomaVar)

Další informace o této možnosti získáte od uživatelské podpory na adrese support@seqone.com.

GermlineVar



GermlineVar se používá k detekci zárodečných mutací na základě sekvenování panelu genů nebo exomu pacienta.

Při provádění analýzy GermlineVar jsou uživatelé vyzváni, aby:

- Vyberte verzi pracovní sady, kterou chcete spustit, a klikněte na tlačítko "Select",
- Vyberte vzorky, které chcete analyzovat, a klikněte na "Next step",
- Spusťte analýzu kliknutím na "Start analysis".

GermVar



GermVar umožňuje detekci zárodečných mutací na základě sekvenování panelu genů, exomu nebo genomu pacienta.

Při provádění analýzy GermVar jsou uživatelé vyzváni, aby:

- Vyberte verzi pracovní sady, kterou chcete spustit, a klikněte na tlačítko "Select",
- Vyberte model, který se má použít,
- Vyberte vzorek (vzorky) k analýze a klikněte na "Next step",
- Spusťte analýzu kliknutím na "Start analysis".

GermlineVarLR



GermlineVarLR se používá k detekci zárodečných mutací v long-read sekvenačních datech ze sekvenování genových panelů.

Při provádění analýzy GermlineVarLR jsou uživatelé vyzváni, aby:

- Vyberte verzi workset, kterou chcete spustit, a klikněte na tlačítko "Select",
- Vyberte model, který se má použít, a vzorek (vzorky), který se má analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start analysis".

GermlineFamily



GermlineFamily slouží k detekci dědičných mutací na základě sekvenování panelu genů, exomu nebo genomu pacienta a členů jeho rodiny. Tuto pracovní sadu lze použít také pro Carrier screening.

Při provádění analýzy GermlineFamily musí uživatelé:

- Vyberte verzi pracovní sady, kterou chcete spustit, a klikněte na tlačítko "Select",
- Vyberte alespoň 2 vzorky
- V části "Set conditions" (Nastavit podmínky) uveďte, zda je vzorek pacientův nebo rodinného příslušníka, a to definováním rodinného odkazu v části "Family link" (Rodinný odkaz). Pacient je automaticky označen jako "affected". V případě screeningu nosičství uveďte u každého vzorku rodinný vztah pomocí "Mother" (Matka) nebo "Father" (Otec), aniž byste zaškrtnuli políčko "Is affected?" (Je postižen?).
- Zkontrolujte, zda parametry odpovídají, a klikněte na "Next step",
- Spusťte analýzu kliknutím na "Start analysis".

GermVar Family



GermVar Family umožňuje detekci zárodečných mutací na základě sekvenování panelu genů, exomu nebo genomu pacienta.

Při provádění analýzy GermVar Family jsou uživatelé vyzváni, aby:

- Vyberte verzi pracovní sady, kterou chcete spustit, a klikněte na tlačítko "Select",
- Vyberte model, který se má použít,
- Vyberte alespoň 2 vzorky,
- V části "Set conditions" (Nastavit podmínky) uveďte, zda je vzorek pacientův nebo rodinného příslušníka, a to definováním rodinného odkazu v části "Family link" (Rodinný odkaz). Pacient je automaticky označen jako "affected". V případě screeningu nosičství uveďte u každého vzorku rodinný vztah pomocí "Mother" (Matka) nebo "Father" (Otec), aniž byste zaškrtnuli políčko "Is affected?" (Je postižen?).
- Zkontrolujte, zda parametry odpovídají, a klikněte na "Next step",
- Spusťte analýzu kliknutím na "Start analysis".

CNVCapture



CNVCapture slouží k detekci zárodečných variací počtu kopií (CNV) z dat získaných z genových panelů.

Při provádění analýzy CNVCapture jsou uživatelé vyzváni, aby:

- Vyberte verzi pracovní sady, kterou chcete spustit, a klikněte na tlačítko "Select",
- Vyberte minimální počet 8 vzorků k analýze a klikněte na "Next step",
- Spusťte analýzu kliknutím na "Start analysis".

CNVExome

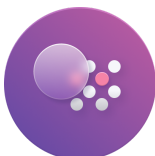


CNVExome slouží k detekci variací počtu kopií (CNV) z dat získaných z exomu.

Při provádění analýzy CNVExome jsou uživatelé vyzváni, aby:

- Vyberte verzi workset, kterou chcete spustit, a klikněte na tlačítko "Select".
- Vyberte model, který se má použít, a vzorek (vzorky), který se má analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start analysis".

SomaVar



SomaVar detekuje somatické mutace ze vzorků nádorů. Tuto pracovní sadu lze použít se záchyťovými a Ampliconovými daty.

Při provádění analýzy SomaVar jsou uživatelé vyzváni, aby:

- Vyberte verzi pracovní sady, kterou chcete spustit, a klikněte na tlačítko "Select".
- V možnostech vyberte z režimů způsobilých pro analýzu příslušný režim zpracování UMI (viz část **Volba režimu zpracování UMI**).
- Vyberte vzorek (vzorky), který chcete analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start analysis".

SomaCNVCapture



SomaCNVCapture slouží k detekci somatických variací počtu kopií (CNV) z dat získaných z panelu capture.

Při provádění analýzy SomaCNVCapture jsou uživatelé vyzváni, aby:

- Vyberte verzi workset, kterou chcete spustit, a klikněte na tlačítko "Select".
- V možnostech vyberte z režimů způsobilých pro analýzu příslušný režim zpracování UMI (viz část **Volba režimu zpracování UMI**).
- Vyberte minimální počet osmi vzorků, které chcete analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start analysis".

SomaRNA



SomaRNA lze použít k detekci genových fúzí z dat zachycených metodou RNA-seq.

Při provádění analýzy SomaRNA jsou uživatelé vyzváni, aby:

- Vyberte verzi pracovní sady, kterou chcete spustit, a klikněte na tlačítko "Select".
- V možnostech vyberte z režimů způsobilých pro analýzu příslušný režim zpracování UMI (viz část **Volba režimu zpracování UMI**).
- vyberte vzorek (vzorky), který (které) chcete analyzovat, a klikněte na "Next step",
- Spusťte analýzu kliknutím na "Start analysis".

SomaMSI



SomaMSI zjišťuje stav mikrosatelitních oblastí (MSS nebo MSI).

Při provádění analýzy SomaMSI jsou uživatelé vyzváni, aby:

- Vyberte minimální počet osmi vzorků, které chcete analyzovat, a klikněte na "Next step".
- V možnostech vyberte z režimů způsobilých pro analýzu příslušný režim zpracování UMI (viz část **Volba režimu zpracování UMI**).
- vyberte minimální počet 8 vzorků k analýze a klikněte na "Next step",
- Spusťte analýzu kliknutím na "Start analysis".

SomaHRD

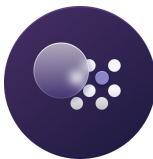


SomaHRD umožňuje detekci HRD statusu z mělkých WGS dat samotných nebo doprovázených panelem genů včetně BRCA1 a BRCA2.

Při provádění analýzy SomaHRD jsou uživatelé vyzváni, aby:

- Vyberte verzi worksetu, kterou chcete spustit, a klikněte na tlačítko "Select".
- Zadejte metadata o buněčnosti (% nádorových buněk) pro každý vzorek (viz oddíl **Metadata vzorku**).
- Vyberte vzorek (vzorky), který chcete analyzovat, a klikněte na "Next step",
- Pro každý vzorek vyberte příslušnou analýzu SomaVar nebo SomaCGP (GRCh37 nebo GRCh38) pomocí možnosti "Select panel analysis (Optional)",
- Spusťte analýzu kliknutím na "Start *analysis*".

SomaVar LF

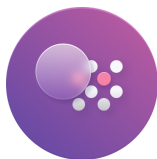


Detekce nízkofrekvenčních krátkých variant a strukturálních variant, jakož i fúzních genů ze vzorků nádorů (solidní nebo tekuté biopsie, cirkulující nádorová DNA) obohacených o záchyt panelu genů.

Při provádění analýzy SomaVar LF jsou uživatelé vyzváni, aby:

- Vyberte verzi worksetu, kterou chcete spustit, a klikněte na tlačítko "Select".
- V možnostech vyberte z režimů způsobilých pro analýzu příslušný režim zpracování UMI (viz část **Volba režimu zpracování UMI**).
- Vyberte vzorek (vzorky), který chcete analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start *analysis*".

SomaCGP



SomaCGP umožňuje detekci somatických mutací, strukturních variant, genových fúzí a také MSI a TMB statusu ze vzorků nádorů. Tato pracovní sada je určena pro analýzu FFPE DNA připravené pomocí panelu SureSelect Cancer CGP Assay společnosti Agilent a je k dispozici pouze tehdy, pokud jsou při vytváření projektu vybrána metadata Comprehensive Genomic Profiling. Tuto pracovní sadu lze použít s daty ze zachytu.

Při provádění analýzy SomaCGP jsou uživatelé vyzváni, aby:

- Vyberte verzi worksetu, kterou chcete spustit, a klikněte na tlačítko "Select".
- Na požádání vyberte „CNV baseline“ a/nebo „MSI baseline“ (pokud není vybrána žádná základní hodnota, analýza bude provedena bez detekce CNV/MSI).
- V možnostech vyberte z režimů způsobilých pro analýzu příslušný režim zpracování UMI (viz část **Volba režimu zpracování UMI**).
- Vyberte vzorek (vzorky), který chcete analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start analysis".

SomaHemato

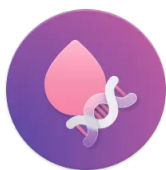


SomaHemato umožňuje detekci somatických mutací ze vzorků s hematologickou specifikou. Tuto pracovní sadu lze použít s daty ze zachytu.

Při provádění analýzy SomaHemato jsou uživatelé vyzváni, aby:

- Vyberte verzi worksetu, kterou chcete spustit, a klikněte na tlačítko "Select".
- V možnostech vyberte z režimů způsobilých pro analýzu příslušný režim zpracování UMI (viz část **Volba režimu zpracování UMI**).
- Vyberte vzorek (vzorky), který chcete analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start analysis".

SomaLBx



SomaLBx umožňuje detekci nízkofrekvenčních somatických mutací, strukturních variant a genových fúzí z ctDNA. Tato pracovní sada je určena k analýze ctDNA a lze ji použít s daty ze zachytu.

Při provádění analýzy SomaLBx jsou uživatelé vyzváni, aby:

- Vyberte verzi workset, kterou chcete spustit, a klikněte na tlačítko "Select".
- Vyberte „CNV baseline“ (pokud není vybrána žádná základní linie, analýza bude probíhat bez detekce CNV).
- V možnostech vyberte z režimů způsobilých pro analýzu příslušný režim zpracování UMI (viz část **Volba režimu zpracování UMI**).
- Vyberte vzorek (vzorky), který chcete analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start analysis".

SomaMethyl



SomaMethyl umožňuje detekci metylovaných CpG ostrovů z cirkulujících vzorků nádorové DNA. Tato pracovní sada je určena k analýze metylace v ctDNA a lze ji použít s daty zachycenými bisulfátem.

Při provádění analýzy SomaMethyl jsou uživatelé vyzváni, aby:

- vybrat verzi workset k provedení a kliknout na "Select",
- vybrat "Methyl baseline" odpovídající kontrolní kohortě (pokud není vybrána žádná kohorta, analýza bude spuštěna bez detekce prahové hodnoty a diferenciální metylace),
- vybrat vzorek (vzorky) k analýze a kliknout na "Next step",
- spustit analýzu kliknutím na "Start analysis".

Soubory analýz dostupné ze souborů sekundárních analýz

Následující worksets jsou k dispozici, pokud jsou při vytváření projektu vybrány *soubory VCF - Secondary Analysis data* nebo pokud jsou vybrána metadata Agilent i CGH MicroArrays (viz část **Secondary Analysis files**). V tomto případě platforma obdrží výstupní data z aplikace vybraného poskytovatele, přistoupí k anotaci a vloží varianty do příslušných vizualizačních rozhraní.

GermVar Tertiary



Analýza variant detekovaných aplikací Illumina Germline nebo aplikací EPI2ME od Oxford Nanopore Technologies z dat WGS a WES nebo variant detekovaných aplikací Agilent Cytogenomics z dat MicroArray CGH.

Při provádění terciární analýzy GermVar jsou uživatelé vyzváni, aby:

- Vyberte verzi workset, kterou chcete spustit, a klikněte na tlačítko "Select",
- Vyberte vzorek (vzorky), který chcete analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start analysis".

GermVar Tertiary Family

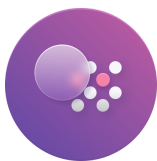


Analýza variant detekovaných aplikací Illumina Germline nebo aplikací EPI2ME od společnosti Oxford Nanopore Technologies z dat WGS a WES.

Při provádění analýzy GermVar Tertiary Family jsou uživatelé vyzváni, aby:

- Vyberte verzi pracovní sady, kterou chcete spustit, a klikněte na tlačítko "Select",
- Vyberte alespoň 2 vzorky,
- V části "Set conditions" (Nastavit podmínky) uveďte, zda je vzorek pacientův nebo rodinného příslušníka, a to definováním rodinného odkazu v části "Family link" (Rodinný odkaz). Pacient je automaticky označen jako "affected". V případě screeningu nosičství uveďte u každého vzorku rodinný vztah pomocí "Mother" (Matka) nebo "Father" (Otec), aniž byste zaškrtnuli políčko "Is affected?" (Je postižen?).
- Zkontrolujte, zda parametry odpovídají, a klikněte na "Next step",
- Spusťte analýzu kliknutím na "Start analysis".

SomaVar Tertiary



SomaVar Tertiary je pipeline terciární analýzy založená na VCF, která zpracovává jak soubory VCF Illumina WES, tak soubory VCF genových panelů od různých poskytovatelů kitů (viz část **Formát vstupního souboru**).

Při provádění SomaVar Tertiary analýzy jsou uživatelé vyzváni, aby:

- Vyberte verzi pracovní sady, kterou chcete spustit, a klikněte na tlačítko "Select",
- Vyberte vzorek (vzorky), který chcete analyzovat, a klikněte na "Next step".
- Spusťte analýzu kliknutím na "Start analysis".

11.1.2. Volba režimu zpracování UMI

Prezentace režimů zpracování UMI

V praxi platforma navrhuje jednu nebo více metod zpracování UMI při provádění analýzy:

1. **Standard mode** (doporučený): konsenzus je generován z PCR duplikátů se stejným UMI, pokud je jejich počet větší nebo roven 2. UMI reprezentované jednou sekvencí (singletony) jsou rovněž zachovány. Minimální alelická frekvence uváděných variant a minimální počet reads podporujících mutaci (AO, Alternate Observation) se liší podle použitého pracovního souboru:
 - U SomaVar, SomaHemato, SomaRNA, SomaCNVCapture, SomaCGP a SomaMSI je minimální alelická frekvence 1 % a AO jsou minimálně 3 reads.
 - U SomaVar LF a SomaLBx je minimální alelická frekvence 0,1 % a minimální AO jsou 2 reads. Tato pracovní sada umožňuje detekci variant s frekvencí alel nižší než 1 %. Doporučuje se, pokud je hloubka sekvenování větší než 5000x, a pro takové aplikace, jako je sekvenování ctDNA.
2. **UMI disabled** (k dispozici pouze v SomaVar, SomaRNA, SomaHemato, SomaLBx, SomaCNVCapture, SomaCGP a SomaMSI): UMI se nepoužívají pro deduplikaci a před analýzou jsou z konce sekvencí oříznuty. Proto jsou data generovaná touto analytickou řadou surová (nededuplikovaná). Minimální alelická frekvence uváděných variant je 1 % a minimální počet reads podporujících mutaci (AO) je 3.

Rozdíly v nastavení těchto režimů jsou shrnuty v následující tabulce:

	UMI disabled	UMI standard
Minimální medián/průměrná kvalita mutované báze (skóre Phred)*	30	30
Minimální počet reads podle konsensu	N/A	2
Minimální kvalita každého základu v konsensu (skóre Phred)	N/A	30
Read without filtered consensus	N/A	no

Tabulka 1: Nastavení různých konfigurací zpracování UMI

* Tyto filtry se vztahují pouze na varianty, které nejsou vybrány modelem SVM po volání variant (pipeline SomaVar a SomaRNA).

Proces ošetření UMI

Struktura UMI závisí na jednotlivých dodavatelských sadách a v současné době jsou na platformě SeqOne k dispozici následující sady:

- QIAGEN QIAseq
- Agilent XTBS/Low input
- Agilent XTBS V2
- Illumina TruSight Oncology 500
- Twist UMI Adapter System
- Roche KAPA HyperCap
- QIAseq targeted DNA Pro

 Warning: Konfigurace UMI "QIAseq Targeted DNA Pro" by se měla používat pouze pro SomaVar s délkou čtení alespoň 150 bp. Pro více informací kontaktujte podporu.

Zpracování UMI je založeno na použití sady nástrojů fgbio, která dokáže zpracovávat UMI a reads označená UMI. Náš postup analýzy UMI se skládá z následujících kroků:

1. V případě potřeby ořízněte UMI na konci sekvencí a z páru souborů FASTQ vygenerujte nezarovnaný BAM,
2. Zarovnání sekvencí bez UMI pomocí BWA-MEM2 (DNA/ctDNA) nebo STAR (RNA)
3. Sloučení souborů BAM obsahujících UMI se zarovnanými sekvencemi pomocí ZipperBam,
4. Seskupte čtení, pokud mají stejnou pozici a sdílejí stejný UMI,
5. Vytvoření konsenzuální sekvence pro každou skupinu UMI,
6. Filtrujte každou bázi podle její kvality: důvěra v konsenzuální sekvenci se zvyšuje s počtem reads na skupinu,
7. Filtrování konsenzuálních sekvencí: pokud je v konsenzuální sekvenci filtrováno více než 20 % bází, konsenzuální sekvence se nezachová,
8. Převod těchto filtrovaných konsenzuálních sekvencí do fastq a konečné zarovnání na referenční genom pomocí BWA-MEM2 (DNA/ctDNA) nebo STAR (RNA).

Kompatibilita s analytickými worksety

SomaVar, SomaHemato, SomaCGP, SomaLBx, SomaMethyl, SomaVar LF, SomaCNVCapture, SomaMSI a SomaRNA Worksets jsou kompatibilní s daty UMI. Možnosti zpracování, které jsou k dispozici při spuštění jednotlivých worksetů, jsou popsány v tabulce níže:

	UMI disabled	UMI standard
SomaVar	Ano	Ano
SomaCNVCapture	Ano	Ano
SomaMSI	Ano	Ano
SomaRNA	Ano	Ano
SomaVar LF	Ne	Ano

SomaCGP	Ano	Ano
SomaLBx	Ano	Ano
SomaHemato	Ano	Ano
SomaMethyl	Ne	Ano

Tabulka 2: Možnosti zpracování UMI dostupné při spouštění kompatibilních pracovních sad.

11.2. Stav analýzy

Po spuštění analýzy používá platforma SeqOne různé stavy, které uživatelům umožňují sledovat průběh analýzy:

- **Bioinformatics Pending:** platforma čeká na ukončení další analýzy
- **Bioinformatics Running:** analýza probíhá; tento stav také ukazuje, v jaké fázi analýzy se nachází.
- **Bioinformatics Finished:** analýza byla dokončena
- **Analysis failed:** během analýzy došlo k chybě; v tomto případě jsou uživatelé vyzváni, aby kontaktovali tým uživatelské podpory.
- **Interpretation - new:** analýza je hotová a připravená k interpretaci, žádný uživatel ji zatím neotevřel.
- **Analysis failed - error:** Analýza narazila na chybu. V tomto případě je uživatel vyzván, aby kontaktoval "User support".
- **Analysis failed - canceled:** Analysis did not start because the sample attached to it encountered an error during loading or verification.

11.3. Kvalita vzorku

Karta kontroly kvality projektu je rozdělena na 2 dílčí karty: "General statistics" a "ID check summary".

General statistics. Na této kartě se zobrazí tabulka s celkovou kvalitou zarovnání a pokrytí pro každý vzorek projektu. Uživatel může tabulku exportovat jako soubor TSV kliknutím na .

Výpočet je založen na nástroji SeQC, což je program vyvinutý společností SeqOne pro výpočet statistik zarovnání a pokrytí ze všech čtení zarovnaných ke genomu.

Uživatel si může vybrat, zda chce použít výchozí metriky kontroly kvality, nebo definovat vlastní prahové hodnoty na kartě "Settings" projektu (viz část Konfigurace výchozích **metrik kontroly kvality projektu**).

Navrhované metriky jsou následující:

- celkový počet přečtení
- průměrná délka reads
- medián velikosti fragmentů DNA v knihovně (inzerty)
- procento kvalitních reads >30
- procento reads zarovnaných s referenčním genomem.
- procento cílových reads
- procento duplikovaných reads
- míra přechodu/převodu
- Průměrné pokrytí
- Podíl cílů pokrytých při různých prahových hodnotách v závislosti na kontextu (Amplicon, Capture).

Cíle panelu nebo exomu jsou určeny ze souboru BED vybraného při vytváření projektu.

ID check summary. Na této kartě jsou pro každý vzorek uvedeny genotypy všech markerů zahrnutých do jeho analýz (viz odstavec **Marker** v části **Správa uživatelských nastavení - definice**). Řádek odpovídá vzorku a každý sloupec představuje marker s přiřazeným ID. Uživatel může tuto tabulku exportovat jako soubor TSV kliknutím na .

12. Kontrola kvality dat

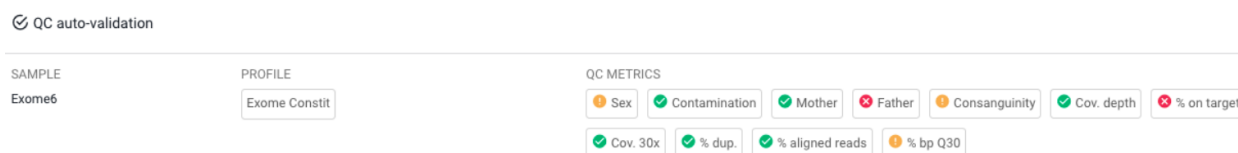
12.1. Metriky a zpráva o kontrole kvality

12.1.1. Metriky a kontroly kvality

Na kartě QC má uživatel přístup ke kontrolám kvality použitým na analýzu ve formě dílčích karet popsaných v níže uvedených oddílech: V těchto podkapitolách jsou uvedeny kontroly kvality: *QC report*, *ID check* (pokud je to relevantní) a *Technical report*.

V případě analýzy MicroArray CGH není karta QC k dispozici. Technická zpráva je však k dispozici na kartě *Files tab* v analýze *reportu* ve formátu *report.html*.

V horní části karty QC je oddíl "*QC auto-validation*", který uživateli umožňuje barevně označit metriky QC jako *Passed*, *Warning* nebo *Failed*. Kliknutím na přidruženou ikonu získá uživatel přístup k použitým prahům a podrobnostem QC: pohlaví, kontaminace, rodičovství a příbuzenské vztahy, hloubka pokrytí, % na cíli, 30X nebo 100X pokrytí, % duplikátů, % čtení zarovnaných s referenčním genomem nebo % bp Q30.



12.1.2. Zpráva QC

Globální *zpráva o kontrole kvality* pro proces sekvenování obsahuje různé metriky popisující kvalitu a zarovnání sekvencí, pokrytí oblastí v manifestu a některé statistiky týkající se variant. Tato zpráva se skládá z následujících oddílů:

Alignment details (Podrobnosti zarovnání) (SomaMethyl). Celkový počet zpracovaných čtení, unikátně a neunikátně zarovnaná čtení a míra neshod.

Deduplication (Deduplikace) (SomaMethyl). Metriky duplikace pro zarovnaná čtení, upozorňující na potenciální problémy s přípravou knihovny.

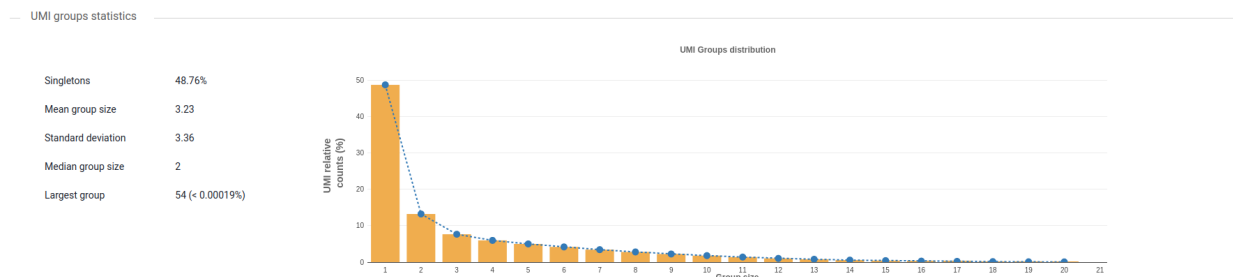
Cytosine coverage & methylation (Pokrytí cytosinu a methylace) (SomaMethyl). Míra konverze cytosinu na uracil pro každý kontext (CpG, CHG, CHH).

UMI groups statistics (pouze SomaVar, SomaVar LF, SomaMethyl a SomaRNA UMI). Viz odstavec **Volba režimu zpracování UMI**.

Zpráva o kontrole kvality analýz SomaVar, SomaVar LF, SomaMethyl a SomaRNA nabízí podrobnější pohled na složení vzorku, zejména na rozložení UMI podle počtu sekvencí, které je obsahují.

Statistiky popisují zejména:

- Podíl singletonů (UMI nesených jedinečnou sekvencí)
- Průměrná a mediánová velikost skupin UMI, stejně jako standard deviation.
- Velikost největší skupiny a podíl všech sekvencí, které zahrnuje.



Obrázek 7: Statistiky týkající se rozdělení UMI z hlediska počtu sekvencí ve vzorku.

Globální metriky kontroly kvality pro projekt jsou také podrobnější a rozdělené do tří oddílů:

- Metriky kontroly kvality po zpracování UMI, které udávají trend každé metriky (v procentech) na základě nezpracovaných údajů.
- Raw QC metrics vypočtené ze všech sekvencí (před zpracováním UMI).
- Rozdělení UMI z hlediska počtu sekvencí v každém vzorku.

Cross-contamination estimate (pouze SomaVar LF). Cross-contamination estimate se odhaduje individuálně pomocí nástroje GATK CalculateContamination. Ten využívá signál referenčních alel zjištěných v homozygotních lokusech k výpočtu podílu sekvencí v jedinci pocházejících z křížové kontaminace.

Sample integrity and identity (Germline analyses only). Tento oddíl uvádí metriky týkající se integrity a identity vzorku:

- **"Family link"** : metadata definovaná uživatelem
- **Original sex**: metadata týkající se pohlaví zadaná uživatelem (muž/žena/neznámý)
- **Zjištěné pohlaví**: zjištěné pohlaví (Muž/žena/neznámý)
- **Koeficient příbuznosti** : koeficient příbuznosti vypočtený pomocí nástroje Somalier pro každý vzorek v analýze rodiny, tato metrika se zobrazuje N/A pro samotný vzorek.
- **Contamination**: odhadované procento kontaminace vzorku cizí DNA pomocí nástroje verifyBamID2.

Dále je uveden počet homozygotních a heterozygotních variant, poměr těchto dvou hodnot (poměr Hom/Het), počet homozygotních a heterozygotních variant na chromozomu X, ekvivalentní poměr na chromozomu X (poměr Het/Hom na chr X) a počet variant na chromozomu Y.

Metriky *Family link* a *Relationship coefficient*, které slouží ke kontrole příbuznosti mezi jedinci v rodině, se zobrazují pouze v analýzách tří rodin (GermlineVar Family nebo GermVar Tertiary Family).

Metriky pohlaví se používají ke kontrole shody mezi zadaným pohlavím a pohlavím zjištěným v datech.

Metrika kontaminace se nepočítá pro worksety GermVar Tertiary a GermVar Tertiary Family.

Na úrovni projektu nebo analýzy jsou k dispozici *kontroly kvality* týkající se shody pohlaví, kontaminace a příbuznosti, které umožňují snadno identifikovat případné shody nebo neshody.

Alignment metrics. V této části jsou uvedeny hlavní metriky vypočtené na konci zarovnání, jako např.:

- Celkový počet sekvencí ve vzorku,
- Medián velikosti inzertů, vypočtený ze vzdálenosti mezi dvěma reads stejného páru, jejichž příslušná délka je zahrnuta,
- Průměrná délka sekvence,
- Podíl sekvencí úspěšně zarovnaných s referenčním genomem,
- Míra duplicit (pouze zachycení),
- Podíl bází, jejichž kvalita je vyšší než 30,
- Průměrné pokrytí cílových oblastí, stanovené na základě analýzy manifestu.

Coverage. Diagram znázorňující podíl cílových oblastí pokrytých při různých prahových hodnotách, vyznačených na ose x. Hodnoty pokrytí lze zobrazit v tabulkové formě kliknutím na tlačítko, které uživateli umožní tyto hodnoty zkopírovat/vložit.

QC graphs. Grafy z analýzy kvality FastQC, které znázorňují zejména obsah GC v sekvencích a kvalitu bází podle jejich umístění v sekvencích.

Variant statistics. Seznam statistických údajů o variantách zjištěných v analýze, které zahrnují zejména:

- Počet přechodů (Ts) a transversí (Tv) mezi zjištěnými substitucemi,
- Poměr Ts / Tv,
- Celkový počet zjištěných mutací,
- Podíl SNV, MNV a indelů na tomto celkovém počtu.

12.1.3. Zpráva ID-Check

Záložka "ID check" je věnována výsledkům sledování identity v záložce QC pracovních sad vyvolávajících varianty.

Výsledky se zobrazí ve formě tabulky s následujícími informacemi:

- Popis každého markeru na základě genomických souřadnic, například: chr-start-ref-alt,
- štítek (rsID, vlastní štítek), pokud jej zadá uživatel,
- sledovaný genotyp (GT) nebo frekvence alel (VAF),
- hloubka pokrytí v místě značky (COV),
- kvalitu databáze (QUAL).

Pokud je do analýzy zahrnuto více vzorků, sloupce pro jednotlivé vzorky (GT/VAF, QUAL, COV) se opakují pro každý vzorek zvlášť.

Výběrem varianty v tabulce se v nové záložce otevře prohlížeč IGV, který se soustředí na genomické souřadnice varianty.

12.1.4. Technická zpráva

V *technické zprávě* jsou uvedeny nástroje použité při analýze a jejich verze. Je přístupná na kartě QC analýzy a poté v části "Technical report".

Popisuje zejména :

- obecné informace o konfiguraci a metadatech projektu a analýzy, jako je typ sekvenování, verze genomu a verze použité pracovní sady;

Technical Report

This report describes the back-end process used for the secondary analysis.



General Infos

Project Name	VerificationTest - ID Marker Multiple samples
Experiment Type	dna-seq
Sequencing Type	paired-end
Sequencing Equipment	Illumina
Sequencing Method	panel
Amplification Method	capture
UMI Mode	standard
Genome Version	GRCh37
Analysis Name	SomaVar - toy_fusions48
Workset Name	SomaVar
Workset Description	Detection of somatic variants based only on the index sample
Workset Version	v2.5.11

- podrobnosti o použitém manifestu, jako je název, velikost, délka (bp) a výplň přidaná do analýzy;
- podrobnosti o použitých databázích a bioinformatických nástrojích.

12.1.5. Zpráva UPD

Zpráva UPD je k dispozici v sekci QC kompatibilních zárodečných analýz (např. GermVar a související pracovní postupy). Shrnuje výsledky detekce UPD a poskytuje vizuální prvky na podporu interpretace.

Zpráva UPD obvykle obsahuje:

- Přehled výsledků: indikace, zda analýza detekovala událost podobnou UPD a typ události (např. úplná nebo segmentální; v analýzách tria může zpráva také podpořit rodičovský původ a podtypy událostí, jako je UPD | Iso / UPD | Het).
- Seznam událostí: seznam detekovaných událostí s jejich chromozomálním umístěním a velikostí a přidruženým skóre, které používá algoritmus.
- Vizualizační grafy: grafy založené na genotypu SNP (např. frekvence B-alely (BAF) a signál homozygotnosti/ROH), které pomáhají potvrdit přítomnost dlouhých homozygotních oblastí a hranic.

Informace o kontrole kvality: metriky týkající se spolehlivosti detekce (např. hustota SNP / počet SNPs) a varování, pokud data nejsou dostatečná pro spolehlivé volání.

Detekované události UPD se také zobrazují v Large Variant Viewer (LVV) se specifickými štítky (např. UPD | Iso a UPD | Het).

Skóre UPD

Skóre UPD je numerický indikátor používaný ke kvantifikaci toho, jak silně vzorek vykazuje signál podobný UPD podél genomu. Detekce UPD v SeqOne spoléhá na vzorec genotypu SNP:

- V singletonové (sólo) analýze algoritmus hledá velké oblasti homozygotnosti (ROH) a náhlé změny v distribuci SNPs podél chromozomů.
- V analýze tria se vzorec SNP navíc používá k určení rodičovského původu události a k rozlišení izodisomie (UPD | Iso) od heterodisomie (UPD | Het).

V praxi skóre UPD odráží sílu a konzistenci signálu podporujícího událost UPD (např. jasný vzorec ROH/BAF a dobře ohraničené hranice). Vyšší skóre naznačuje, že událost je s větší pravděpodobností skutečnou událostí UPD, zatímco nižší hodnoty naznačují slabší nebo hlučnější signály.

Skóre UPD musí být vždy interpretováno společně s prvky uvedenými na kartě Zpráva UPD (sekce QC), včetně:

- vizualizačních grafů UPD (např. frekvence B-alely / signál homozygotnosti),
- velikosti a chromozomálního umístění události,
- metrik kvality analýzy (např. hustota SNP / počet SNPs).

Graf frekvence B-alely (BAF) ze zprávy UPD

Graf frekvence B-alely (BAF) je vizualizace genotypů SNP podél genomu a používá se k hodnocení alelické nerovnováhy a ztráty heterozygotnosti (LOH).

Pro každou pozici SNP odpovídá BAF podíl sekvenačních čtení podporujících alelu "B" a pohybuje se od 0 do 1:

- Hodnoty blízké 0 odpovídají homozygotnímu genotypu (AA).
- Hodnoty blízké 0,5 odpovídají heterozygotnímu genotypu (AB).
- Hodnoty blízké 1 odpovídají homozygotnímu genotypu (BB).
-

V normální diploidní oblasti graf BAF obvykle ukazuje tři pásy kolem 0, 0,5 a 1. V dlouhých oblastech homozygotnosti (ROH) nebo isodisomie UPD je heterozygotní pás (kolem 0,5) snížen nebo chybí, což podporuje přítomnost signálu podobného UPD.

Důležitá omezení:

- Signál podobný UPD lze také pozorovat v kontextu příbuznosti rodičů (více oblastí ROH napříč genomem).
- Volání UPD je založeno především na distribuci SNPs a nemusí samo o sobě stačit k odlišení některých vzorců UPD od událostí souvisejících s počtem kopií; v případě potřeby proveďte křížovou kontrolu s dalšími dostupnými důkazy (např. pokrytí / výsledky CNV).

13. Interpretace výsledků analýzy

13.1. Výsledky analýzy

Uživatelé mohou získat přístup k analýzám projektu kliknutím na *ikonu* "Analyses" v přidruženém poli nebo kliknutím na stejnojmennou ikonu v levém horním rohu uživatelského panelu.

Pokud byla analýza dokončena a je k dispozici k prohlížení, zobrazí se se stavem "*Interpretation - new*".

Výsledky analýzy jsou pak k dispozici v podobě několika sestav, které lze otevřít přímo kliknutím na ně.

Existuje několik variant stavu "*Interpretace*" v závislosti na dosaženém pokroku při interpretaci výsledků:

- **New** - tabulka variant ještě nebyla zobrazena
- **In progress** - tabulka variant byla zobrazena, ale zatím nebyly ověřeny žádné varianty.
- **Evaluating** - alespoň jedna varianta byla ověřena
- **Reporting validated** - pro analýzu byla vygenerována zpráva.

13.1.1. Správa pracovního postupu interpretace analýzy

13.1.1.1. Zadání analýz pro interpretaci

Správci entit mohou na kartě Analýzy projektu přiřadit jednu nebo více analýz, které mají být interpretovány různými uživateli entit. To do this, the administrator can select one or more

analyses and then click on the button  Edit assignment at the bottom of the page.

Okno "Assignment" umožňuje přiřadit jednoho nebo více uživatelů k vybraným analýzám, a to vyhledáním identifikátoru uživatele (uživatelského jména) pomocí vyhrazeného vyhledávacího panelu. Přiřazení je také možné upravit přidáním nebo odstraněním uživatelů. Uživatelé, kterým je analýza přiřazena, jsou uvedeni ve sloupci "Assigned To." (Přiřazeno). Uživatelé, kteří se nikdy nepřihlásili do SeqOne, nemohou být přiřazeni, dokud si nevytvoří své přihlašovací údaje.

Uživatelé mají přístup k analýzám, které jim byly přiděleny (záložka To do), a k ověřeným analýzám (záložka Done) ve svém seznamu To do, který je přístupný prostřednictvím ikony  v pravém horním rohu stránky. U každého nového přiřazení se zobrazí oznámení.

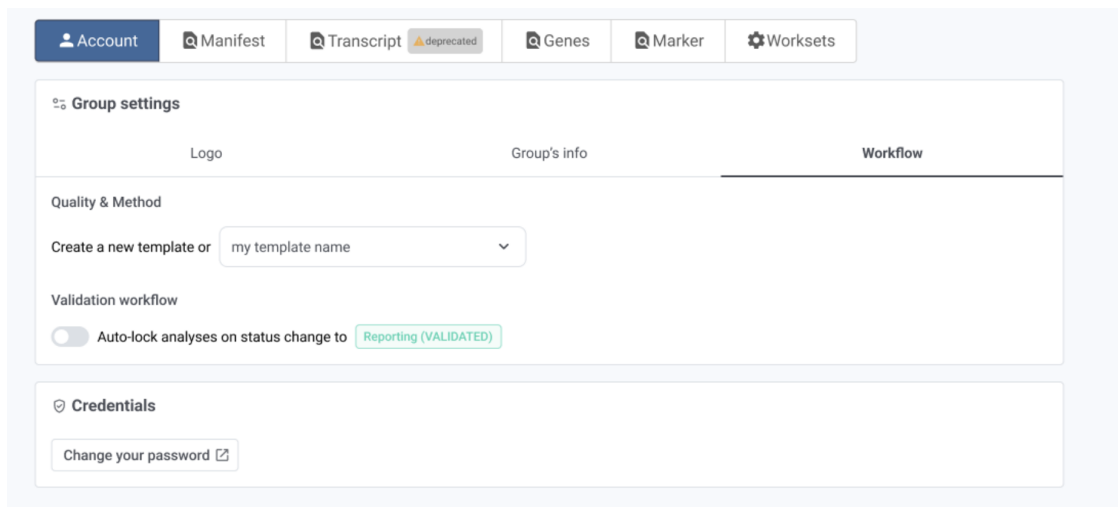
Uživatelé mohou k přiřazeným analýzám přistupovat přímo kliknutím na ně. Svůj úkol mohou potvrdit kliknutím na příslušné zaškrtnávací políčko na kartě "K provedení" a přesunout jej na kartu "Done". Tato akce je vratná.

To do list		
To do	Done	
Assigned 10 days ago		
 somaHRD - sample_pos_with_BRCA TEST_HRD_sWGS_UMI_xths2		☐
Assigned 7 days ago		
 somaHRD - sample_pos_with_BRCA TEST_HRD_sWGS_UMI_xths2		☐

Analýzy se zobrazují v následujícím pořadí podle svého stavu: dokončené, spuštěné nebo neúspěšné. U každé analýzy je podrobně uvedeno datum zadání, název analýzy a název projektu.

13.1.1.2. Analýza uzamčení

Správce subjektu může analýzu uzamknout nebo odemknout. V nastavení entity může na kartě "Account" a poté "Workflow" aktivovat automatické zamykání validovaných analýz pomocí funkce "Auto-lock analyses on status change to Reporting (VALIDATED)":



Když uživatel na konci své interpretační práce změní stav analýzy v části "Reporting - validated", analýza se automaticky uzamkne. U uzamčených analýz se zobrazuje stav "Reporting - Validated & Locked" (Hlášení - validováno a uzamčeno), který je viditelný pomocí ikony .

 Reporting VALIDATED & LOCKED

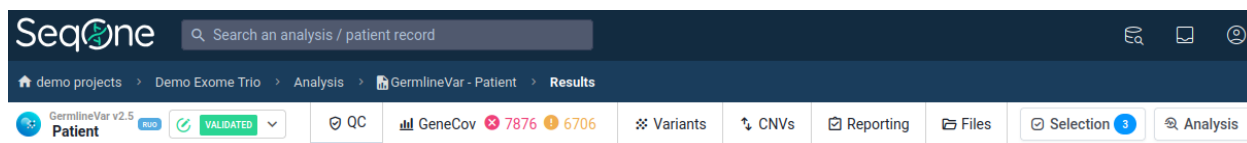
Pokud je analýza uzamčena, uživatelé nemohou z rozhraní provádět následující akce :

- Add or modify variant evaluations in VKB,
- Add nebo upravit komentáře související s variantami,
- vygenerovat novou zprávu z okna *Selection*,
- vygenerovat novou zprávu HRD,
- ručně upravit stav analýzy,
- proveďte jednu z následujících akcí na panelu analýzy:
 - upravit přiřazení testu (v části "Přiřazení").
 - Modifier pohlaví pacienta
 - přidat nebo upravit HPOs
 - Modifier (
 - smazat klinickou zprávu

Pokud se uživatel pokusí o některou z těchto akcí, bude vyzván, aby analýzu odemkl, pokud k tomu má oprávnění, nebo aby kontaktoval svého správce, pokud oprávnění nemá. Uživatel, který chce analýzu odemknout, musí mít oprávnění "Unlock analyses".

13.1.1.3. Boční panel pro správu analýzy

Boční panel analýzy alias "Analysis Drawer" je přístupný z kterékoli karty analýzy.



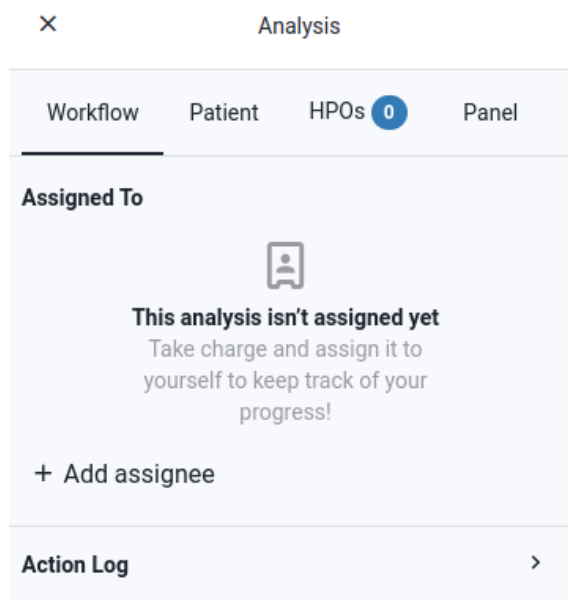
Kliknutím na tlačítko  Analysis v pravém horním rohu v banneru karet analýzy může uživatel přímo zobrazit a spravovat data analýzy a HPO pacienta, aniž by opustil aktuální obrazovku.

"Analysis Drawer" obsahuje 4 zobrazení:

- **Workflow:** umožňuje organizovat a sledovat pracovní postup pro analýzu.
- **Patient:** umožňuje zobrazit a upravit metadata pacienta.
- **HPOs:** umožňuje přidat nebo upravit termíny HPO v analýze, dostupné pouze pro analýzy v ústavním kontextu.
- **Panel:** umožňuje použít seznam genů pro analýzu a filtrovat výsledky analýzy na základě tohoto seznamu.

Modální zobrazení Workflow

Toto zobrazení umožňuje organizovat a sledovat pracovní postupy uživatelů pro interpretovanou analýzu.



Assigned To. Toto zobrazení zobrazuje úkoly přiřazené pro aktuální analýzu. Úlohy mohou přiřazovat pouze správci nebo uživatelé s oprávněním "Manage assignments" (viz **část Správa správců - prezentace rozhraní**).

Po přiřazení úkolů k aktuální analýze se zobrazí seznam všech přiřazených uživatelů spolu s datem a časem dokončení. Tento seznam se automaticky aktualizuje, jakmile uživatel úkol dokončí.

Uživatelé mohou upravovat pouze své vlastní úkoly zaškrtnutím nebo zrušením zaškrtnutí příslušného políčka. Po najetí na úkol určený pouze pro čtení se zobrazí nápověda "Můžete dokončit pouze svůj".

Pokud nebyly přiřazeny žádné úkoly, zobrazí se v zobrazení zpráva "Tato analýza ještě není přiřazena".

Pokud se neautorizovaný uživatel pokusí přiřadit úkol kliknutím na tlačítko "+ Add assignee", zobrazí se chybová zpráva "K přiřazení analýzy musíte mít oprávnění správce. V případě potřeby se obraťte na správce platformy a požádejte o příslušný přístup."

Action Log. V banneru "Analysis Action Log (Audit Trail)" se zobrazuje historie všech akcí provedených na stavu HRD v analýze SomaHRD, filtrech, profilech filtrů a panelech omezených genů v tabulce variant s uvedením typu akce, uživatele, časového razítka a popisu akce.

Nejnovější akce se zobrazují v horní části seznamu. Stránkování umožňuje uživatelům efektivně procházet historii akcí. Celou historii lze stáhnout ve formátu .csv pomocí ikony "Download Log".

Modální zobrazení Patient

Toto zobrazení umožňuje zobrazit a upravit metadata pacienta, například pohlaví (Female, Male nebo Unknown), genetická skupina předků GnomAD (pouze Germline) nebo indikace (pouze Somatic).

Zobrazení a úprava indikace je k dispozici pouze pro somatické analýzy:

- Pokud byl údaj zadán před spuštěním analýzy (viz **Metadata vzorku**), je pro vzorek definován jako výchozí, je pak označen hvězdičkou a v seznamu se zobrazuje na prvním místě.
- Pokud je indikace během interpretace změněna, zobrazí se varovné hlášení, že vybraná indikace se liší od původní indikace, a podbarvení tablet v podsloupci "A" ve *sloupci Clinical Evidences* bude aktualizováno (viz "Selection").
- Pokud není k dispozici žádná indikace pro analýzu, zobrazí se zpráva "No available indication".

Analysis

Workflow **Patient** HPOs Panel

Active sample: **sample2**

Sex

Female Male Unknown

Indication

Ovarian endometrial cancer (Female reprod... ×

Unsaved changes detected. Save now to avoid losing your updates.

Cancel Save modifications

Zobrazení a úprava "GnomAD genetic ancestry group" je k dispozici pouze pro analýzy **germline**. Výchozí hodnota je prázdná a je považována za kombinovanou populaci GnomAD.

Změny provedené v tomto zobrazení je třeba uložit nebo zrušit kliknutím na příslušné ikony "Save modifications/Cancel" v dolní části panelu, aby bylo možné změnit "Tabs".

GnomAD genetic ancestry group

African/African American

Amish

Ashkenazi Jewish

Admixed American

East Asian

European (Finnish)

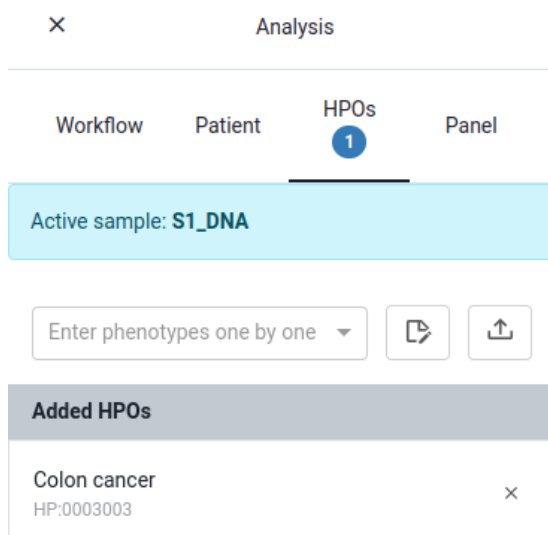
Modální zobrazení HPO

Toto zobrazení, které je k dispozici pouze v projektech Germline, umožňuje zobrazit, upravit a přiřadit termíny HPO týkající se vzorku, a to buď ručně, nebo z klinického popisu pacienta.

Uživatel si může zobrazit seznam termínů HPOs, které již byly se vzorkem spojeny před analýzou. V případě nepřítomnosti HPOs kódů se zobrazí zpráva "No HPOs yet" (Zatím žádné HPOs).

Pokud chce uživatel přidat termíny ručně, může je zadat z vyhledávacího panelu "Enter phenotypes one by one" (Zadat fenotypy po jednom) vyhledáním fenotypu nebo HPO kódu.

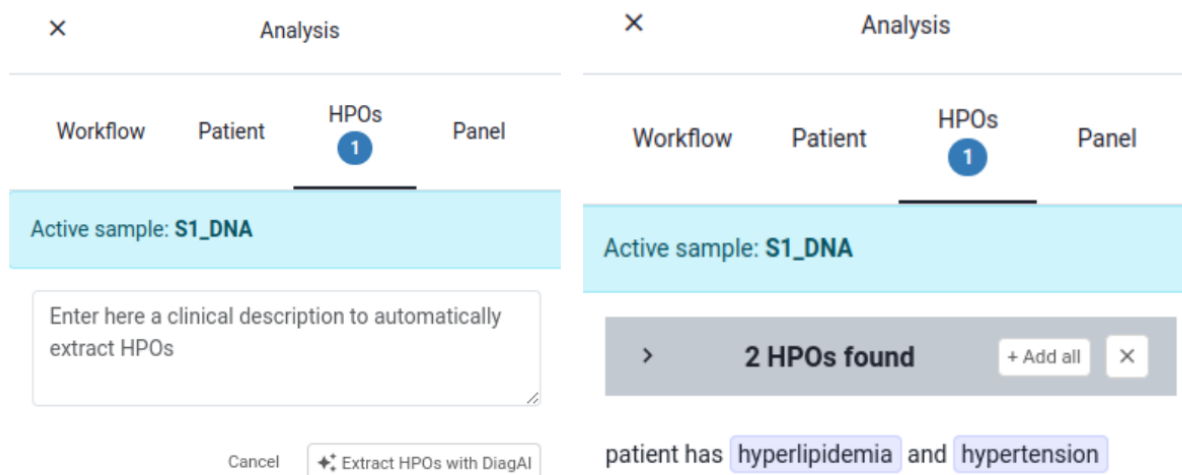
Může také importovat textový soubor obsahující HPO kódy kliknutím na ikonu  (viz také oddíl **Metadata vzorku**).



Obrázek 8: Pohled na panel analýzy HPOs

Uživatel může generovat termíny HPO na základě klinického popisu pacienta pomocí nástroje

DiagAI. Kliknutím na ikonu , se zobrazí dialogové okno, které uživatele vyzve k popisu klinického profilu pacienta (tento popis je omezen na 4000 znaků) a poté kliknutím na "Extract HPOs with DiagAI" (Odvodit HPO pomocí DiagAI) vygeneruje seznam nalezených HPO. Termíny HPO jsou barevně zvýrazněny a uživatel je může přiřadit k pacientovi tak, že na ně najede myší a klikne na "Add" (Přidat), nebo přidá všechny nalezené termíny HPO kliknutím na "Add all" (viz také část Editace podle metadat vzorku **s panelem analýzy**).

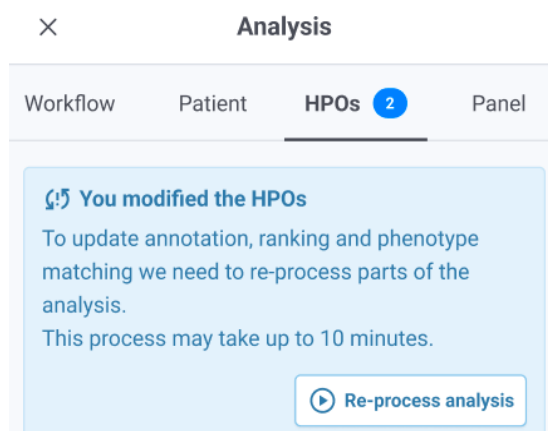


Obrázek 9: Extrakce termínů HPO pomocí nástroje DiagAI ze zobrazení HPO v "Analysis Drawer"

Pokud jsou v tomto zobrazení přidány nebo změněny HPO, zobrazí se modrý banner, který navrhuje, aby byla analýza znovu zpracována s novými HPO.

Výběrem možnosti "Re-process analysis" může uživatel aktualizovat skóre DiagAI a klasifikaci variant (Short list nebo SmartPick) podle nově zadaných HPO.

Tento krok opětovného zpracování na několik minut zmrazí tabulku variant, aby bylo možné aktualizovat klasifikaci variant a sloupec DiagAI.

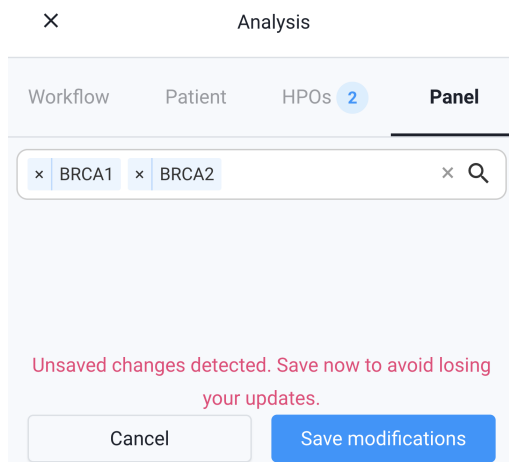


Modální zobrazení Panel

Toto zobrazení je k dispozici pro všechny worksety kromě worksetů pro analýzu CNV (CNVExome, CNVCapture a SomaCNVCapture).

Umožňuje přístup k funkci "Restricted panels", která umožňuje omezit a uzamknout analýzu na vybrané geny na způsob panelu in silico.

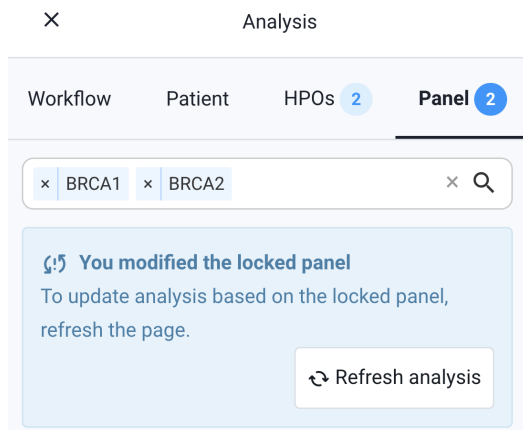
Za tímto účelem může uživatel prostřednictvím rozbalovací nabídky vybrat jeden nebo více seznamů genů, které se mají použít, z dříve importovaných (viz odstavec Gene list, v části **Správa uživatelských nastavení**). Uložení vybraných panelů prostřednictvím "Save modifications" (Uložit úpravy) se na celou analýzu použije filtr podle vybraného panelu (vybraných panelů) genů.



Restricted panels" lze vybrat ihned po spuštění analýzy (viz část **Spuštění nové analýzy**) nebo v zobrazení Panel po analýze.

Při přidání nebo úpravě genových panelů upozorní uživatele na změnu modrý banner ("You modified the locked panel") a nabídne obnovení analýzy s novými panely pomocí funkce "Refresh analysis".

Kliknutím na tlačítko "Refresh analysis" si uživatel zvolí aktualizaci tabulky variant, záložky CNV (pokud je k dispozici) a záložky GeneCov podle nového panelu.



13.1.1.4. Boční panel pro správu hodnocení

Panel  **Selection** je přístupný z kterékoli karty v rámci analýzy. Umožňuje uživateli zobrazit malé varianty a CNV anotované ve VKB pro aktuální analýzu, vygenerovat soubor JSON ("Generate file") a upravit klinickou zprávu ("Generate report").

CNV validované ve VKB jsou rovněž k dispozici v této sekci, pouze v případě CNV z analýz exomu, genomu a CGH Microarray.

Malé varianty validované ve VKB jsou odlišeny logem , zatímco CNV jsou zobrazeny logem  a slovy "gain" nebo "loss".

×

Selection 3

↓ Generate file

☑ Generate report

Variants

❖ SLC25A22

5 NM_001191061.2 | chr11:g.792146_792147del

The frameshift variant c.813_814del (p.Ala272GlnfsTer144) in the gene SLC25A22 (NM_001191061.2) is not found in control...

❖ SLC25A22

5 NM_001191061.2 | chr11:g.792142C>T

The missense variant c.818G>A (p.(Arg273Lys)) in the gene SLC25A22 (NM_001191061.2) is extremely rare, with onl...

CNVs

↕ 6q14.3 Loss

4 103.25kb | SNX14, SYNCRIP

13.1.2. Přehled výsledků analýzy na kartě Dashboard

U některých worksetů je uživatelům přístupná karta "Dashboard", která shrnuje všechny výsledky a metriky kontroly kvality a která se ve výchozím nastavení zobrazuje při otevření analýzy. Karta "Dashboard" poskytuje přehled výsledků analýzy v několika sekcích. Dostupné oddíly se liší v závislosti na typu pracovní sady. Tato karta je k dispozici pro SomaCGP, SomaVar Tertiary a SomaLBx.

Patient. Tato část shrnuje metadata vzorku: název vzorku (Sample ID), *pohlaví* a patologickou indikaci (Indication).

Quality Controls. Tato část obsahuje informace o několika klíčových ukazatelích týkajících se kvality vzorků a sekvenování:

- Čistota nádoru,
- Vzorek ploidie,
- procento bází, jejichž kvalita je větší nebo rovna 30 (% bází Q30),
- průměrné pokrytí zájmových oblastí (průměrné pokrytí ROI) před a po zpracování UMI,
- procento oblastí zájmu s pokrytím větším než 200X (ROI > 200X) před a po zpracování UMI.

Signatures. V části věnované signaturám jsou uvedeny klíčové údaje pro analýzy TMB, MSI a HRD:

- Výpočet TMB (Tumor Mutational Burden): počet mutací/Mb.
- Microsatellite instability: procento nestabilních míst.
- HRD status (Homologous Recombination Deficiency).

Stav skóre TMB a MSI si může uživatel vybrat pomocí rozevíracích nabídek umístěných pod každým skóre.

Selection. V této části jsou zobrazeny varianty hodnocené v databázi VKB pro aktuální analýzu.

Suggested by SeqOne. Genomové změny zjištěné při analýze jsou prezentovány jako redukovaný seznam klinicky relevantních variant. Variant: je vybrán, pokud je jejich třída ComPerMed 5, nebo pokud mají tvrzení úrovně I odpovídající indikaci pacienta, nebo tvrzení úrovně II pro experimentální terapie související s touto indikací, nebo tvrzení úrovně I pro jinou indikaci. Uživatel může přidat varianty z této kategorie do sekce Selection kliknutím na tlačítko "Add".

Soubor JSON lze vygenerovat kliknutím na "Generate JSON file" a obsahuje všechny informace ze "Dashboard" stránky.

13.1.3. Podrobné pokrytí podle exonů

Na kartě GeneCov mohou uživatelé zobrazit míru pokrytí genů zahrnutých do analytického manifestu pro každý exon, ať už se jedná o panel genů nebo exom.

Metodika

Tento modul protne uživatelský manifest s anotacemi RefSeq pro všechny geny zahrnuté do panelu nebo exomu, přičemž na každé straně ponechá 10bp, aby pokryl splice junctions, a poté vypočítá hloubku pokrytí pro každou bázi zahrnutou do tohoto průsečíku pomocí nástroje mosdepth.

Jsou nastaveny různé prahové hodnoty, aby bylo možné každému exonu přiřadit stav (pokrytý, varování, neúspěšný) podle procenta bází pod těmito prahovými hodnotami pokrytí.

Status: je definován na základě Coverage v exonu včetně přilehlých 10bp.

Pokud jsou použity UMI, vypočítá se pokrytí po deduplikaci.

Prahové hodnoty platné pro různé typy analýz jsou podrobně uvedeny v následující tabulce*:

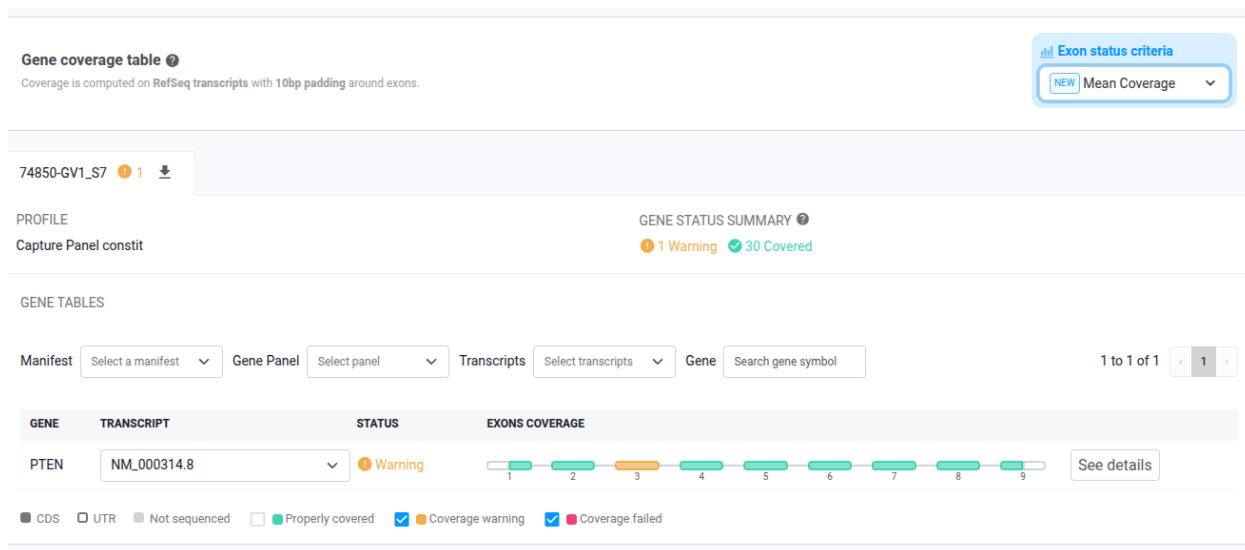
	GermlineVar / Germline Family/ GermVar / GermVar Family				SomaVar		SomaRNA		SomaVar LF		SomaCGP	
	Panel		Exome		Panel		Panel		Panel		Panel	
	Warning	Failed	Warning	Failed	Warning	Failed	Warning	Failed	Warning	Failed	Warning	Failed
Capture	Cov. 100x < 100%	Cov. 30x < 100%	Cov. 30x < 100%	Cov. 10x < 100%	Cov. 500x <100%	Cov. 200x <100%	Cov. 200x <100%	Cov. 50x <100%	Cov. 5000x <100%	Cov. 2500x <100%	Cov. 100x < 100%	Cov. 50x < 100%
Amplicon	Cov. 100x < 100%	Cov. 30x < 100%			Cov. 1000x < 100%	Cov. 500x < 100%	Cov. 200x <100%	Cov. 50x <100%				

Tabulka 3: Prahové hodnoty pokrytí pro modul Coverage podle typu analýzy.

*Pokud uživatel předdefinoval vlastní metriky QC z parametrů projektu, budou použity na kartě GeneCov (viz část Konfigurace **výchozích metrik QC projektu**).

Zobrazení a výklad

Každý gen lze vizualizovat v grafickém znázornění a výsledky z nástroje Gene Coverage lze filtrovat na základě genů, in silico panelů a seznamů transkriptů. Cílené exony jsou zobrazeny buď zeleně (pokryto), oranžově (varování) nebo červeně (Failed) podle výše popsaných prahových hodnot.



Obrázek 10: Podrobnosti karty GeneCov

Možnosti filtrování

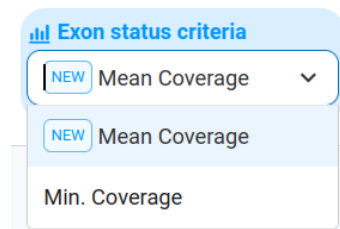
Nabídka filtrů umožňuje:

- omezit zobrazení na podmnožinu oblastí pomocí souboru .bed, který byl dříve importován do rozhraní, pomocí filtru "Manifest",
- použít seznam genů pomocí filtru "Gene panel",
- použít seznam prepisů prostřednictvím filtru "Transcripts",
- vyhledat konkrétní gen pomocí filtru "Gene".

Metódy výpočtu Coverage.

Uživatel může vybrat metodu výpočtu pokrytí, která se použije, pomocí rozevírací nabídky "Exon status criteria" v pravém horním rohu stránky:

1. **Mean Coverage** (výchozí): K určení stavu pokrytí se používá průměrná hloubka pokrytí celého exonu.
2. **Minimum Coverage** : Používá minimální podíl bází, které splňují určitý práh pokrytí. Tím je zajištěno, že určité procento bází má alespoň minimální požadovanou hloubku pokrytí.



Rozbalovací nabídka "**Gene status criteria**" umožňuje uživatelům použít tyto metody výpočtu na úrovni genů:

- **Mean coverage (default):** K určení stavu pokrytí používá průměrnou hloubku pokrytí u všech exonů genu.
- **Worst exon status:** Určuje stav pokrytí genu na základě nejhůře fungujícího exonu. Pokud kterýkoli exon nesplní práh pokrytí, gen je považován za nepokrytý. V předchozích verzích to byla jediná podporovaná metoda.

Zobrazená data

Exony, které nejsou nebo jsou částečně pokryty návrhem (< 80 % překrytí mezi manifestním intervalem a hranicemi exonu RefSeq), jsou považovány za necílené. Tyto exony jsou v grafu znázorněny šedě (*Not sequenced*).

Podrobné zobrazení každého genu lze použít k zobrazení metrik pokrytí exonů kliknutím na "See details": průměrné pokrytí, medián, procento každé oblasti pokryté 10x, 30x, 50x, 100x, 200x, 300x a 500x.

Okno IGV umožňuje pro každý exon podrobněji zobrazit pokrytí sekvence a zarovnání.

Export dat GeneCov

Kompletní data analýzy pokrytí jsou k dispozici jako soubor `sample_genecov.v2.tsv.gz`, který je přístupný na kartě Soubory analýzy.

Na kartě GeneCov může uživatel exportovat analýzu pokrytí podle metody výpočtu pokrytí (Mean Coverage nebo Min Coverage)

prostřednictvím ikony  vedle názvu vzorku v levém horním rohu karty vzorku.

Download genes from S1_DNA as .tsv file ×

Genes status

☒ Failed ☒ Warnings ☐ Covered

Transcripts

☒ Canonical ☐ All

Close Download

Uživatel si může vybrat, zda chce exportovat data podle stavu pokrytí genů (Failed, Warning, Covered) a zda chce exportovat pouze kanonický transkript (Canonical) nebo všechny dostupné transkripty (All), než klikne na Download, aby získal požadovaný soubor.

Data se exportují jako soubor .tsv obsahující následující sloupce:

- Genes (gene_symbol),
- Číslo genu podle HGNC (gene_id),

- Kanonický přepis (is_canonical_transcript),
- Transcript: Status (transcript_status)
- Exon (exon_number)
- Type exonu CDS nebo UTR (exon_type)
- Stav pokrytého nebo nepokrytého exonu (exon_status)
- Region, na který se manifest vztahuje, nebo ne (cíl)
- Genomická pozice (interval_id)
- Číslo chromozomu (Chr)
- Výchozí pozice (start)
- Koncová poloha (konec)
- Forward nebo reverzní vlákno (pramen)
- % pokrytí při 1, 10, 20, 30, 50, 100, 200, 300, 500 a 1000x (cov_)
- Průměrné pokrytí (cov_avg)
- Medián Coverage (cov_med)

13.1.4. Malé varianty

Metodika

Detekce malých variant se provádí kombinací několika nástrojů, které se mohou v jednotlivých worksetech lišit.

Ať už jde o analýzu exomu nebo genového panelu, je volání variant určeno manifestem (ve formátu BED) přidruženým k projektu, včetně intervalů oblastí zájmu a výplně aplikované na tyto oblasti (viz část Create project).

Kromě toho jsou varianty volány s ohledem na kvalitu zarovnání reads v lidském genomu (verze GRCh37 nebo GRCh38). Takzvaná čtení "MAPQ0" s kvalitou zarovnání (mapování) rovnou 0, protože se zarovnávají na několika místech genomu, se při vyvolávání variant pro pracovní soubory somatické analýzy neberou v úvahu. V analýzách "Germline" jsou čtení MAPQ0 brána v úvahu pro vyvolání variant, pokud je při vytváření projektu v metadatech Low-Quality Mapping Region (Oblast mapování s nízkou kvalitou) zadán manifest specifikující oblasti zarovnání s nízkou kvalitou. Varianty zjištěné ze čtení MAPQ0 jsou v tabulce variant označeny nápoředou "_mapq0" viditelnou ve sloupci CAL (Caller).

SNVs / indels. Pro detekci SNV a indelů :

- *Freebayes*¹ je používán jako jediný volací kód variant v pracovních sadách GermlineVar (exom), GermlineFamily (exom), SomaRNA a SomaVar LF. Dolní detekční prahy jsou následující:
 - VAF $\geq 10\%$ v GermlineVar pro exom,
 - VAF $\geq 5\%$ v GermlineVar pro genového panelu,
 - VAF $\geq 1\%$ v SomaRNA,
 - VAF $\geq 0,1\%$ v přípravku SomaVar LF,
- Je doplněn nástrojem *GATK HaplotypeCaller*² v GermlineVar - Panel workset (panel).
- Jeho somatický protějšek, *Mutect2*³, se používá spolu s Freebayes v pracovních sadách SomaVar, SomaHemato a SomaCGP. Minimální práh detekce je u těchto dvou volajících v systému SomaVar VAF $\geq 1\%$.
- *Sentieon*⁴ je používán jako hlavní volací kód variant v pracovních sadách GermVar, GermVar Family a SomaLBx.
- *Clair3*⁵ se používá jako jediný kód volání variant v workset GermlineVar LR.

¹ <https://github.com/freebayes/freebayes>

² <https://gatk.broadinstitute.org/hc/en-us/articles/360037225632-HaplotypeCaller>

³ <https://gatk.broadinstitute.org/hc/en-us/articles/360037593851-Mutect2>

⁴ <https://github.com/Sentieon>

⁵ <https://github.com/HKU-BAL/Clair3>

Analýzy GermVar a GermVar Family spoléhají na bioinformatické komponenty Sentieon pro klíčové kroky sekundární analýzy. Zejména:

- Zarovnání se provádí pomocí optimalizované implementace BWA-MEM2 od společnosti Sentieon a duplicitní čtení jsou označena.
- Zárodečné volání (SNVs/Indels) se provádí pomocí Sentieon DNAScope⁶.
- Sentieon TNScope⁷ lze použít k detekci mozaikových variant s nízkým VAF (v závislosti na konfiguraci GermVar).

Poznámka k regionům s nízkou mapovatelností (MAPQ0): ve vysoce homologních regionech (např. SMN1/SMN2) mohou být varianty odfiltrovány filtrováním strojového učení vyvolávající. V případě potřeby poskytuje SeqOne volitelný modul MAPQ0 (modul regionů s nízkou kvalitou) založený na Freebayes s přizpůsobenými prahy pro zvýšení citlivosti v uživatelem definovaných regionech.

Analýzy SomaLBx spoléhají na bioinformatické komponenty Sentieon pro klíčové kroky sekundární analýzy tekuté biopsie. Zejména:

- Zarovnání čtení se provádí pomocí Sentieon (BWA-MEM) a pipeline zahrnuje zpracování/konsenzus UMI založené na Sentieon pro podporu vysoce citlivých pracovních postupů ctDNA.
- Somatické volání krátkých variant (SNVs/Indels) se provádí pomocí Sentieon TNScope⁷ v režimu pouze pro nádor (tumor-only), s podporou velmi nízké alelové frakce (VAF $\geq 0,1$ % v závislosti na konfiguraci).
- Poznámka: genotypy variant jsou ve výsledcích SNV/indel SomaLBx standardně hlášeny jako heterozygotní, protože Sentieon TNScope neprovádí predikci genotypu.

SVs - Structural Variants. Inzerce a delece, jejichž délka je blízká nebo větší než délka sekvenčního readu, jsou detekovány pomocí SV calleru GRIDSS⁸ v pracovních sadách GermlineVar (panel/exom), SomaVar, SomaVar LF SomaHemato a SomaCGP worksets. Tyto události jsou detekovány pomocí SV caller Delly⁹ v pracovních GermVar a GermVar Family worksets. V somatických worksets jsou detekovány pouze inzerce a delece menší než 300bp, zatímco v pracovní sadě GermlineVar (pro verze ≥ 2.0) GermVar a GermVar Family worksets (versions ≥ 3.5) je použit limit 10kb. Varianty detekované pomocí GRIDSS jsou detekovány bez ohledu na VAF.

FLT3-ITD. Specifický detekční modul pro interní tandemové duplikace genu FLT3 (FLT3-ITD) je k dispozici pouze ve worksets SomaVar a SomaHemato, od verze v2.2. Tento modul je založen na nástroji *FILT3r*¹⁰.

⁶ <https://github.com/Sentieon/sentieon-dnascopy-ml>

⁷ <https://github.com/Genomic-Medicine-Linkoping/TNScope>

⁸ <https://github.com/PapenfussLab/gridss>

⁹ <https://github.com/dellytools/delly>

¹⁰ <https://gitlab.univ-lille.fr/filt3r/filt3r>

Alu Insertions. Společnost SeqOne vyvinula také modul pro detekci inserce transponovatelných elementů (AluMEI), jako jsou Alu elementy. Využívají ho pracovní sady GermlineVar (panel/exom), GermlineFamily (panel/exom), GermVar (panel/exom/genom), GermVar Family (panel/exom/genom), SomaHemato a SomaVar. Pro varianty detekované pomocí AluMEI se nepoužívá žádný spodní práh VAF.

Kombinované varianty *Freebayes*, *HaplotypeCaller* (nebo *Mutect2*) a *FiLT3r* caller jsou k dispozici v souboru *final.vcf.gz* analýzy. Varianty specifické pro *GRIDSS*, *FiLT3r* nebo *AluMEI* jsou k dispozici v příslušných souborech VCF (*gridss.vcf.gz*, *filt3r.vcf.gz* a *variants.alumei.vcf.gz*), **které jsou rovněž k dispozici ke stažení na kartě Soubory analýzy.**

V konkrétním případě *FiLT3r* :

- Varianty *FiLT3r* společné pro *Freebayes* a *Mutect2* jsou k dispozici ve finálním VCF analýzy (*final.vcf.gz*),
- varianty jedinečné pro *FiLT3r* a varianty společné pro *FiLT3r* a *GRIDSS* jsou uvedeny ve *FiLT3r* VCF (*filt3r.vcf.gz*).

Dva další moduly vyvinuté společností SeqOne pak umožňují :

- Zachování vícenukleotidových variant neboli MNV, které jsou kódované a omezené na jeden kodon během procesu rozkladu variant podobných delinám.
- Rekalibrace alelické frekvence variant indelů a delinů zjištěných pomocí *Freebayes* a *GATK* (nástroj Kraf).

Tyto dva moduly se týkají pouze pracovních sad GermlineVar (panel/exom), SomaVar, SomaRNA a SomaVar LF.

Kraf používá strategii s jedním k-merem bez zarovnání. Rekalibrovaná alelická frekvence indelů a delinů je zachována pouze v případě, že je vyšší než původní alelická frekvence varianty. Varianty jsou pak anotovány především pomocí nástroje *Variant Effect Predictor*¹¹ (VEP) a další informace jsou doplněny z následujících databází:

- ClinVar
- OMIM (komerční verze) + PanelApp Australia + PanelApp England
- CKB
- dbSNP
- COSMIC (komerční verze)
- GnomAD (data exomu + genomu)
- Dfam
- DGV
- LOVD

Kromě toho jsou k dispozici následující výsledky *in silico*:

¹¹ <https://www.ensembl.org/info/docs/tools/vep/index.html>

- Scores. (*MaxEntScan, GeneSplicer, dbSCSNV*)
- Skóre zachování (*GERP++, PhastCons100way, PhastCons20way, PhyloP100way obratlovci, PhyloP20way savci, SiPhy*)
- Functional Prediction" (*REVEL, DANN, FATHMM, FATHMM-MKL, LRT, MetaLR, MetaSVM, MutationAssessor, MutationTaster 2, PROVEAN, SIFT, PLI*).

Vizualizace a interpretace

Karta Varianty  umožňuje zobrazit, filtrovat a anotovat varianty získané během analýzy. Ty jsou zobrazeny ve formě tabulky, přičemž každý řádek odpovídá jiné variantě. Počet vybraných variant po filtrování (vybrané varianty) a celkový počet zjištěných variant (celkový počet variant)

se zobrazuje v pravém horním rohu záhlaví tabulky variant pomocí následující ikony:  9/22 . Stránkování ve spodní části stránky umožňuje přecházet z jedné stránky výsledků na další. Pokud analýza používá omezený panel, bude záhlaví tabulky obsahovat také celkový počet variant v omezeném panelu (celkový počet variant v omezeném panelu) .

Tuto navigaci usnadňují klávesové zkratky pomocí kláves ← a → na klávesnici.

Ve výchozím nastavení je v tabulce zobrazeno několik popisných sloupců. Uživatelé mohou tabulku exportovat jako soubor TSV kliknutím na  .

Pro každou variantu může být k dispozici několik Transcripts, jejichž počet je zobrazen vlevo na každém řádku. Varianty jsou anotovány na transkriptech RefSeq, s výjimkou mitochondriálních variant, které jsou anotovány na *Ensembl*.

Transcripts MANE Select a MANE Clinical Plus jsou označeny ikonou



, respektive  , doplněnou nápovědou.

TRANSCRIPT	HGVS
MANE Select transcript	
NM_005962.5	c.536T>G p.V179G
NM_015252.5	c.-296+2T>G
MANE Plus Clinical transcript	
NM_000132.4	c.5375T>G p.V1792G

Pro stejnou variantu, s výjimkou mitochondriálních variant, jsou transkripty seřazeny podle níže uvedeného pořadí:

- ['high' > 'moderate' > 'low' > 'modifier']
- Type: ["NM", "XM", "NR", "XR"] (pro Transcripts RefSeq)
- Type: [Mane Select > MANE Clinical Plus > non MANE]
- alfanumerické pořadí

U mitochondriálních variant jsou transkripty seřazeny v následujícím pořadí:

- impact ['high' > 'moderate' > 'low' > 'modifier']
- pokud je dopad "modifier" :
 - 'non_coding_transcript_exon_variant' efekt
 - [protein coding > Mt_rRNA > Mt_tRNA]
- alfanumerické pořadí

DiagAI. Tento sloupec zobrazuje skóre od 0 do 100 na základě sady DiagAI využívající strojové učení. DiagAI Score je založeno na čtyřech složkách:

- Algoritmus UP², který předpovídá patogenitu varianty.
- matice PhenoGenius, která hodnotí shodu mezi genem a HPO pacienta (pokud byly zadány během analýzy).
- Shoda mezi pozorovaným genotypem a způsobem přenosu onemocnění popsáním v databázích GenCC, OMIM a PanelApp (Austrálie a Anglie).
- Kvalita varianty

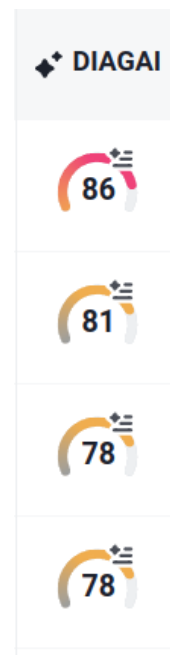
Čím více se skóre blíží hodnotě 100, tím více se předpokládá, že daná varianta je patogenní a diagnosticky zajímavá. Čím blíže je skóre k 0, tím více se předpokládá, že je varianta benigní.

Na určité zájmové varianty upozorňují také dvě ikony:

- Ikona  znamená, že daná varianta má vyšší pravděpodobnost, že je příčinou onemocnění (AI Smart Pick).
- Ikona  označuje, že se varianta nachází na krátkém seznamu variant, u nichž je vysoká pravděpodobnost, že jsou příčinou onemocnění pacienta.

Termíny HPO lze zadat před spuštěním analýzy úpravou metadat vzorku (viz část **Metadata vzorku**). Lze je také upravovat během interpretace pomocí modálního zobrazení HPO v postranním panelu analýzy (viz oddíl **Modální zobrazení HPO**). V takovém případě může uživatel znovu zpracovat skóre DiagAI a klasifikaci variant (Short list nebo SmartPick) podle nově zadaných HPO pomocí funkce "Re-process with new HPOs" (Znovu zpracovat s novými HPO).

Pokud uživatel klikne na ikonu skóre DiagAI, otevře se stránka varianty přímo na kartě DiagAI s podrobným popisem skóre (viz DiagAI v části Variantní stránka níže).

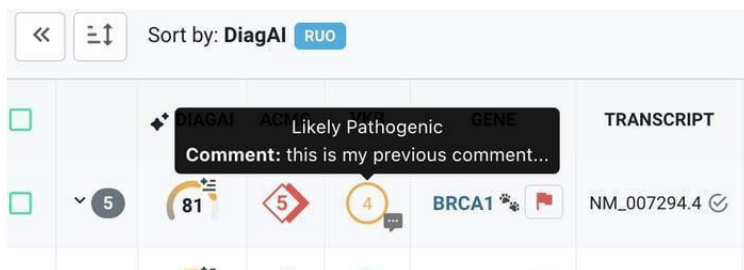


ACMG. Klasifikace patogenity každé varianty se vypočítává v rámci poloautomatického procesu podle doporučení vydaných *American College of Medical Genetics and Genomics*. Uživatelé mohou klasifikaci revidovat ve vyhrazené záložce na stránce varianty, která se pak použije k vygenerování konečného hodnocení varianty.

ComPerMed. Klasifikace patogenity každé varianty se vypočítává v rámci poloautomatického procesu podle doporučení vydaných belgickým výborem odborníků ComPerMed (Personalised Medicine Commission). Uživatelé mají přístup k podrobným informacím o klasifikaci na zvláštní kartě na stránce Variant.

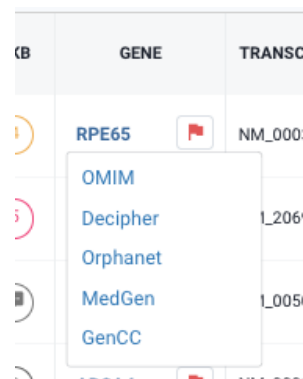
AMP. (Pouze verze analýzy SomaVar v1.8 nebo novější, SomaVar LF nebo SomaCGP). Tento sloupec se ve výchozím nastavení zobrazuje v analýzách, které obsahují anotace CKB, a zobrazuje úroveň AMP každé varianty, pokud je k dispozici (úroveň I nebo II). Zobrazená úroveň AMP vychází z nejlepších klinických důkazů u všech typů bez ohledu na vybranou indikaci.

VKB. Evaluations provedené uživatelem během předchozích analýz jsou uloženy ve znalostní bázi uživatele. Toto hodnocení se objeví ve sloupci VKB tabulky pro každý nový výskyt varianty. V nápovědě se rychle zobrazí třída a komentář související s hodnocenou nebo dříve hodnocenou variantou.



Gene. Pojmenování genů podle [HGNC](#) a související příznak OMIM :

- Červená vlajka: "OMIM gene" s fenotypy a klinickými charakteristikami ",
- Red obrysový praporek: "OMIM gene" s fenotypy, ale **bez** klinických příznaků ",
- Červený bod: "OMIM gene' **without** phenotypes or clinical features "



Genes, o nichž je známo, že jsou podle aplikace PanelApp (Austrálie) imprintingovými geny, jsou zvýrazněny ikonou . Související tooltip označuje, zda je gen mateřsky exprimovaný, paternálně exprimovaný, neznámý, nebo zda se vztahuje k několika imprintingům.

Kliknutím na vlajčku vedle názvu genu se dostanete přímo na stránku OMIM. Kliknutím na

název Gene se zobrazí rozbalovací nabídka, která poskytuje odkazy na databáze OMIM gene¹², Decipher gene¹³, Orphanet gene¹⁴ MedGen gene a GenCC gene¹⁵¹⁶. Každý odkaz vede přímo na stránku odpovídající vybranému genu.

Informace týkající se OMIM a databází jsou k dispozici pouze pro zárodečné analýzy.

HGVSC. V tomto sloupci je uvedena anotace HGVS varianty na komplementární DNA (cDNA) a na proteinu.

HGVSC

c.10580G>A
p.R3527Q

Effect. Tento sloupec popisuje dopad varianty na protein a vliv na transkript.

Dopad se vypočítává podle 4 úrovní:

- **High** (high impact, in red): Variant má vysoký rušivý dopad na protein, pravděpodobně způsobuje zkrácení proteinu, ztrátu funkce nebo spouští jeho rozpad zprostředkovaný nesmyslem.
- **Moderate** (moderate impact, žlutě): nenarušující varianta, která pravděpodobně změní účinnost proteinu.
- **Low** (low impact, šedě): předpokládá se, že varianta je neškodná nebo že pravděpodobně nezmění chování proteinu.
- **Modifier** (modifying impact, "blue"): obecně nekódující varianty nebo varianty ovlivňující nekódující oblasti nebo geny, u nichž jsou předpovědi obtížné nebo není prokázán dopad.

EFFECT	
Moderate	missense
Low	splice polypyrimidine tract + 1
Moderate	missense
Moderate	missense
Moderate	missense
Moderate	missense

Pro každou variantu může být ve sloupci Effect k dispozici několik účinků na transkript. V takovém případě je zobrazen pouze účinek s nejvyšším důsledkem podle VEP¹⁷, za nímž následuje číslo (např. +2) pro ostatní dostupné účinky. Podrobnosti o nich jsou zobrazeny v příslušném tooltipu.

¹² <https://www.omim.org/>

¹³ <https://www.deciphergenomics.org/genes>

¹⁴ <https://www.orpha.net/fr/disease/gene>

¹⁵ <https://www.ncbi.nlm.nih.gov/medgen/>

¹⁶ <https://search.thegencc.org/>

¹⁷ http://www.ensembl.org/info/genome/variation/prediction/predicted_data.html

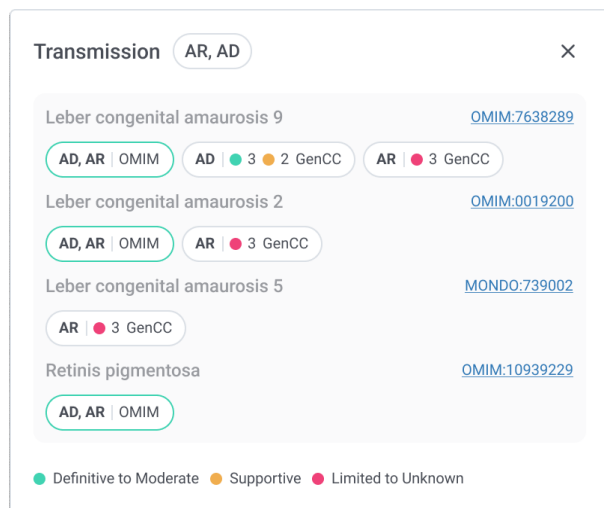
Transmission. Tento sloupec označuje způsob přenosu onemocnění spojených s geny uvedenými v databázích GenCC a OMIM. Barevné odznaky jsou viditelné pro rychlou identifikaci způsobů přenosu na základě předpovědi DiagAI:

- "De Novo"
- "Compound Het - Verified"
- "Recessive Hom"
- "Compound Het - Suspected"

De Novo, Compound Het - Verified a Recessive Hom se zobrazují pouze v rodinných analýzách.

K výběru variant s těmito odznaky lze použít filtr (viz **část Filtrování variant - Transmission**).

TRANSMISSION	
AD AR	CompHet Verified Nail disorder, nonsyndromic ... +1
AD	CompHet Suspected Coffin-Siris syndrome 2
AD	De Novo Advanced sleep phase syndrome..
AR	Recessive Hom Advanced sleep phase syndrome..
AD	Tarsal-carnal coalition svndrome +3



Kliknutím na odkazy na způsoby přenosu v tomto sloupci se otevře vyskakovací okno, kde uživatelé mohou najít nemoci, které jsou uvedeny v seznamu, s jejich názvy a odkazy na databáze (OMIM, MONDO - pokud není k dispozici OMIM - nebo GenCC).

Nemoci jsou uspořádány ve specifické hierarchii: nejprve se objevují nemoci nalezené v databázích OMIM i GenCC, poté záznamy pouze v databázi OMIM a nakonec záznamy pouze v databázi GenCC.

Každá položka obsahuje barevně odlišené štítky: Značky OMIM mají zelené okraje, zatímco značky GenCC používají systém semaforu (zelená/žlutá/červená) pro označení síly důkazů:

- Zelená: Definitivní + Silný + Střední
- Žlutá barva: Podpůrné
- Red: Červená: omezená, zvířecí, neznámá

Clinical Evidences. (pouze Analysis SomaVar v1.8 nebo novější, SomaVar LF nebo SomaCGP). Tento sloupec se ve výchozím nastavení zobrazuje v analýzách obsahujících anotace CKB. Je rozdělen do tří podsloupců, z nichž každý obsahuje celkový počet daného typu důkazů: A (Actionability), D (Diagnosis) a P (Prognosis). Ve sloupci A je také zobrazen počet citlivých (vlevo) a rezistentních (vpravo) terapeutických tvrzení jako odznak.

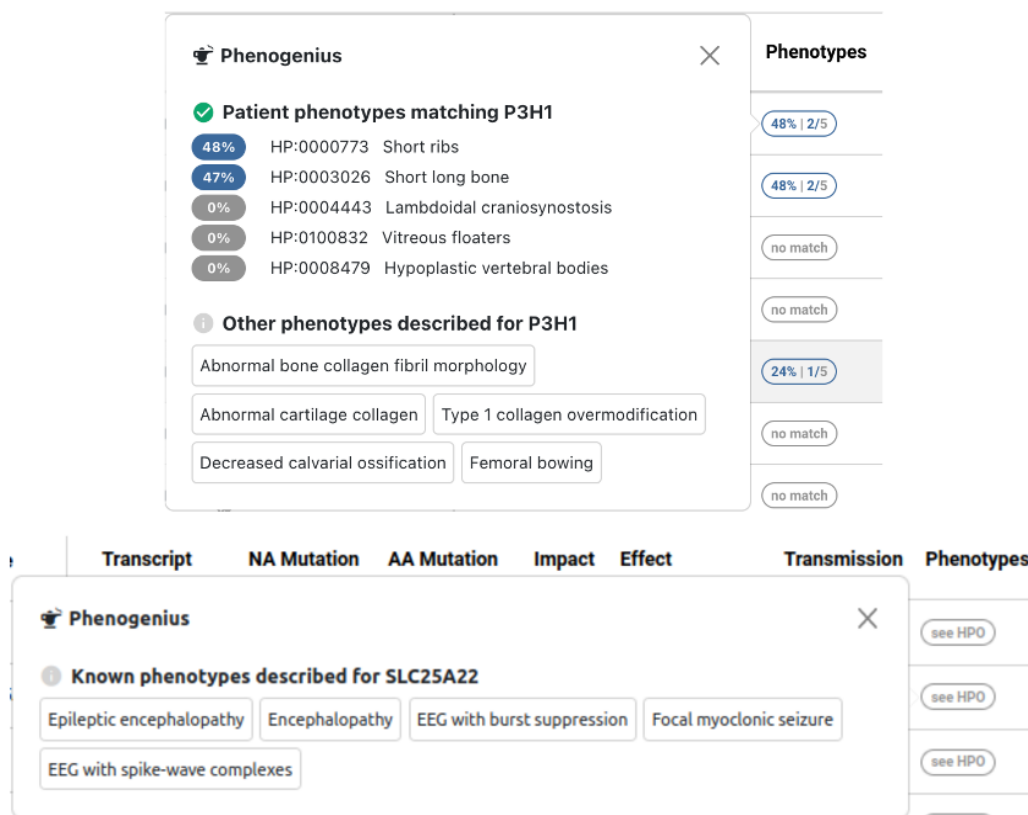
CLINICAL EVIDENCES		
A	D	P
81 81 8	0	0
54 53 2	0	0
54 53 2	0	0

Hodnoty v těchto dílčích sloupcích jsou barevně odlišeny:

- Pokud je nalezena rezistence (> 0 tvrzení), je vždy zobrazena červeně bez ohledu na indikaci a jinak šedě.
- Citlivost je zvýrazněna zeleně pouze v případě, že je vybrána příslušná indikace, a šedě v opačném případě.

Pokud se alespoň jeden důkaz tohoto typu shoduje s indikací vybranou v tabulce variant (Panel indikací), zobrazí se v nápovědě před označením indikace zatržítko. Pokud není nalezena žádná shoda, popis se nezobrazí.

Phenotypes. Tento sloupec, dostupný pro projekty Germline, umožňuje přiřazení genu k termínům HPO pomocí matice PhenoGenius¹⁸. Pokud jsou se vzorkem spojeny termíny HPO, zobrazí se nejvyšší procentuální shoda mezi genem a fenotypem spolu s poměrem termínů, s nimiž je gen spojen. Pokud se zadaným HPO neshoduje žádný gen, sloupec vrátí zprávu "no match"; pokud vzorek nemá žádné HPO termíny, sloupec zobrazí zprávu "see HPO". V obou případech jsou nicméně známé termíny pro daný "Gene" navrženy v souvisejícím vyskakovacím okně.

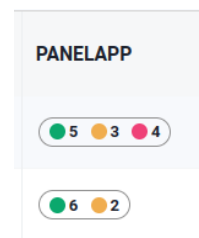


Obrázek 11 : Zobrazení informací o programu PhenoGenius

¹⁸ Kevin Yauy a kol., " Learning phenotypic patterns in genetic diseases by symptom interaction modeling ", medRxiv, p. 2022.07.29.22278181, janv. 2022, doi: 10.1101/2022.07.29.22278181.

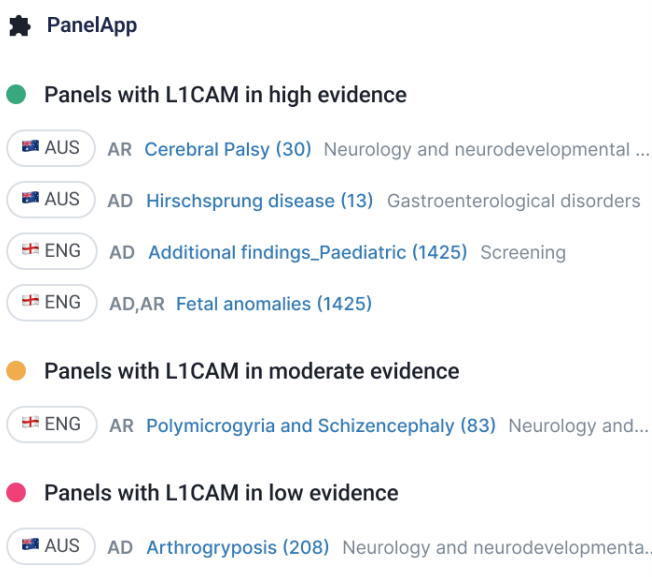
PanelApp. Tento sloupec integruje anotace z databáze panelů *PanelApp Australia*¹⁹ *PanelApp England*²⁰. Indikuje počet panelů spojených s genem, pro který byla zjištěna varianta, podle níže popsané klasifikační úrovně.

Pokud se gen v panelu nenachází, zobrazí se ve sloupci "N/A".



V souvisejícím vyskakovacím okně se zobrazí podrobnosti o asociaci genu a fenotypu, jakož i způsob přenosu spojený s fenotypem, způsob přenosu spojený s fenotypem a databáze, ze které je panel odvozen, PanelApp Australia (AUS) nebo PanelApp England (ENG) (Obrázek 12).

Každému genu bude přiřazen jeden nebo více panelů podle třístupňového klasifikačního systému založeného na důkazech o souvislosti genu a fenotypu: Vysoká důkaznost v **zelené barvě** **žlutou barvou** jsou označeny střední důkazy a **červenou barvou** jsou označeny nízké důkazy.

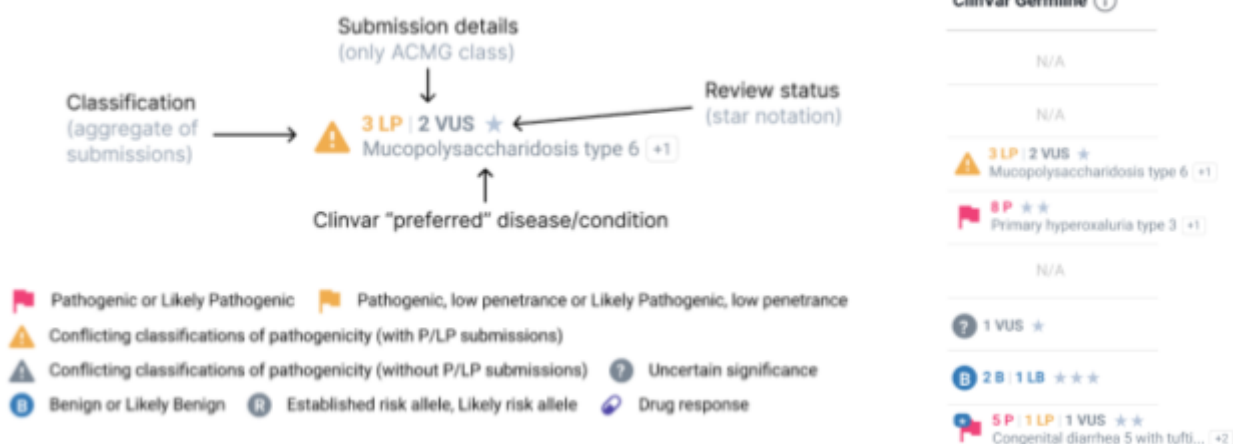


Obrázek 12: Vizualizace informací aplikace PanelApp.

¹⁹ <https://panelapp-aus.org>

²⁰ <https://panelapp.genomicsengland.co.uk/>

ClinVar Germline. Tento sloupec zobrazuje informace získané z "ClinVar" během analýzy. Poskytuje řadu informací, včetně třídy patogenity, podrobností a počtu podání pro každou třídu ACMG, stavu revize a preferovaného stavu/nemoci ClinVar. Kliknutím na buňku v tabulce Variant bude uživatel přesměrován na příslušnou stránku "ClinVar".



Další informace naleznete na adrese https://www.ncbi.nlm.nih.gov/clinvar/docs/review_status/.

GnomAD. Tento sloupec zobrazuje 6 prvků z databáze GnomAD pro konkrétní variantu v populaci: alelickou frekvenci (AF), počet alel (AC), počet homozygotů (HOM), počet hemizygotů (HEM), frekvence hlavní alely (MAF) a pro projekty GRCh37 verzi GnomAD, pro kterou se tato varianta nachází (VER).

Kliknutím na hodnotu zobrazenou ve sloupci AF je uživatel přesměrován do databáze GnomAD ve verzi v4, případně ve verzi v2. Pokud varianta není v databázi uvedena, bude hodnota zobrazená v tomto sloupci N/A.

GNOMAD				PROJ
	AC	HOM	VER	
e-1	47775	9503	v2	4 / 4

Pokud je tato varianta přítomna v databázi GnomAD, ale neprojde všemi filtry kvality alespoň pro jeden ze dvou souborů dat (exom nebo genom), zobrazí se ve sloupci VER upozornění. V takovém případě se zobrazené hodnoty přizpůsobí hodnotám, které jsou pro tuto variantu k dispozici, a zobrazí se filtry, které výstrahu vyvolaly, spolu s krátkou zprávou uvádějící porušující datovou sadu

(sady).

Score UP². Skóre funkční predikce UP² (Universal Pathogenicity Prediction) v rozsahu -1 až 1 klasifikuje patogenitu variant pomocí přístupu strojového učení podle tří kategorií: **P** (Pathogenic : skóre > 0,1), VUS (Variant of **Unknown** Significance : skóre ≥ -0,1 et ≤ 0,1) nebo B (Benign : skóre < -0,1). Skóre UP² se v tabulce variant zobrazuje ve výchozím nastavení, ale uživatel může

sloupec UP² deaktivovat na panelu správy sloupců kliknutím na tlačítko



CI-SpliceAI. **RUO** Tento sloupec zahrnuje skóre predikce vypočtené pomocí přístupu strojového učení pro identifikaci variant s vlivem na alternativní sestřih po vzoru *CI-SpliceAI*.²¹

SpliceAI				
AG	DG	AL	DL	Link
0.01	0.02	0.62	0.84	↗
0.01	0.22	0.01	0.09	↗
0.01	0.02	0.12	0.01	↗
0.01	0.62	0.02	0.01	↗
N/A	N/A	N/A	N/A	↗
No splicing impact				
SpliceAI score was not significant (Delta score <0.2)				

Scores se vypočítá pro podmnožinu variant, které splňují následující kritéria: splice varianty podle VEP (splice_*) s frekvencí GnomAD menší než 10⁻³ a splňující následující kritéria kvality (QA > 20, AO > 10, DP > 20).

Pokud je vypočteno, sloupec uvádí 4 delta skóre pro danou variantu, a to pro zisk akceptoru (AG), ztrátu akceptoru (AL), zisk donoru (DG) a ztrátu donoru (DL).

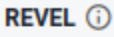
Skóre delta varianty, které se pohybuje od 0 do 1, odpovídá pravděpodobnosti, že varianta modifikuje splicing. Pro prahové hodnoty delta skóre se používá barevný kód: zelená (>0,2 ; vysoká odvolatelnost), žlutá (>0,5; doporučená) a červená (>0,8; vysoká přesnost). Pokud je delta skóre nižší než 0,2, považujeme skóre za nevýznamné. Ve sloupci je pak uvedeno "No splicing impact".

Pro každou variantu je uveden odkaz na příslušné rozhraní *SpliceAI* pro vyhledávání²²

²¹ https://github.com/YStrauch/CI-SpliceAI__Annotation

²² <https://spliceailookup.broadinstitute.org/>

REVEL. Skóre predikce REVEL je založeno na celkové metodě predikce patogenicity missense variant pomocí kombinace skóre z 13 jednotlivých nástrojů: MutPred, FATHMM v2.3, VEST 3.0, PolyPhen-2, SIFT, PROVEAN, MutationAssessor, MutationTaster, LRT, GERP++, SiPhy, phyloP a phastCons.

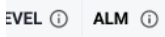


B

P

Skóre se může pohybovat od 0 do 1 pro missense variantu, přičemž vyšší skóre odráží větší pravděpodobnost, že varianta způsobuje onemocnění. Použijí se následující prahové hodnoty: REVEL skóre > 0,5 pro P (Pathogenic) a REVEL skóre < 0,5 pro: **B** (Benign). Skóre může uživatel zobrazit v tabulce variant zaškrtnutím políčka REVEL ve vložce "Columns" (Sloupce) v části "Functional Prediction" (Funkční predikce).

ALM. Skóre predikce AlphaMissense²³ je k dispozici ve sloupci ALM. AlphaMissense je predikční model, který posuzuje patogenicitu missense variant pomocí hlubokého učení. Model je postaven na strukturálních předpovědích AlphaFold, aby se zvýšila jeho přesnost.



P

P

B

AlphaMissense integruje evoluční konzervaci a strukturní kontext pro posouzení potenciálního dopadu missense variant na funkci proteinu.

Skóre se pohybuje v rozmezí od 0 do 1, přičemž vyšší hodnoty naznačují větší pravděpodobnost, že varianta je škodlivá. Použijí se tyto prahové hodnoty: ALM skóre > 0,564 pro P (Likely Pathogenic) a ALM skóre < 0,34 pro B (Likely Benign). V ostatních případech (0,34 < ALM skóre < 0,564) bude označena jako VUS (Ambiguous).

Sample Freq. Tento sloupec zobrazuje pro každou variantu její četnost v projektu (podsloupec PROJECT) a ve všech projektech subjektu s podobnými metadaty (podsloupec "ALL").

Sample frequency. v projektu (PROJECT) odpovídá počtu vzorků, u nichž byla zjištěna varianta, ve vztahu k celkovému počtu vzorků analyzovaných v projektu.

Sample Freq. ALL odpovídá procentu analyzovaných vzorků, ve kterých se daná varianta vyskytuje u dané entity, v rámci podobných projektů

SAMPLE FREQ.	
PROJECT	ALL
2 / 13	7.41%
2 / 13	7.41%

²³ <https://alphamissense.hegelab.org/>

metadat. Kliknutím na hodnotu zobrazenou ve sloupci "ALL" získá uživatel přístup k názvům projektů a počtu vzorků, v nichž se varianta vyskytuje, a také k podrobnostem o všech podobných metadatových projektech (možnost Zobrazit všechny projekty).

Ve výchozím nastavení je podsloupec ALL v tabulce variant skrytý, ale lze ji znovu aktivovat na

panelu správy sloupců kliknutím na ikonu  a aktivací tlačítka „Sample Frequency Entity (Entity)“.

Exon. V tomto sloupci je uveden celkový počet exonů v genu podle příslušného transkriptu a počet exonů ovlivněných mutací ve formě zlomku. V nápovědě je uvedena vzdálenost mezi variantou a nejbližším místem splicing. Pokud se varianta nachází před místem sestřihu, bude hodnota kladná, a pokud se nachází za ním, bude záporná.

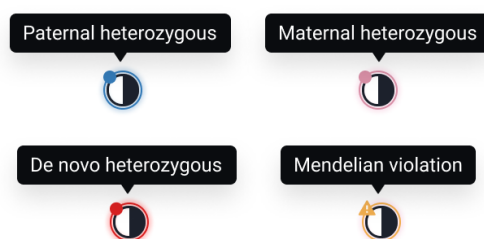
GT (Genotype). V tomto sloupci je uveden genotyp předpovězený programy pro detekci variant. Různé možné piktogramy jsou popsány v části **Filters**.

Pokud lze v analýze trojic (GermlineFamily) určit původ heterozygotní varianty, bude označen následujícím barevným kódem:

- **pink** → heterozygotní původ matky
- **blue** → heterozygotní původ otce
- **red** → De novo heterozygotní varianta
- **orange** → Mendelovo porušení

Barva se nepoužije, pokud je přenos neurčitý (např. pokud jsou oba rodiče heterozygoti), nebo u homozygotních variant.

Patient				Father				Mother			
GT	VAF	QUAL	COV	GT	VAF	QUAL	COV	GT	VAF	QUAL	COV
	50.8%	34	61		N/A	N/A	82		44.6%	33	83
	53%	34	83		N/A	N/A	49		N/A	N/A	100
	46.9%	34	64		48.8%	33	84		N/A	N/A	86



Obrázek 13: Barevný kód použitý pro heterozygotní genotypy na základě přenosu GermlineFamily

Fázování variant je k dispozici pro analýzy GermVar Tertiary a GermVar Tertiary Family pomocí souborů VCF ze sekvenátorů Oxford Nanopore Technologies. Když jsou varianty rozfázovány, zobrazí se ikony genotypů takto : . Zaškrtnutí znamená, že varianta je fázovaná. Fázované varianty s černou částí na stejné straně a patřící ke stejnému "Phase ID" jsou ve fázi, tj. na stejném vlákně.

Kliknutím na genotyp získáte seznam cis nebo trans variant, pokud jsou přítomny. Zobrazeno je pouze 5 variant s nejvyšším hodnocením podle DiagAI, jak je uvedeno na protější straně.

Kompletní seznam variant je k dispozici po kliknutí na "Browse all variants".

Variants in *trans*
Phase set ID 283928 X

! Compound heterozygous candidates for

4 SLC25A22 chr11:g.792142C>T

5 SLC25A22 chr11:g.792146_792147del

GT

VAF (Variant Allele Frequency). Tento sloupec udává frekvenci variantních alel, tj. podíl čtení podporujících alternativní alelu (AO) k celkovému počtu čtení na dané pozici (COV), vyjádřený v procentech.

Analýzy používající worksety SomaVar verze ≥ 2.5 , SomaCGP verze ≥ 0.7 a SomaVar LF verze ≥ 2.1 zobrazují následující doplňující informace:

- nápopěda zobrazuje hodnoty VAF, AO a COV vypočtené dvěma *variant callers* Freebayes a Mutect2, které danou variantu detekují.
- varovná ikona  upozorňuje uživatele na případný nesoulad mezi 2 variant callers u variant detekovaných s nízkou hodnotou VAF. U variant, které ve sloupci VAF vykazují VAF nižší než 5 %, se tedy varování zobrazí v případě, že mezi VAF vypočtenými Freebayesem a Mutect2 je rozdíl > 10 %.

AO (Alternate Observation). Tento sloupec udává počet reads nesoucích danou variantu.

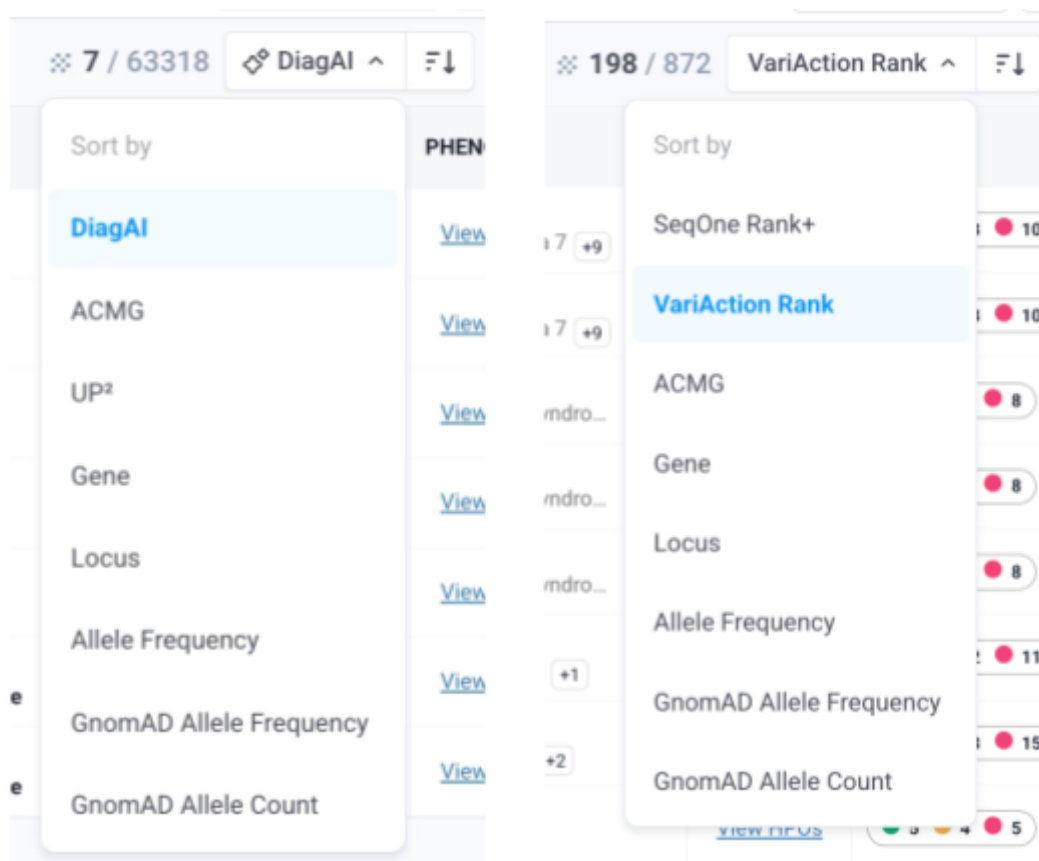
CAL (Caller). V tomto sloupci je uvedeno jméno *variant caller* (s) který (kteří) danou variantu zjistil (zjistili).

Cov (Coverage). Tento sloupec udává počet celkových reads na pozici varianty.

SOR. Strand Odds Ratio (SOR) je metrika vláknové předpojatosti (strand-bias) odvozená z testu symetrického poměru šancí (Symmetric Odds Ratio), která měří nerovnováhu čtení podporujících variantu mezi přímými (forward) a zpětnými (reverse) vlákny. Vyšší SOR naznačuje vyšší pravděpodobnost artefaktu, zatímco nižší SOR naznačuje vyváženější podporu vláken a je konzistentnější se skutečnou variantou.

Třídění variant

Na kartě Variant: je několik možností třídění variant. Ty jsou dostupné z rozbalovací nabídky v pravém horním rohu tabulky. U každé možnosti řazení je možné seřadit vzestupně nebo sestupně kliknutím na ikonu řazení . Ve výchozím nastavení jsou všechny možnosti seřazeny v sestupném pořadí, s výjimkou možnosti "Locus", která je seřazena vzestupně.



Obrázek 14: Okno pro výběr možností třídění variant

DiagAI Score. RUO Tato možnost řazení je ve výchozím nastavení použita pro všechny analýzy germline a umožňuje seřadit varianty podle jejich skóre DiagAI (nahrazuje SeqOne Rank+ v analýzách provedených po 2. říjnu 2024).

SeqOne Rank+. RUO Algoritmus prioritizace SeqOne Rank+, založený na starším modelu umělé inteligence než DiagAI Score, zůstává k dispozici a je standardně aplikován na všechny analýzy provedené před 2. říjnem 2024.
Je založen na 4 hlavních složkách:

- Nová implementace klasifikace ACMG založená na strojovém učení vycvičeném na ClinVar (skóre UP²).
- Fenotyp pacienta (pokud jsou se vzorkem spojeny termíny HPO).
- Shoda mezi zjištěným genotypem a způsobem přenosu patologie popsáním v databázích GenCC, OMIM a PanelApp (Austrálie).
- Kvalita Variant

VariAction Rank. RUO Algoritmus VariAction Rank pro prioritizaci somatických variant je navržen v analýzách SomaVar od verze 1.8. Tento systém využívá anotaci JaxCKB k výpočtu skóre prioritizace variant na základě jejich klinické aktivity. Tato prioritizační skóre zohledňují dopad varianty, její patogenitu pomocí skóre ComPerMed, AMP Tier, specifickou variantu s molekulárním profilem a také shodu s indikací, pokud bylo onemocnění zadáno před zahájením analýzy.

ACMG. RUO Výběrem této možnosti lze varianty seřadit podle jejich klasifikace ACMG. Ve výchozím nastavení jsou varianty řazeny sestupně od patogenní třídy 5 po benigní třídu 1.

Locus. Výběrem této možnosti lze varianty seřadit podle jejich genomové pozice.

Allele frequency. Výběrem této možnosti lze varianty seřadit podle frekvence alel. Ve výchozím nastavení budou varianty seřazeny sestupně. Seřadit vzestupně je možné kliknutím na ikonu

řazení. 

GnomAD Allele Frequency, GnomAD Allele Count. Varianty lze seřadit podle jejich frekvence alel (AF) nebo počtu alel (AC) v GnomAD ve vzestupném nebo sestupném pořadí (pokud jsou data GnomAD k dispozici).

UP². Varianty lze řadit podle skóre patogenity UP², což je složka DiagAI Score.

Gene. Varianty lze řadit vzestupně nebo sestupně podle abecedy podle symbolu genu. U variant ovlivňujících překrývající se geny je pro řazení použit symbol genu spojený s variantou s nejvyšším dopadem na základě pravidel prioritizace transkriptů popsaných výše. Při rozšíření seznamu transkriptů může uživatel vidět druhý překrývající se gen, který se může jevit jako nezařazený. Řazení se také nemusí projevit v případě, že je na tabulku variant aplikováno více filtrů najednou.

Konfigurace sloupců tabulky Variant

Sloupce zobrazené v tabulce variant lze konfigurovat v nabídce *Organise your columns*, která je

přístupná prostřednictvím tlačítka  v pravém horním rohu tabulky variant. Zobrazení

sloupců může uživatel modifikovat kliknutím na ikony  pro aktivaci požadovaného sloupce (sloupců) v levé části. Pravá část zobrazuje aktivní sloupce v analýze a umožňuje uživatelům vytvářet a ukládat profily sloupců z rozbalovací nabídky kliknutím na ikonu "Select profile".

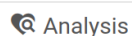
Kliknutím a přetažením aktivních sloupců na požadované místo lze změnit jejich pořadí. Ve výchozím nastavení jsou sloupce uspořádány podle "Germline" a "Somatic" profilů.

Když uživatel najde na název profilu sloupce v rozbalovací nabídce předvolby, zobrazí se nápověda, která workset daný profil vytvořil. To uživatele upozorní, pokud by jejich profil mohl mít sloupce nekompatibilní s aktuální pracovní sadou.

Chcete-li vytvořit nový profil sloupce, jednoduše zadejte název do pole "Create a new preset" a klikněte na tlačítko "Save".

Úprava metadat vzorku z panelu analýzy

Uživatel může upravovat metadata vzorku přímo z analýzy a zadat pohlaví pacienta nebo HPO (pouze pro zárodečné analýzy) ze zásuvky analýzy, která je přístupná kliknutím na tlačítko



v pravém horním rohu stránky (viz část Postranní **panel pro správu analýzy**).

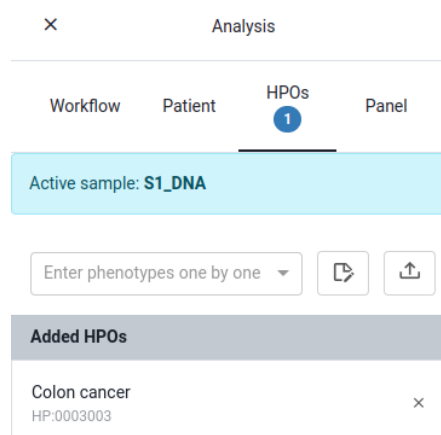
V zobrazení Pacient má uživatel přístup k úpravě pohlaví (viz **část Patient modal view**).

V analýzách projektu Germline má uživatel v zobrazení HPOs přístup k několika funkcím pro úpravu termínů HPOs:

- Zobrazení již zadaných termínů před spuštěním analýzy
- Ruční přidání kódů HPO nebo fenotypů pomocí vyhledávacího panelu "Enter phenotypes one by one".
- Import textového souboru obsahujícího kódy HPO
- Generování termínů HPO na základě klinického popisu pacienta vyvoláním DiagAI pomocí ikony



. Tato funkce vyzve uživatele k popisu klinického profilu pacienta v italštině (popis je omezen na 4000 znaků). Poté se po kliknutí na "Extract HPOs with DiagAI" (Výpis HPO pomocí DiagAI) vygeneruje barevně



zvýrazněný seznam HPO. Nakonec může uživatel přiřadit všechny navržené termíny HPO k pacientovi kliknutím na "Add all" (Přidat všechny) nebo je vybrat jednotlivě najetím myši na vybrané termíny a následným kliknutím na "Add" (Přidat).

Přidání a úprava HPO z tohoto zobrazení dává uživateli možnost znovu zpracovat skóre DiagAI a klasifikaci variant (Short list nebo SmartPick) podle nových HPO zadaných prostřednictvím funkce "Re-process analysis".

Varianty filtrování

Na kartě *Variant*: můžete vyhledávání zpřesnit pomocí níže uvedených kategorií filtrů. Panel filtrů je přístupný prostřednictvím ikony  umístěné v levém horním rohu tabulky variant nebo stisknutím klávesy F na klávesnici. Filtry jsou k dispozici na kartě "Basic". V horní části panelu filtrů se v části „Active Filters“ zobrazuje počet aktuálně použitých filtrů.

Filtry Genes & Variants

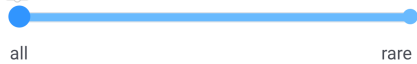
Název a přehled filtru	Description
Genes ⓘ Search gene symbols	Omezí seznam variant na jeden nebo více genů.
Variants ⓘ Search by HGVS	Omezuje seznam variant na jednu nebo více variant podle nomenklatury HGVS _c nebo HGVS _p .

Sekce Samples & Quality

Název a přehled filtru	Description
Coverage ⓘ 1 1 836	Definuje minimální prahovou hodnotu pro hloubku pokrytí (COV). Maximální hodnota zobrazená tímto filtrem odpovídá nejvyšší hodnotě pokrytí identifikované mezi variantami.
Allele Frequency ⓘ 0% 0 100 100%	Filters varianty podle četnosti mutantní alely v analyzovaném vzorku.
Quality ⓘ Low Medium High	Průměrná kvalita (phred skóre) každé báze nesoucí mutovanou alelu, která udává důvěryhodnost detekce varianty: <i>Medium</i> ($\geq \text{QUAL}=20$): 1 šance na nesprávné volání ze 100, <i>High</i> ($\geq \text{QUAL}=30$): 1 šance na nesprávné volání z 1000. Tento filtr se vztahuje na všechny vzorky.
Quality ⓘ All Low $\geq Q8$ Medium $\geq Q20$ High $\geq Q30$	U analýz GermVar Tertiary a GermVar Tertiary Family na datech od společnosti Oxford Nanopore Technologies indikuje kvalita založená na skóre Phred důvěru v detekci variant. $\geq \text{QUAL}=8$ (Low): ($\text{QUAL}(\text{QUAL})$): 1 šance na nesprávné volání z 8, $\geq \text{QUAL}=20$ (Medium): 1 šance na nesprávné volání ze 100, $\geq \text{QUAL}=30$ (High): 1 šance na nesprávné volání z 1000. Tento filtr se vztahuje na všechny vzorky.

Genotype ⓘ Homozygous 387 Heterozygous 1163 None 1942 Biallelic Variant 222 Biallelic Variant 84 Unknown 76 Unknown 0	Filtry variant podle sledovaného genotypu.
Alternate Observations ⓘ 0 0 458	Definuje počet sekvencí nesoucích mutaci (alternativní pozorování nebo AO). Maximální hodnota zobrazená tímto filtrem odpovídá nejvyšší hodnotě AO identifikované mezi variantami.
Strand Odds Ratio ⓘ 0 0 1 2 3 4 5 ≥5 min 0 — max 5	Filtruje varianty podle skóre strand odds ratio (SOR) vypočítaného v místě varianty.

Sekce Frequencies

Název a přehled filtru	Description
▼ Sample frequency ⓘ all  all rare	Vyloučí varianty na základě jejich opakovaného výskytu ve zkoumaném projektu, od nejčastějších (všechny) po nejméně časté (vzácné).
▼ Relative pop frequency ⓘ all 10-2 10-3 10-4 10-5 10-6 absent 	Vybírá varianty podle četnosti jejich mutantní alely v kombinované populaci projektů ExAc a gnomAD.
▼ Absolute pop frequency ⓘ 0 1609917 	Vybírá varianty na základě počtu alel v kombinované populaci ExAc a GnomAD.
▼ Absolute homozygous pop frequency ⓘ 0 803576 	Vybírá varianty na základě jejich homozygotních frekvencí v kombinované populaci ExAc a GnomAD.

Sekce Genetics

Název a přehled filtru	Description
▼ Impact ⓘ Low Moderate High ☒ Keep <i>Modifier</i>	Vliv mutace na protein. <i>Modifier</i>: ovlivňuje nekódující geny
▼ Effects ⓘ > Loss of Function 10 • Transcript ablation 10 • Splice acceptor 10 • Splice donor 10 • Stop gained 10 • Frameshift 10 • Stop lost 10 • Start lost 10 • Transcript amplification 10 ↺ ☒ ☒ ☒ ☒ ☒ ☒ ☒ ☒	SO (Sequence Ontology) termíny, které popisují nejvíce rušivé následky na úrovni genů. Tento filtr umožňuje vybrat dopad mutace na gen a možnosti jsou pro lepší orientaci uspořádány do logických kategorií: <i>Loss of Function, Coding, Splicing region, Silent variant, Intronic, UTR, ncRNA, Upstream/Downstream, Regulatory, Others</i>.
▼ Gene panel ⓘ Select panel ▼	Vyberte a použijte jeden nebo více genového panelu. Tento filtr zobrazí všechny Gene panels dostupné pro danou entitu. Tento filtr není k dispozici, pokud byl na analýzu použit omezený panel z karty <i>Panel</i> v podokně správy analýzy (viz část Modální zobrazení panelu). Zpráva informuje uživatele, že filtr "Gene panel" nelze použít s omezeným panelem.
▼ Panel in-silico ⓘ Dyslipidemia_panel_1_Twist★ ▼	Vybere a použije panel in silico. Tento filtr zobrazí všechny dostupné soubory BED pro danou skupinu. Název souboru použitého při analýze je zvýrazněn hvězdičkou. Tento filtr zachovává pouze varianty, jejichž počáteční pozice (start) se nachází v jednom z intervalů vybraného souboru lůžka.

PanelApp ⓘ Epi ENG Blepharophimosis ptosis and epicanthus ENG Early onset or syndromic epilepsy ENG Epidermodysplasia verruciformis	Umožňuje filtrovat varianty podle seznamu in-silico panelů v katalozích PanelApp Australia a PanelApp England. Odznak označuje databázi, ze které panel in silico pochází: AUS (Austrálie) nebo ENG (Anglie). V obou katalozích nejsou k dispozici superpanely pro výběr, stejně jako panely, které neobsahují alespoň jeden gen. Například panel obsahující pouze STRs nebude možné vybrat.
Gene occurrences ⓘ Usable only if they are fewer than 2000 selected variants.	Definuje prahovou hodnotu pro počet výskytů daného genu v tabulce variant.
Chromosomes & MT ⓘ X 25 Y 0 M 0	Zachovává varianty na pohlavních chromozomech a/nebo mitochondriálním chromozomu (M).
RefSeq transcripts ⓘ BARD1_2transcripts.txt	Vybere seznam transkriptů RefSeq pro filtrování variant.

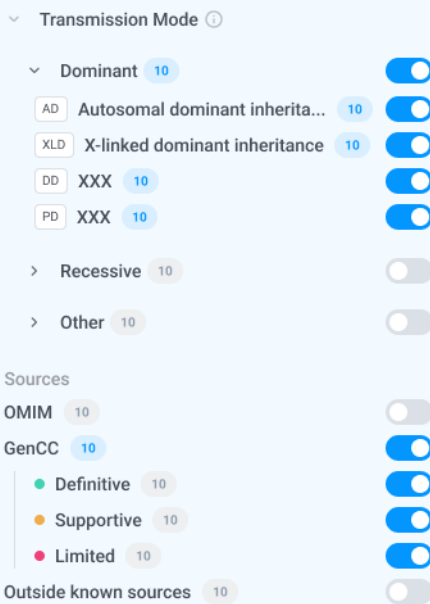
Sekce Databases

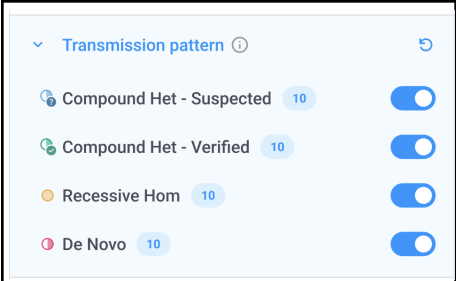
Název a přehled filtru	Description
Variant Databases ⓘ dbsnp 31 cosmic 31 civic 30 lovd 10 clinvar 3	Vybírá varianty anotované v následujících databázích: dbsnp, cosmic, lovd, clinvar a civic.

▼ Clinvar ⓘ P or LP 3 P/LP low penetrance 0 Conflicting (with P/LP subs) 1 Conflicting (without P/LP subs) 0 Uncertain significance 4 B or LB 18 Risk allele 0 Drug response 2 ☐ Keep others	Umožňuje filtrovat varianty podle jejich stavu patogenity na ClinVar. Pole "Ponechat ostatní" slouží k ponechání variant bez anotace ClinVar.
▼ OMIM ⓘ omim morbid clinic 111 omim without phenotypes 23 omim morbid 6	Umožňuje filtrovat podle morbidní kliniky OMIM, OMIM bez fenotypů a morbidních variant OMIM. Tento filtr má tlačítko in/out, které lze použít k výběru nebo odstranění stavu.
▼ SeqOne Artifacts ⓘ ☐ Remove artifacts	Tlačítko k odstranění opakujících se variant přítomných v 5 % kohorty 500 exomů považovaných za artefakty podle algoritmu SeqOne.

Sekce Transmission & Phenotypes

Název a přehled filtru	Description
▼ Description keywords ⓘ	Pomocí klíčových slov charakterizuje fenotyp(y) spojený(é) s variantami pomocí dotazů do databází OMIM, Medgen a Litvar. Tato funkce nebere v úvahu HPOs, protože těm je věnován zvláštní filtr (viz níže). K dispozici jsou dva režimy vyhledávání: sjednocení a průnik.

	Umožňuje filtrovat varianty na základě dříve zadaných termínů HPO, a to buď před analýzou vzorku (viz metadata vzorku), nebo po analýze na kartě HPO v podokně analýzy (viz modální zobrazení HPO). Shody jsou identifikovány pomocí našeho přístupu PhenoGenius, který se při výběru genů opírá o matici HPO-genu. V režimu filtru jsou k dispozici 2 možnosti : • Výchozí filtrování, které vyvažuje přesnost a počet vrácených genů. • Rozšířené filtrování maximalizující přesnost s více geny Možnost "See gene list" poskytuje seznam genů spojených s HPO vybranými v tomto filtru. Tato funkce je k dispozici pouze v případě, že jsou zaškrtnuty podmínky HPO. Otevře se vyskakovací okno s ID každého genu, symbolem, hodnocením v programu Phenogenius a skóre v programu PhenoGenius. Seznam genů si můžete stáhnout pomocí ikony Stáhnout nebo zkopírovat pomocí tlačítka Kopírovat/vložit.
	Umožňuje uživatelům filtrovat varianty pomocí odkazů na dědičnost z nemocí souvisejících s lidskou lékařskou genetikou. Varianty lze filtrovat podle způsobů přenosu (dominantní, recesivní, jiné) a zdrojů přenosu (OMIM, GenCC nebo jiné databáze).

	Umožňuje výběr variant podle heterozygotního složeného (předpokládaného nebo ověřeného), homozygotně recesivního nebo de novo způsobu přenosu na základě DiagAI. Upozorňujeme, že filtry "Compound Het - Verified", "Recessive Hom" a "De Novo" jsou k dispozici pouze v rodinných analýzách.
---	---

Sekce Classification

Název a přehled filtru	Description
VKB ⓘ NewUnknown variant 5Pathogenic 4Likely Pathogenic 3Uncertain Significance 2Likely Benign 1Benign Possible artifact ☐ Analysis only ⓘ	Umožňuje vybrat neznámé varianty, které nebyly vyhodnoceny (New - Unknown variant) nebo ověřeny ve VKB na všech analýzách subjektu. Pokud je zaškrtnuto políčko "Analysis only", budou varianty vyhledávány pouze na analýze, která je právě anotována.
ACMG ⓘ 5Pathogenic5 4Likely Pathogenic1 3Unknown significance (VUS)56 2Likely Benign49 1Benign6	Pathogenicity classification' pro zárodečné varianty, vypočtená na stupnici od 1 (benigní) do 5 (pathogenic).
ComPerMed ⓘ 5Pathogenic7 4Likely Pathogenic6 3VUS102 2Likely Benign1 1Benign7	"Select" somatic variants according to their ComPerMed pathogenicity class on a scale of 1 (benign) to 5 (pathogenic).
Evidence Type (CKB) ⓘ Actionable1346 Diagnostic0 Prognostic0	Vybírá varianty podle 3 typů důkazů (CKB): Akční, diagnostické, prognostické.

AMP/ASCO/CAP Tier ⓘ I II 1346 0	"Select" according to their assigned AMP level (levels I, II and N/A) in somatic context analysis only.
---	---

Sekce Scores

Název a přehled filtru	Description
CI-SpliceAI ⓘ all > 0.2 > 0.5 > 0.8 ☒ Keep unknown	Vyberte skóre delta. Pole "Keep unknown" ponechává varianty s hodnotou "N/A", ale vylučuje varianty bez vlivu na sestřih (varianta "No splicing impact").
Conservation ⓘ Low Conserved Highly Conserved	Stupeň ochrany lokality vypočtený z více skóre ochrany.
Metasvm ⓘ Tolerated Damaging Highly Damaging	Bere v úvahu skóre 10 aplikací pro predikci variant a maximální frekvenci pozorovanou v populaci 1000genomes.
Universal Pathogenicity Predictor (UP²) ⓘ Benign VUS Pathogenic -1 0 1 Min -1 Max 1	Vybírá varianty podle univerzálního prediktoru patogenity (UP²).
LOEUF ⓘ Intolerant Tolerant 0 1 2 Min 0 Max 2 ☒ Keep unknown	Vybírá varianty podle skóre LOEUF.

▼ CADD ⓘ Benign Pathogenic 1 99 Min 1 Max 99 ☒ Keep unknown	Vybírá varianty podle skóre CADD Phred.
▼ REVEL ⓘ Benign Pathogenic 0 0.5 1 Min 0 Max 1 ☒ Keep unknown	Vybírá varianty podle skóre REVEL.

Vyberte indikaci

V systémech SomaVar, SomaVar LF a SomaCGP je možné vybrat předdefinovanou patologickou indikaci z databáze JaxCKB z rozbalovací nabídky dostupné na *kartě* Patient v postranním panelu správy analýz (viz **část** Modální **zobrazení** pacienta).

Ontologický systém je založen na *ontologii nemocí*²⁴ a obsahuje termíny specifické pro JaxCKB. Informace o citlivosti v podsloupci 'A' sloupce Clinical Evidences (viz oddíl Stránka s variantami) **jsou** zvýrazněny zeleně pouze tehdy, pokud indikace pro danou variantu odpovídá vybrané indikaci. Pokud jsou pro variantu nalezena tvrzení o rezistenci, příslušná buňka ve sloupci "A" bude zvýrazněna červeně bez ohledu na indikaci. V opačném případě bude šedá.

Pokud bylo onemocnění vybráno během načítání vzorku a před analýzou (viz **Metadata vzorku**), stane se výchozí indikací definovanou pro vzorek a na kartě Patient bočního panelu správy analýzy. Je označen ikonou hvězdičky a v seznamu se zobrazuje jako první. Změna indikace v postranním panelu správy analýz neaktualizuje předem vybrané onemocnění pro vzorek (metadata onemocnění).

Analysis

Workflow Patient HPOs Panel

Active sample: CTDNA_toy

Sex

Female Male Unknown

Indication

★ Colorectal cancer (Colorectal cancer) × ▾

²⁴ <https://disease-ontology.org/>

Použití předdefinovaného profilu filtru

Uživatelé mohou vybrat předdefinovaný profil filtru z rozbalovací nabídky kliknutím na ikonu "Select profile". SeqOne platform nabízí ve výchozím nastavení několik předdefinovaných profilů filtrů: *GermVar*, *My Selection*, *Research*, *DiagAI*.

Úprava a ukládání profilů filtrů

Uživatelé mohou uložit aktuální polohu kurzorů pro každý filtr jako profil filtru tak, že profil pojmenují a kliknou na "Save as new". Profil filtru pak bude k dispozici prostřednictvím stejné rozbalovací nabídky v kategorii *Private*.

Při ukládání profilu filtru se otevře se formulář pro vytvoření entity a uživatel může vybrat tlačítko s názvem entity, aby ji zpřístupnil ostatním uživatelům entity. Pokud uživatel vybere tlačítko "Private" (Soukromý), bude profil přístupný pouze uživateli, který jej vytvořil.

Pouze uživatel, který profil filtru původně vytvořil, jej bude moci upravit kliknutím na tlačítko "Update"; pokud profil upraví jiný uživatel v entitě, bude vyzván, aby jej uložil pod jiným názvem. Pokud uživatel nemá oprávnění k úpravě veřejného profilu filtru, bude vyzván, aby kontaktoval správce svého subjektu. Profily filtrů SeqOne nelze upravovat a ukládat pod jiným profilem filtru.

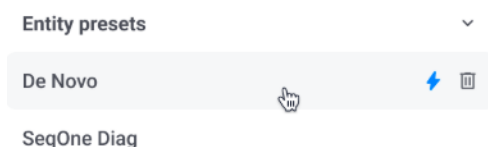
Lze resetovat jeden nebo více filtrů. Stačí kliknout na ikonu  vpravo od filtru nebo v pravém horním rohu pole "Filters".

Oblíbené profily filtrů

Uživatelé si mohou nastavit profily filtrů jako oblíbené, a to kliknutím na ikonu  vpravo od profilu, takže se zobrazí jako zástupce přímo v panelu rychlého přístupu nad tabulkou variant.

Chcete-li profil filtru odebrat z oblíbených položek, stačí

kliknout na stejnou ikonu  . Ve výchozím nastavení je profil *My Selection* definován jako oblíbený.



Kombinace profilů filtrů

Profilů filtrů lze kombinovat pod podmínkou **NEBO** kliknutím na kartu Kombinace na panelu filtrů a poté na "Create a combo". Ve výchozím nastavení se ve formuláři pro vytvoření zobrazí 2 skupiny profilů filtrů, ale uživatelé mohou přidat další kliknutím na "Add a group".

Po výběru profilů jsou uživatelé vyzváni, aby si svůj kombinovaný profil uložili kliknutím na "Save as new". Stejně jako u profilů filtrů může uživatel ve formuláři pro vytvoření zpřístupnit svůj kombo profil všem uživatelům subjektu. Pouze uživatel, který kombi profil původně vytvořil, jej bude moci upravit kliknutím na "Update".

Pokud je profil upraven jiným uživatelem subjektu, bude tento uživatel vyzván, aby jej uložil pod jiným názvem. Pokud uživatel nemá oprávnění upravovat veřejný kombo profil, bude vyzván, aby kontaktoval správce svého subjektu.

The screenshot shows a web interface for managing filters. At the top, there's a 'Filters' header with a count of 2 and a refresh icon. Below it, there are two tabs: 'Basic' and 'Combo', with 'Combo' being the active tab. A search bar is located below the tabs. Under the search bar, it says 'Active groups 2' and there's a 'Save as new' link. The main content area shows two groups of filter profiles. 'Group 1' contains a dropdown menu with 'SeqOne-diag' selected. Below this is an 'OR' button. 'Group 2' contains a dropdown menu with 'Filter_ACMG' selected. At the bottom of the groups, there is an 'Add a group' button.

Profilů filtrů SeqOne

My Selection. Tento profil filtru umožňuje, aby se varianty vyhodnocené jako 3 (VUS), 4 (Likely Pathogenic) a 5 (Pathogenic) zobrazovaly přímo ve VKB v interpretované analýze.

DiagAI. **RUO** Tento profil filtru, který je k dispozici v zárodečných analýzách (GermlineVar, GermlineFamily, GermVar Tertiary, GermVar Tertiary Family), zobrazuje redukováný seznam variant s vysokou pravděpodobností, že jsou příčinou onemocnění pacienta. Tento profil využívá algoritmus prioritizace DiagAI a výskyt varianty v projektu (filtr Frekvence vzorků).

Lze ji použít, pokud jsou nebo nejsou se vzorkem spojeny výrazy HPO, ale její citlivost se zvyšuje v přítomnosti výrazů HPO. V případě rodinné analýzy se profil filtru automaticky přizpůsobuje velikosti analyzované rodiny, protože bere v úvahu, že výskyt varianty v projektu může být větší. Uživatelé mohou filtr Frekvence vzorků v projektu také kdykoli upravit.

Zadejte HPOs a zobrazte varianty SmartPick s nejvyšší pravděpodobností, že způsobují onemocnění.

Family profiles. V analýzách GermlineFamily a GermVar Tertiary Family budou profily filtrů De Novo (novinka), Recessive Hom (novinka) a Compound Het (novinka) k dispozici pouze pro analýzy trojic zahrnující postižený indexový případ, nepostiženého otce a nepostiženou matku. V opačném případě budou tyto profily deaktivovány.

Při vytváření a ukládání osobního profilu se nedoporučuje upravovat *profily filtrů De Novo (nový)*, *Recessive Hom (nový)* a *Compound Het (nový)*. Při vytváření varianty těchto profilů se doporučuje začít od začátku.

Carrier screening. RUO Tento profil filtru je v GermlineFamily k dispozici pouze pro duo analýzy (rodiče bez indexového vzorku) a identifikuje recesivní geny s potenciálně patogenní heterozygotní variantou u obou "nepostižených" rodičů pro prekoncepční testování.

Somatic FFPE. Tento profil filtru zobrazuje redukovaný seznam somatických variant v somatických analýzách (SomaVar, SomaCGP) provedených na FFPE DNA.

Somatic LF. Tento profil filtru zobrazuje redukovaný seznam somatických variant v analýzách SomaVar LF.

Označit jako zobrazenou funkci

"Mark as viewed" je funkce, která umožňuje uživateli selektivně skrýt určité varianty. K tomu stačí vybrat požadovanou variantu (varianty) a kliknout na ikonu  Mark as viewed zobrazenou ve spodní části tabulky variant. Výběr variant se poté zobrazí šedě. Pro odstranění výběru může uživatel buď kliknout na ikonu  Reset all viewed zobrazenou v horní části tabulky variant, čímž zobrazení obnoví, nebo vybrat variantu (varianty) a kliknout na  Mark as not viewed.

Uživatel může vyskakovací okno skrýt kliknutím na křížek. Tím se zruší výběr všech variant, které však zůstanou šedé.

Výběr variant

Varianty lze přidat do panelu "Selection" panel všech vyhodnocených variant (viz část níže). Chcete-li tak učinit, vyberte varianty zaškrtnutím políček na levé straně tabulky variant a poté klikněte na ikonu „Přidat do výběru“ v dolní části tabulky.

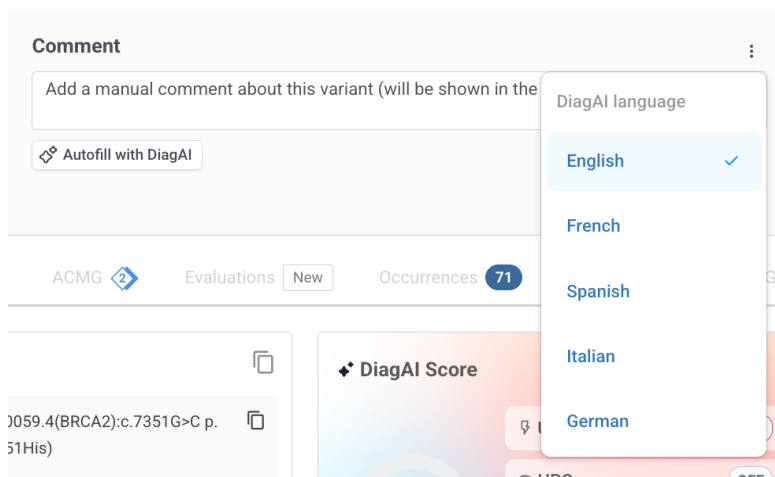
Hodnocení variant

Varianty lze rychle vyhodnotit z tabulky variant kliknutím do sloupce VKB a zadáním jejich třídy patogenity a volitelného komentáře. Tato hodnocení a komentáře jsou přístupné všem

uživatelům platformy v rámci jednoho subjektu prostřednictvím ikony "Selection" v analýze (viz oddíl Panel **pro správu hodnocení**). Variant: lze také hodnotit z její vyhrazené stránky.

Komentáře lze zadávat ručně nebo je generovat automaticky. Rozhraní obsahuje prázdné pole "Classification" pro komentáře a tlačítko "Autofill with DiagAI" pro automatické generování. Během generování komentářů se zobrazuje indikátor. U stávajících komentářů umožňuje možnost "Improve with DiagAI" jejich vylepšení. Každý vygenerovaný komentář lze schválit, upravit nebo resetovat. Po schválení se komentáře uloží s třídou patogenity, datem a obsahem. Po validaci je lze upravit nebo odstranit, přičemž zůstane zachována jejich historie.

Chcete-li vybrat jazyk pro DiagAI, klikněte na tři tečky v pravém horním rohu pole pro komentáře a z rozevřacího menu vyberte preferovaný jazyk.

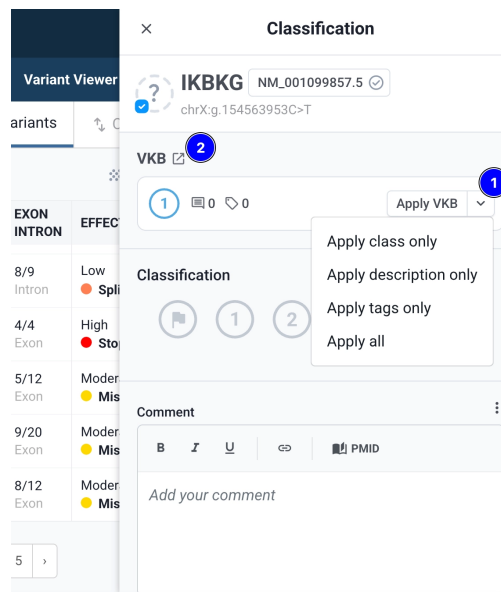


Varianty lze také klasifikovat kliknutím do sloupce klasifikace, jehož záhlaví má vlaječku. Chcete-li variantu klasifikovat, klikněte na ikonu +. Otevře se postranní panel klasifikace, který uživateli umožňuje přidat třídu patogenity, komentář a tag.

Chcete-li tuto klasifikaci přidat do databáze VKB, aktivujte přepínač "Add to VKB". Pokud již klasifikace v databázi VKB existuje, kliknutím na "Apply VKB" ji aplikujte na aktuální záznam. Další možnosti jsou k dispozici kliknutím na šipku vedle "Apply VKB" (viz krok 1 na obrázku níže).

Klasifikaci potvrďte kliknutím na "Save" v pravém horním rohu postranního panelu.

Chcete-li upravit klasifikaci přímo v databázi VKB, klikněte na přesměrovací odkaz VKB (viz krok 2 na obrázku níže) a otevřete okno "VKB Entry".

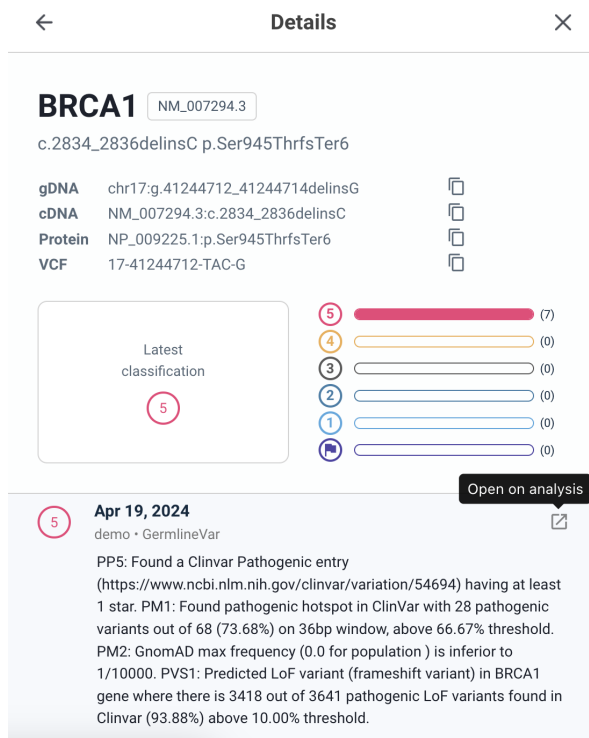


V okně "VKB Entry" klikněte na "Edit" vpravo nahoře pro úpravu hodnocení, poté klikněte na "Save" pro uložení změn. Historie VKB a výskyty varianty jsou přístupné procházením vyhrazených karet.

VKB Search. Boční panel je k dispozici po kliknutí na tlačítko "VKB" ve vyhledávacím panelu v záhlaví SeqOne. Ten umožňuje vyhledávat anotované varianty v databázi VKB podle názvu genu, transkriptu, anotace HGVS nebo identifikátoru SeqOne (Genome_variant_chromosome_location_REF_ALT). Vyhledávání lze filtrovat podle referenčního genomu (GRCh37 nebo GRCh38) a/nebo třídy patogenicity.

Obrázek 15: VKB Search.

Po kliknutí na hodnocenou variantu se otevře podrobnější stránka s informacemi o variantě, jako je její genomická pozice, její anotace HGVS na komplementární DNA (cDNA) a na proteinu, třídy patogenity, analýzy a související komentáře, kde byla hodnocena, a poté odkaz na analýzy.



Obrázek 16 : Okno s podrobnostmi o variantě ve VKB

Stránka variant

Kliknutím na variantu v tabulce variant se otevře postranní panel s několika záložkami, který obsahuje další informace. Uživatelé mohou otevřít stránku s variantami na jiné kartě kliknutím na "Open in new window" v pravém horním rohu. Boční panel lze zavřít pomocí klávesy "Esc".

Uživatelé mohou v tomto okně hodnotit a komentovat variantu. Stejně jako hodnocení a komentáře provedené z hlavní tabulky variant mohou tato hodnocení a komentáře vidět všichni členové skupiny používající platformu.

Na této stránce mohou uživatelé také získat přístup k informacím o variantách, včetně genotypu (GT), alelické frekvence (VAF), hloubky pokrytí (COV) a phred skóre (QA).

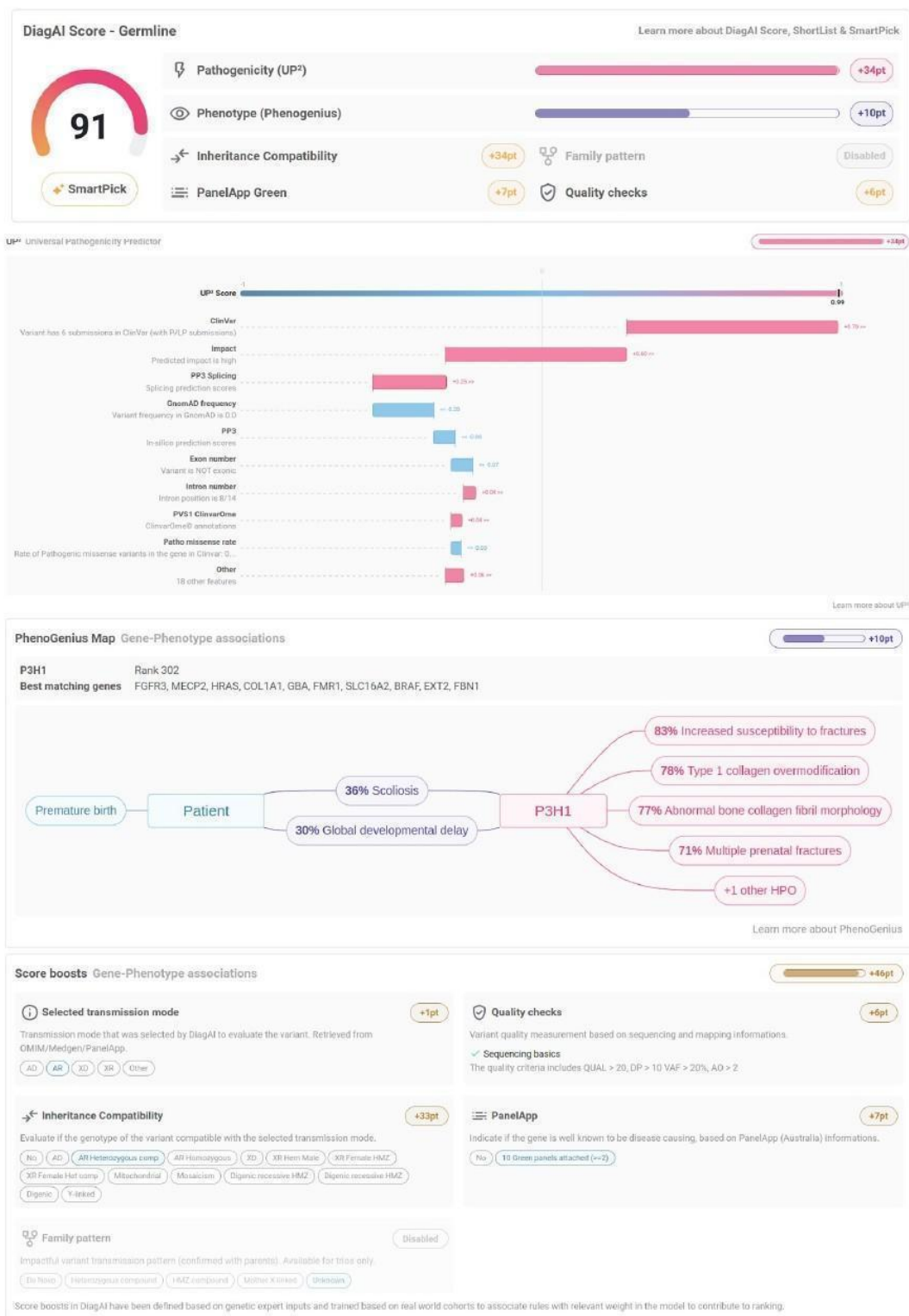
Karty na stránce Variant jsou popsány níže.

Variant. Karta je rozdělena do několika bloků, které jsou popsány níže.

- **Nomenclature:** jsou uvedeny různé nomenklatury (g., c., p., položka VCF) s možností kopírování informací z celého bloku nebo z každého řádku do schránky pomocí ikony .
- **DiagAI Score:** používá blok pro souhrnný výpočet skóre DiagAI, který je k dispozici na kartě Skóre DiagAI na stránce varianty (viz odstavec DiagAI Score).
- **ClinVar:** informace převzaté ze sloupce ClinVar Germline tabulky variant, které uvádějí třídu patogenity přiřazenou variantě, její klinický stav a počet záznamů v ClinVar (viz Vizualizace a **interpretace v části Malé varianty**). Ikona  otevře stránku ClinVar na nové kartě.
- **Annotations:** dopad transkriptu, příslušný exon, vzdálenost od nejbližšího místa sestřihu a typ mutace.
- **Transmission:** odkazy na databáze OMIM a GenCC, pokud je s genem spojeno jedno nebo více onemocnění.
- **Složení Heterozygous Candidates,** tento blok je rozdělen na 2 části:
 - "Short Variants" seskupuje všechny kandidátní SNV nacházející se ve stejném genu (na základě ID genu) a řadí je podle jejich skóre DiagAI. Variant. vybraný na stránce Variant. se zobrazí jako první a v seznamu je šedá.
 - V části CNVs je uvedeno, zda byly nalezeny CNVs, seřazené podle třídy ACMG. Tento oddíl je k dispozici pouze v případě, že byla provedena analýza CNV, jinak se zobrazí varovná zpráva "CNVs not available" (CNVs nejsou k dispozici).

Kliknutím na jeden z výsledků se otevře nová karta se stránkou varianty (pokud se jedná o malou variantu) nebo stránkou CNV (pokud se jedná o kopírovanou variantu). Pokud nejsou nalezeny žádné CNV, zobrazí se varovné hlášení se slovy "no CNVs found".

- **Tools:** odkazy na další databáze (LitVar, LOVD, VarSome, dbSNP, Cosmic, GnomAD, CKB, UCSC, Decipher, Ensembl).
- **GnomAD:** pokud je varianta v této databázi známa, zobrazí se tabulka shrnující četnost varianty v různých populacích a přímý přístup do GnomAD je umožněn pomocí ikony .



Obrázek 18: Vizualizace karty DiagAI na stránce varianty

DiagAI. Karta obsahuje vysvětlení DiagAI Score přiřazeného variantě. Tato karta je k dispozici pouze pro analýzy GermlineVar, GermlineFamily, GermVar Tertiary a GermVar Tertiary Family (od verze 1.8).

Karta DiagAI je rozdělena do 4 hlavních částí:

- **DiagAI Score - Germline** Summary Block Showing the Score, SmartPick Status, and Listing the Main Characteristics Contributing to the Overall Score and the Associated Number of Points.
- Vodopádový graf vysvětlující **skóre UP² (Universal Pathogenicity Predictor)**.

Varovná zpráva  se zobrazí, když uživatel vyhledá stránku varianty, jejíž transkript se liší od transkriptu použitého pro výpočet UP².

- Blok vysvětlitelnosti **mapy PhenoGenius**, který je barevně označen, pokud jsou ke vzorku přiřazeny termíny HPO. Blok obsahuje podrobnosti o pořadí genu, seznam 10 nejlépe asociovaných genů a mapu PhenoGenius, barevně označenou podle shody mezi HPOs pacienta a genem: modrá barva pro HPOs pouze pacienta, fialová barva pro HPOs pacienta a odpovídající genu, červená barva pro HPOs pouze genu. Sekce je šedá, pokud pro pacienta nejsou zadány žádné termíny HPO.
- Sekce **Score boosts** zahrnuje 5 dílčích funkcí, které mohou, ale nemusí přispívat (Disabled) ke skóre: Režim Transmission: Kompatibilita dědictví, Rodinný vzor, PanelApp a Kontrola kvality.

ACMG. Karta umožňuje uživatelům zobrazit a zkontrolovat kritéria použitá při automatické klasifikaci ACMG. Mohou upravit výběr zaškrtnutím různých požadovaných kritérií a poté použít klasifikaci k hodnocení varianty kliknutím na "Use for Evaluation" (*Použít k hodnocení*). Je také možné upravit úroveň důkazů (*podpůrná*, *střední*, *silná* a *velmi silná*) a přiřadit každému kritériu jiný počet bodů, které budou zohledněny při výpočtu konečného skóre ACMG. Za každou úroveň důkazů bude uděleno: 1 bod za *podpůrné*, 2 body za *středně silné*, 4 body za *silné* a 8 bodů za *velmi silné*. Body za jednotlivá kritéria se poté sečtou a variantě se přidělí status ACMG:

- *Pathogenic*: 10 bodů
- *Likely Pathogenic*: 9 až 6 bodů
- *Uncertain Significance*: 5 až 0 bodů
- *Likely Benign*: -6 až -1 bod
- *Benign*: od -7 bodů

U kritéria BA1 nemůže uživatel zvolit úroveň důkazu, protože se jedná o samostatnou třídu.

TP53 protein coding

NM_000546.6 c.1083del p.S362AfsTer8 ● Frameshift

sample
GT / VAF ☐ 1.25%
COV / QA 319 34

1 2 3 4 5 Add a manual comment about this variant (will be shown in the output report) ... Add comment Cancel

Variant **ACMG** TP53 Compermed Evidences Evaluations New Occurrences 6 Scores IGV Gene

☒ Use ClinGen point-based system

Predicted:
Likely Pathogenic Use for evaluation

Automatic tags: PVS1 | Strong PM2 | Moderate PP5 | Moderate

First level tags: No selected tags.

Second level tags: No selected tags.

Tags that do not apply: No selected tags.

Automatic tags (3 / 3)
These tags are automatically selected.

☒ **PVS1** Strong ⓘ

Predicted null variant in a gene where LOF is a known mechanism of disease
Non-sense or frameshift variant, NOT predicted to undergo NMD, but truncated region is critical to protein function (11 non-truncating pathogenic variants found in Clinvar).

Obrázek 19: Vizualizace karty ACMG na stránce Variant.

TP53. Záložka je k dispozici pouze v případě, že je varianta anotována v genu TP53. Tato záložka poskytuje informace o mutacích genu TP53 a jejich klinickém významu podle databáze IARC²⁵. Karta TP53 je rozdělena do několika bloků, které jsou popsány níže.

Blok anotací (*Annotations*) zobrazuje následující informace spojené s variantou:

- Mutation Rate: míra substituce spojená s pozicí varianty
- Hotspot (yes/no): přítomnost varianty v místě opakované mutace (hotspot).
- Structural Motif: strukturní motiv ovlivněný mutací
- Domain: funkční doména proteinu ovlivněná mutací.

Blok Funkční údaje (*Functional Data*) zobrazuje klasifikace a funkční informace spojené s variantou:

- DNE Class: dominantně negativní účinek (DNE) na transaktivaci divokého typu proteinu p53
- DNELof Class: funkční klasifikace pro ztrátu potlačení růstu a negativní dominantní aktivity
- Transactivation Class: Funkční klasifikace založená na globální transkripční aktivitě (TA)
- Structure Function Class: Functional Prediction.

²⁵ <https://tp53.cancer.gov/>

Bloky "Graphics" a "Sample Details" poskytují podrobnosti o vzorcích, na které se odkazuje v databázi IARC pro danou variantu. Uživatel si může vybrat, zda chce zobrazit údaje spojené se somatickými nebo zárodečnými vzorky, a to výběrem možnosti Somatické nebo Zárodečné.

Blok "Graphics" zobrazuje dva grafy s rozbalovacími nabídkami, které ukazují rozdělení vzorků podle jejich vlastností.

Blok Podrobnosti o vzorku zobrazuje podrobné informace o vzorku v tabulkové formě.

ComPerMed. (pouze somatická analýza). Třída patogenity každého Variantu je vypočítána podle doporučení belgické odborné komise ComPerMed (Komise pro personalizovanou medicínu). Uživatel má přístup k podrobným informacím o této klasifikaci ze speciální záložky na stránce Variant.

Evidences. Na kartě je uveden seznam terapií spojených s vybranou variantou z databáze JaxCKB. Uživatel může procházet jednotlivé řádky a najít podrobnosti o terapiích nebo indikacích.

BRCA2 protein coding

NM_000059.4 c.1310_1313del p.K437IfsTer22 frameshift

sample

GT / VAF 43.7%

COV / QA 126 29

1 2 3 4 5

Add a manual comment about this variant (will be shown in the output report) ...

Add comment Cancel

Variant ACMG 5 Compermed 6 **Evidences** Evaluations New Occurrences 6 Scores IGV i Gene

© The following information is provided for informational purposes only

Actionable 25 Prognostic & Diagnostic 0 Risk factor & emerging

Therapy	Tier	Response	Indication
Olaparib	IA	Sensitive	Breast cancer Her2-receptor negative breast cancer ✓FDA +6
Select indication Breast cancer Evidence type Actionable Molecular profile BRCA2 inact mut Approval N/A Drugs Olaparib (PARP Inhibitor (Pan)) Description Lynparza (olaparib) is included in guidelines as systemic treatment for patients with recurrent or metastatic breast cancer harboring BRCA1/2 germline mutations (NCCN.org).			
Rucaparib	IA	Sensitive	Ovarian cancer ✓FDA Prostate cancer ✓FDA +3

Obrázek 20: Vizualizace karty Evidence na stránce varianty

Evaluations. Tato karta umožňuje přístup k historii hodnocení a interpretací varianty v průběhu času:

SEQ-027_IFU_CZ_V6

Zveřejněno 2026-04-29

130/182

- V aktuální analýze (levá část karty)
- V dalších analýzách prováděných v rámci skupiny (pravá část karty)

Occurrences. Karta identifikuje ostatní vzorky obsahující danou variantu bez ohledu na hodnocení. Pro každý výskyt se zadává název vzorku, alelická frekvence, odkaz na příslušnou analýzu a název projektu.

Publications (Publikace). Tato karta poskytuje přístup k vědecké literatuře spojené s variantou a genem. Kliknutím na dílčí kartu "Variant" (Varianta) nebo "Gene" (Gen) zobrazíte publikace PubMed odkazující na variantu nebo gen. Odkaz PMID přesměruje na záznam PubMed a odkaz na plný text otevře celý článek, pokud je k dispozici.

Scores. Záložka je k dispozici pro varianty s chybějícím, ztraceným startem a získaným stop efektem. S výjimkou těchto případů bude skóre šedé. Tato karta umožňuje zobrazit výsledky predikčních skóre několika nástrojů, jako jsou *REVEL*, *DANN*, *FATHMM*, *FATHMM-MLK*, *LRT*, *MetaLR*, *MetaSVM*, *MutationTaster 2*, *PROVEAN*, *SIFT* a *PLI*. Poskytuje také přístup k skóre zachování, jako jsou *GERP++*, *PhastCons100way Vertebrate*, *PhastCons20way Mammalian*, *PhyloP100way Vertebrate*, *PhyloP20way Mammalian* a *SiPhy*.

IGV. Tato záložka umožňuje uživateli zobrazit reads zarovnané k pozici varianty prostřednictvím webové verze prohlížeče genomu IGV. V oblastech s vysokým pokrytím čtení jsou čtení ve výchozím nastavení podvzorkována, tj. nejsou zobrazena všechna čtení. Jsou použity výchozí parametry prohlížení, tj. maximálně 100 reads na 50nukleotidové okno. Tyto oblasti lze identifikovat podle černé čáry pod profilem pokrytí (downsampling).

Po kliknutí na ni se zobrazí počet reads odstraněných ze zobrazení (# downsampling). Z ikony  je k dispozici nabídka, která nabízí tři možnosti:

- *Otevřít na nové kartě:* otevře zobrazení varianty IGV na nové kartě.
- *Track setting:* umožňuje přidat do zobrazení další vzorky ze stejného nebo jiného projektu.
- *Open local IGV:* přesměruje aplikaci IGV Desktop na souřadnice aktuálně interpretované varianty. Chcete-li zobrazit SNV nebo jiné varianty, musí být vzorek BAM nejprve načten do aplikace IGV Desktop. Tato možnost je kompatibilní pouze s aplikací IGV Desktop verze ≥ 1.7 .

IGV podporuje rozdělené zobrazení, například zkopírováním dvou sad souřadnic do panelu pro vyhledávání lokusů oddělených mezerou. Různá nastavení jsou k dispozici po kliknutí na ozubená kolečka vpravo od prohlížeče genomu. Chcete-li například přímo zobrazit nezarovnané čtení (soft clips), vyberte možnost "Show soft clips".

Gene. Karta obsahuje další informace o genu a související bibliografii.

Vytvoření klinické zprávy

Uživatel může upravovat exportovatelnou klinickou zprávu ve formátu PDF a .docx kliknutím na tlačítko "Generate report" (Generovat zprávu) na kartě "Selection" (Výběr) při procházení analýzy.

V horní části stránky pro úpravy může uživatel vybrat šablonu, nastavit stav zprávy na novou nebo koncept, vybrat vzorek (pro rodinné analýzy lze vybrat více vzorků) a vybrat jazyk úprav. Uživatel může také přepínat mezi zobrazením "Edition" (Úprava) a zobrazením "Preview" (Náhled) před uložením konceptu nebo přechodem na další krok pomocí tlačítek "Save Draft" a "Next Step".

Šablony zpráv lze vytvářet v sekci "Reporting", pod "Template list" v nastavení subjektu (Entity settings).

V zobrazení pro úpravy může uživatel upravovat různé části zprávy a vybírat více možností. V zobrazení náhledu může uživatel zkontrolovat zprávu před pokračováním v úpravách, uložením konceptu nebo ověřením zprávy.

Uživatel může uložit koncept a znovu k němu přistoupit z karty "Selection", a poté kliknout na "Generate report", stejně jako při vytváření nové zprávy. Po vygenerování zprávy ve formátu PDF se tato zpráva automaticky zobrazí na kartě "Reporting" (Zprávy) příslušné analýzy. Až po ověření lze zprávu exportovat jako .PDF nebo .DOCX. Uživatel může zprávu smazat nebo vytvořit další zprávy pro stejnou analýzu.

Obrázek 21 : Zobrazení karty klinické zprávy.

13.1.5. CNV

13.1.5.1. CNV v panelech genů

Metodika

Detekce CNV v datech genového panelu, v programech GermlineVar, GermVar/GermVar Family, SomaVar a SomaHemato je založena na výběru kontrolních vzorků z koanalyzované kohorty podle jejich korelace.

Korelace se vypočítá z tabulky pokrytí, která se zase vypočítá ze všech oblastí v manifestu, ale bez pohlavních chromozomů.

Kritéria pro to, aby byl vzorek způsobilý pro analýzu CNV, jsou následující:

- Aby byla oblast zachována v analýze CNV, je minimální pokrytí 50 reads.
- Prahová hodnota korelace pro výběr kontrolních vzorků je 0,90 pro zárodečný kontext a 0,75 pro somatický kontext.
- Minimální počet kontrol požadovaný pro každý vzorek je šest.

Před spuštěním CNVCapture nebo SomaCNVCapture je třeba definovat pohlaví vzorků (metadata vzorku), aby bylo možné provádět volání genů lokalizovaných na pohlavních chromozomech. A konečně, k provedení analýzy CNVCapture nebo SomaCNVCapture v systému SeqOne je zapotřebí alespoň osm vzorků.

Kritéria, která se uplatňují na uživatelský manifest, aby byl gen panelu způsobilý pro analýzu CNV, jsou následující:

- manifestní oblasti jsou anotovány na základě kanonického transkriptu určeného podle zdroje NCBI MANE Select. Pokud není k dispozici žádný transkript MANE Select, je anotace založena na zdroji NCBI RefSeq.
- analyzovány jsou pouze exony, u nichž se manifestní interval a exon překrývají na $\geq 80\%$. Pokud se žádný exon genu nepřekrývá $\geq 80\%$ s manifestním intervalem, nebude gen do analýzy CNV zahrnut.
- analyzovány jsou pouze introny, u nichž se manifestní interval a intron překrývají z $\geq 60\%$.

CNVcapture Analysis (verze ≥ 2.4) zahrnuje oblasti nacházející se 20 kb před a za geny, na které se vztahuje manifest, pokud jsou tyto oblasti zahrnuty v návrhu manifestu. Konkrétně je analyzován každý z 20 segmentů o délce 1 kb tvořících oblast 20 kb, pokud se alespoň 20 % segmentu překrývá s původním manifestem.

Zobrazení a výklad

Na kartě CNVs se zobrazí tabulka se seznamem genů, u nichž byly identifikovány odchylky v kopiích. Tato karta je k dispozici v pracovních souborech popsanych v části **Metodika** výše.

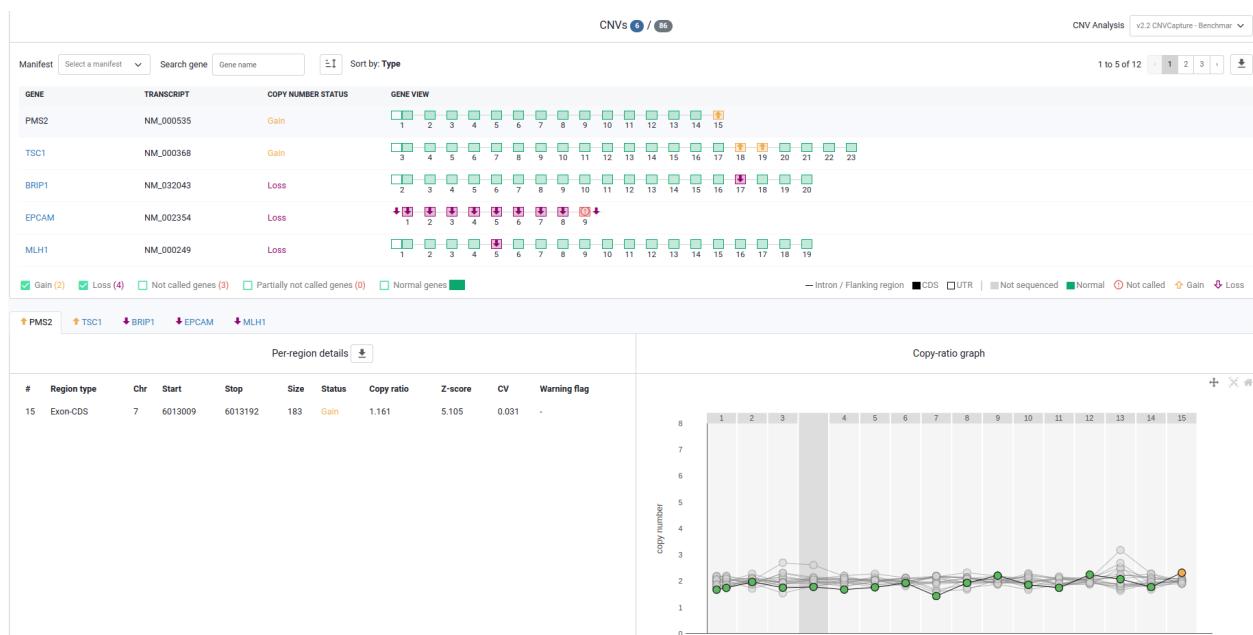
Terminologie používaná k popisu stavu volaných genů je následující:

- Gain: Gene je ovlivněn částečnou nebo úplnou amplifikací.
- Loss: v genu je částečná nebo úplná delece.
- Normal: Gene nenese žádnou CNV
- Partially not called: Gene je částečně zavolán
- Not called: gen je z analýzy CNV vyloučen.

Partially not called: je genu přiřazen, pokud je stav jeho kopie normální, ale alespoň jeden exon nebyl zavolán (se stavem "Failed").

Navzdory způsobilosti pro CNV analýzu je gen z analýzy vyloučen (Not called), pokud :

- Gene je na pohlavním chromozomu a metadata o pohlaví nejsou vyplněna,
- všechny exony Failed podle níže uvedených kritérií: Number of controls, percentage of variation. Status: "Failed" status genu je definován pouze na základě CDS. Intronicke a upstream/downstream oblasti jsou vyloučeny, pokud jsou volány exony CDS.



Obrázek 22: Vizualizace karty CNVs v analýze GermlineVar.

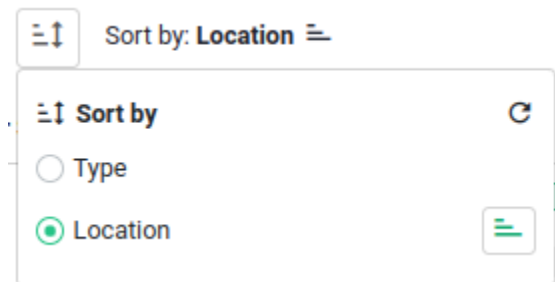
Filtry a možnosti třídění

Nabídku filtrů lze použít k:

- Omezení zobrazení na podmnožinu oblastí/geny pomocí souboru BED, který byl předtím nahrán do rozhraní pomocí funkce "Manifest".
- Vyhledání konkrétního genu pomocí funkce "Search Gene"

K dispozici jsou dvě možnosti třídění CNV:

- **Type** (default): logika třídění proudu (Gain a pak Loss)
- **Location**: slouží k seřazení CNV podle chromozomů a jejich výchozí pozice, vzestupně nebo sestupně, kliknutím na ikonu řazení.



Uživatel může také vybrat geny podle jejich statusu: Genes may be selected according to state: Gain, Loss, Normal genes, Partially not called genes or Not called genes.

☒ Gain (0) ☒ Loss (0) ☐ Not called genes (69) ☐ Partially not called genes (0) ☐ Normal genes

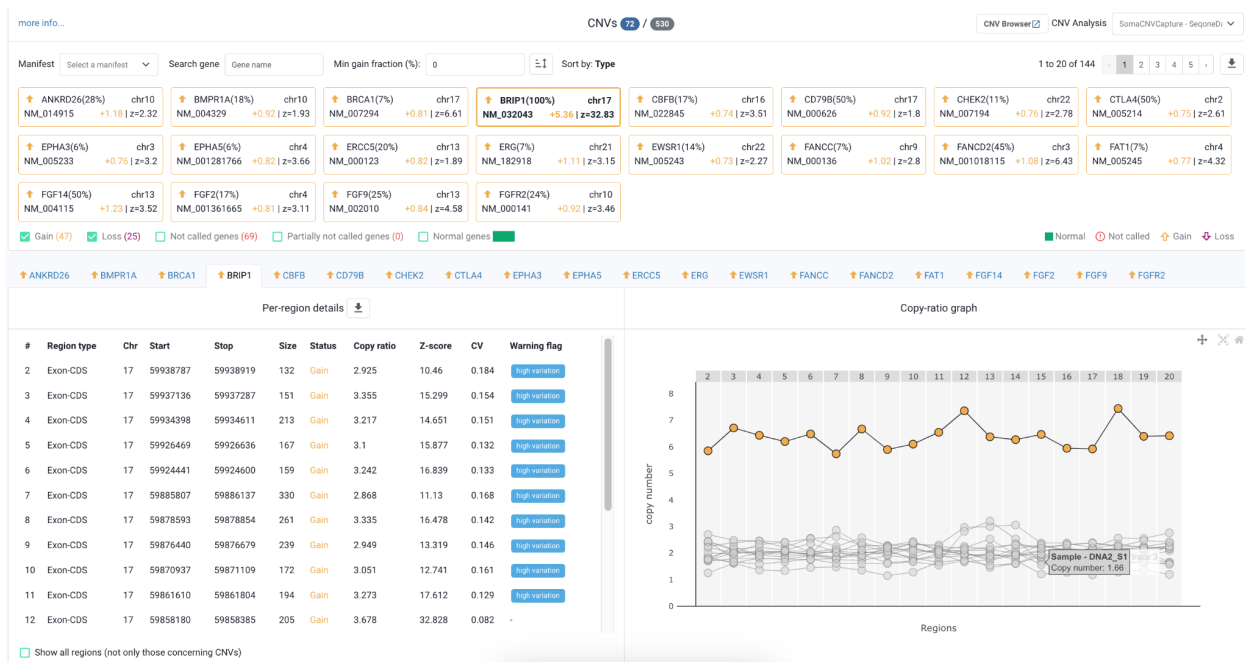
Podrobnosti podle regionu

Tabulka s podrobnostmi pro jednotlivé oblasti zobrazuje podrobnosti pro každou oblast pro každý Gene:

- #: číslo oblasti (exony a introny)
- **Region type**: typ regionu (upstream, utr-exonic, coding-exonic, intronic, downstream).
- **Chr**: chromozom
- **Start/Stop**: počáteční a koncová poloha CNV
- **Size**: velikost CNV (bp)
- **Status**: stav oblasti (gain = duplikace, ztráta = delece, normální = žádná CNV, nevolat = oblast nebyla zahrnuta do analýzy CNV).
- **Copy-ratio**: poměr počtu kopií k normálu (n=2).
- **Z-score**: počet směrodatných odchylek, které dělí výsledek od průměru. Vysoké kladné nebo záporné hodnoty ukazují na silnou důvěru ve volání CNV.
- **CV**: koeficient variability měřící relativní rozptyl v rámci kohortních vzorků.
- **Warning flag**: Příležitostně se může zobrazit chybové hlášení podle následujících kritérií:
 - **Low zscore**: Absolutní hodnota Z-score: <4
 - **High variation**: koeficient variability (CV) > 0,1
 - **Low coverage**: pokrytí oblasti pro tento vzorek <100
 - **Number of controls**: Počet kontrol je nedostatečný, aby bylo možné provést volání v této oblasti.
 - **Percentage of variation**: Variance je příliš velká napříč vzorky na to, aby tato oblast byla schopna provést volání.

- **Read depth:** nedostatečné pokrytí oblasti (<50) a Copy-ratio >0,1
- **Missing Gender:** oblast se nachází na pohlavním chromozomu a metadata o pohlaví nebyla zadána.

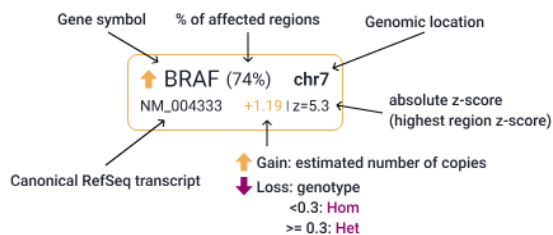
Warning: oblast bohatá na Soft clips, jako je velký indel nebo Alu inzerce, může být interpretována jako pokles pokrytí worksety CNVCapture a SomaCNVCapture a detekována jako ztráta kopie, pokud je velikost inzerce stejného řádu jako velikost volané oblasti.



Obrázek 23: Vizualizace karty CNVs v analýze SomaVar

V případě analýz **SomaCNVCapture** je sice metoda detekce podobná metodě CNVCapture, ale pro grafické znázornění variant se používá měřítko genu, nikoliv exonu. Zobrazení je znázorněno jako řada rámečků. To poskytuje symbol genu, genomickou pozici, kanonický Refseq transkript, absolutní z-score a v případě zisku počet kopií nebo v případě ztráty genotyp.

Vyskakovací okno s vysvětlením těchto polí je k dispozici nad filtry v části **"more info..."**.



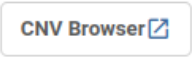
Procento zobrazené v každém rámečku vyjadřuje podíl genu (min *gain fraction*) ovlivněný zjištěnou CNV. Poměr odštěpení lze použít k filtrování výsledků. Ve výchozím nastavení je filtr nastaven na 0 % a vztahuje se pouze na zisky kopií.

Na vymazání zjištěná touto potrubní linkou se tento filtr nevztahuje a měla by být interpretována s opatrností. Schopnost tohoto nástroje detekovat změny kopie genu totiž závisí na povaze a kvalitě vzorku, zejména na frakci nádoru.

Warning: Pokud je u téhož genu zjištěna amplifikace i delecce, bude gen standardně považován za amplifikovaný, a proto se na něj bude vztahovat výše popsany filtr min. podílu zisku.

CNV Browser

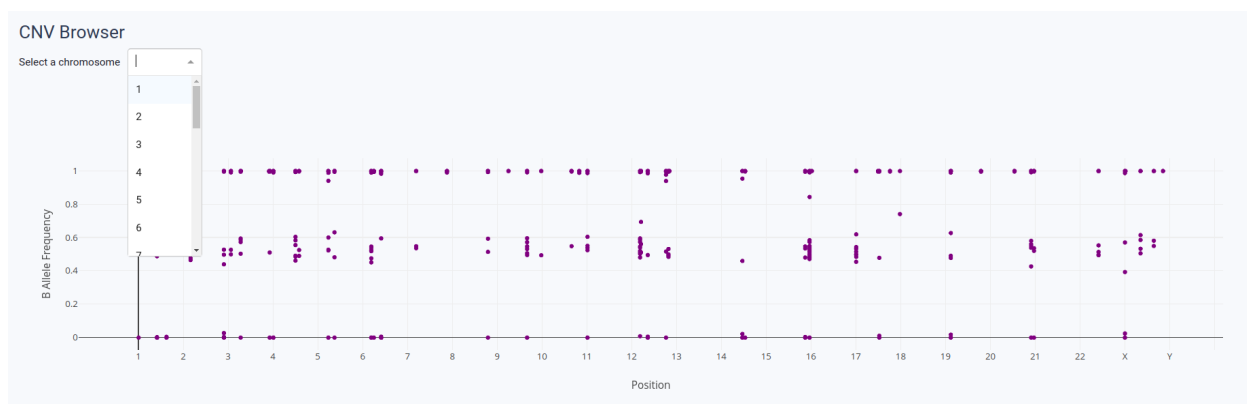
Na kartě zobrazení CNV v analýzách GermlineVar, GermlineFamily, SomaVar nebo SomaHemato

je k dispozici ikona  pro otevření prohlížeče CNV v nové kartě. Graf CNV Browser zobrazuje frekvenci alel B (BAF) SNP (Single Nucleotide Polymorphism) přítomných v manifestu analyzovaného vzorku a přítomných v datové sadě SNP array od společnosti Illumina a SNP z jiných páteří, jako je například OneSeq od společnosti Agilent, aby se usnadnila vizualizace určitých typů událostí, jako je LoH (Loss of Heterozygosity).

BAF je frekvence nereferenční alely (alela B) pozorovaná u zárodečného heterozygotního SNP.

Ve výchozím nastavení graf znázorňuje BAF variant (osa y) v celém genomu (osa x), rozdělený podle chromozomů (1-22, X a Y). Filtr umožňuje uživateli zobrazit varianty na daném chromozomu v levém horním rohu.

V kontextu analýz CNV lze tyto informace o BAF použít k prokázání změny počtu kopií specifických pro alelu, ztráty heterozygotnosti nebo alelické nerovnováhy.



Obrázek 24: Vizualizace karty CNV Browser v analýze GermlineVar.

Export dat CNV

Ikonu  vpravo nahoře lze použít k exportu dat analýzy CNV ve formě souboru .tsv obsahujícího následující sloupce:

- Gene:
- RefSeq transkript (refseq_id),
- Status: Status (Status),
- chromozom (chr),
- genomické souřadnice CNV: počáteční a koncová pozice,
- podíl genu ovlivněného CNV (gene_fraction_cnv),
- poměr počtu kopií (copy_ratio) k normálu (n=2),
- počet směrodatných odchylek, které dělí výsledek od průměru (zscore).
- počet exonů překrývajících CNV (exony) v případě analýzy CNVCapture.

Pro každý gen lze ikonou  vedle položky "Per-region details" použít ke konkrétnímu exportu dat analýzy CNV pro vybraný gen ve formě souboru .tsv obsahujícího následující pole:

- Gene.
- chromozom (chr),
- genomické souřadnice: počáteční a koncová pozice,
- číslo exonu nebo intronu (exon_intron_number),
- Status:
- Type (exon nebo intron, typ),
- související chybová hlášení (warning_flags, fail_flags),
- upravená kritická hodnota pravděpodobnosti (padj),
- kritická hodnota pravděpodobnosti (pvalue),
- poměr počtu kopií k normálu (ratio),
- míra relativního rozptylu v rámci vzorků (variabilita),
- standard deviation výsledku od průměru (zscore).

13.1.5.2. CNV v analýze SomaCGP a SomaLBx

Metodika

SomaCGP a SomaLBx používají PureCN²⁶ k detekci CNV a LOH (loss of heterozygosity). PureCN identifikuje somatické varianty z dat detekce variant za účelem posouzení ploidie a čistoty nádoru vzorku. Tyto informace se používají k určení počtu kopií každé události.

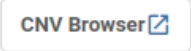
Vizualizace a interpretace

Karta vizualizace CNV v SomaVar Tertiary Analysis funguje stejně jako karta v SomaCGP a SomaLbx, jak je popsáno níže.

CNV jsou prezentovány v tabulce, kterou lze filtrovat. Každá CNV zobrazuje následující informace:

- **Gen:** Gene:
- **Location:** cytoband
- **Souřadnice události / Velikost události:** souřadnice a velikost události
- **Transcript:** Reference transcript
- **Exons:** Exons:
- **Type:** typ události (LoH, Loss or Gain)
- **Počet kopií:** absolutní počet kopií
- **Function:** funkce genu, je-li to relevantní, onkogen nebo Tumor Suppressor Gene (TSG)

Uživatel může vyhledat CNV ovlivňující daný gen pomocí vyhledávacího pole a zaškrtnutím políčka "Show genes in LOH regions" zvolit, zda se mají zobrazit geny přítomné v oblastech LOH. Tabulku výsledků lze exportovat pomocí ikony pro stažení v pravém horním rohu.

Ikona  otevře prohlížeč CNV Browser na nové kartě. Na této kartě se zobrazí graf zobrazující frekvenci alel B (BAF) SNP v analyzovaném vzorku a LOH SNP z vybraných publikací a páteře OneSeq Agilent. Druhý graf představuje data logRatio generovaná PureCN. Další podrobnosti jsou k dispozici v části **CNV v panelech genů** v části **CNV Browser**.

²⁶ <https://github.com/lima1/PureCN>

13.1.5.3. CNV v exomech a genomech

Karta vizualizace CNV pro exomová data je k dispozici v analýzách GermlineVar, GermlineFamily, GermVar a GermVarFamily při použití souborů FASTQ jako primárního vstupu a spuštění analýzy CNVExome. Tato karta také zobrazuje výsledky CNV pro data genomu v analýzách GermVar a GermVar Family při použití souborů FASTQ a v analýzách GermVar Tertiary a GermVar Tertiary Family při použití souborů VCF nebo dat CGH MicroArrays.

Metodika pro exome

Tato metodika se vztahuje konkrétně na nástroj CNVExome. Stejně jako u pipelines pro analýzu CNV na datech genových panelů závisí proces identifikace variant v tomto případě na analýze hloubky pokrytí. Analýza CNVExome pomocí [nástrojeGermlineCNVCaller](#) GATK je však individuální, protože každý vzorek je porovnáván s modelem HHMM (Hierarchical *Hidden Markov Model*) vypočteným nástrojem SeqOne.

Aby se zohlednily technické odchylky při přípravě vzorků, je tento model založen na velké kohortě od jednoho do několika stovek vzorků sekvenovaných laboratoří za použití:

- Stejný protokol přípravy knihovny
- Stejná sada pro zachycení
- Stejná sekvenační platforma a délka read

Tyto požadavky a proměnlivý počet vzorků potřebných k sestavení modelu lze vysvětlit potřebou kohorty, která nabízí homogenní úroveň pokrytí. Je třeba vzít v úvahu několik faktorů, jako je hloubka sekvenování, míra variability v rámci kohorty a reprodukovatelnost protokolů přípravy vzorků.

Před zahájením analýzy CNVExome je třeba vyplnit metadata o pohlaví (viz **Metadata vzorku**), aby bylo možné ověřit ploidii pohlavních chromozomů. Pokud není vyplněno, zobrazí se v rozhraní pro zobrazení CNV varovná zpráva "Ploidy".

Omezení: Dokumentace GATK GermlineCNVCaller doporučuje vyloučit oblasti PAR (pseudoautozomální oblasti), tj. oblasti PAR1 a PAR2, pokud je jejich podíl na celkovém pokrytí X neúměrný. Ve skutečnosti je ploidii pro konce chromozomu X odpovídající oblastem PAR u mnoha modelů vypočtených nástrojem SeqOne obtížné vypočítat. Výsledkem je, že ploidie není správně vypočtena pro tyto oblasti, které jsou detekovány arteficiálně v 1 kopii, zatímco celý chromozom X je detekován ve 2 kopiích.

Uživatelé se mohou obrátit na tým uživatelské podpory na adrese support@seqone.com, aby se dozvěděli více o postupu nastavení CNVExome ve své skupině.

Metodology pro genom

U verzí do v3.6 integrují analýzy GermVar a GermVar Family na genomových datech nástroj Sentieon²⁷ CNVScope pro detekci CNV. Od verze v3.7 používají worksets GermVar a GermVar Family nástroj Cue²⁸ pro volání CNV na genomových datech.

Data o segmentální duplikaci jsou detekována pomocí Sentieon's SegDup Caller, který je specificky navržen k identifikaci událostí počtu kopií ve vysoce homologních genomických oblastech na datech WGS pouze pomocí GermVar. Tento vyvolávač se zaměřuje na vybraný soubor klinicky relevantních genů, o kterých je známo, že jsou náročné pro standardní přístupy k detekci variant. Výsledné události jsou automaticky integrovány do analytického pipeline a vizualizovány v prohlížeči CNVs & SVs jako CNV (zisk i ztráta), což umožňuje identifikaci variant v genomických oblastech, které byly dříve pro standardní analýzu nepřístupné.

Zobrazení a výklad

Na kartě **CNV** je k dispozici interpretační rozhraní pro analýzy CNVExome, GermVar a GermVarFamily, GermVar Tertiary nebo GermVar Tertiary Family. Rozhraní je rozděleno na dvě části: vlevo je tabulka s výsledky CNV a vpravo je podrobná stránka CNV s několika záložkami.

Tabulka výsledků CNV

Tabulka výsledků zobrazuje identifikované CNV spolu s nabídkou filtrů založených na jejich přítomnosti v morbidních genech OMIM [<https://omim.org/>], velikosti, frekvenci výskytu CNV v kohortě nebo LogRatio v případě analýzy MicroArray CGH.

V horní části tabulky CNV se zobrazí počet filtrovaných CNV (vybrané CNV) a celkový počet zjištěných CNV (prvky celkem).

392 elements in total
CNVs 66 / 392

Ikona visacího zámku označuje, že analýza používá panel s omezeným přístupem, a podrobnosti o CNV se zobrazují v souvisejících popiscích nástrojů.

0 in locked panel
392 elements in total
CNVs 0 / 0

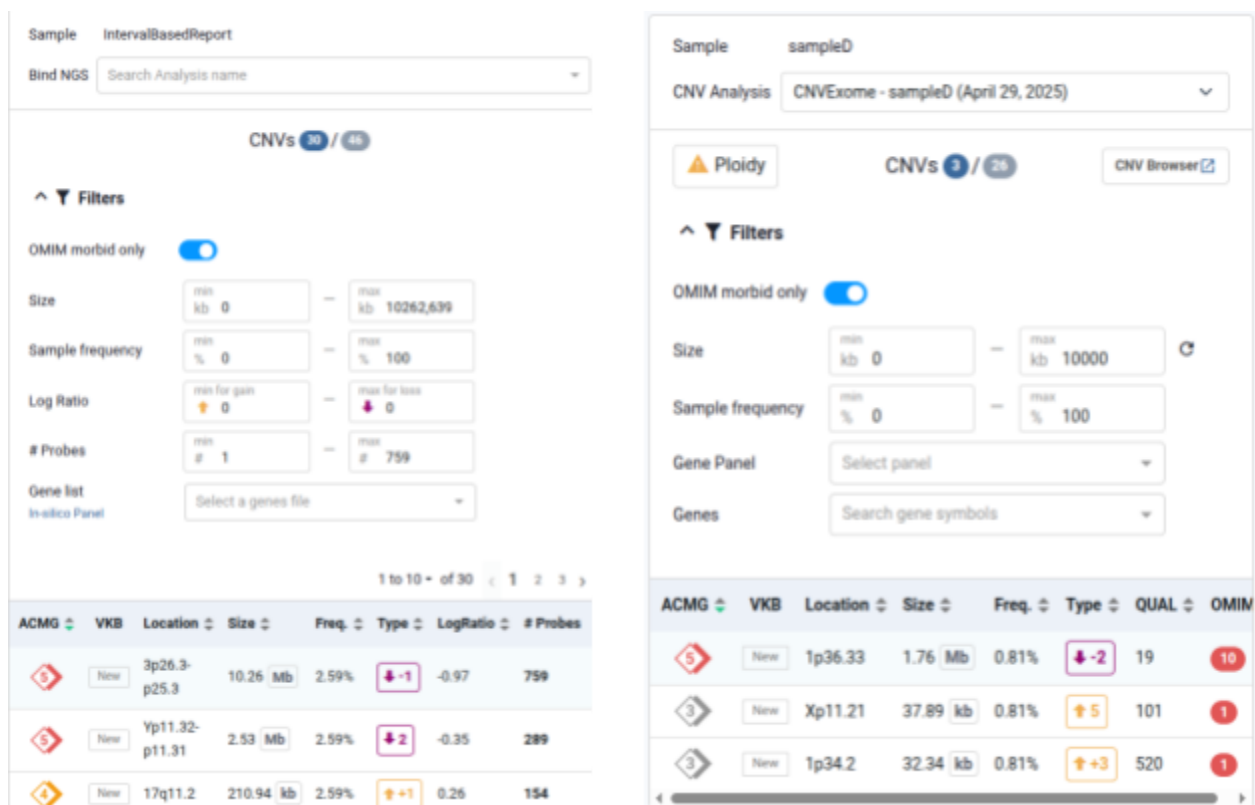
Při zjištění abnormální hodnoty ploidie a/nebo v případě, že před spuštěním analýzy nebyla vyplněna **metadata** o pohlaví (**viz** Metadata vzorku), se zobrazí varovná zpráva .

⚠ Ploidy

Kliknutím na ikonu "Ploidy" si uživatel může zobrazit ploidii pro každý chromozom a pohlaví pacienta. Abnormální hodnoty jsou zvýrazněny žlutě.

²⁷ <https://support.sentieon.com/appnotes/cnvscope/>

²⁸ <https://github.com/PopicLab/cue>



Obrázek 25: Overview of detected CNVs and filter menu of a CGH analysis (left) and a standard analysis (right).

V případě analýzy CGH MicroArrays v kombinaci s maticí SNP (CGH+SNP) jsou zjištěné případy ztráty heterozygotnosti (LOH) uvedeny v tabulce výsledků CNV. Tyto události odpovídají ztrátě heterozygotity neutrální z hlediska kopií (CN-LOH). Jinými slovy, oblast si zachovává dvě kopie, které jsou však identické, charakterizované přítomností homozygotních SNP.

V případě analýzy CGH MicroArrays je možné v rozevírací nabídce "Bind NGS" vybrat analýzu SNV, která není CGH (GermlineVar, GermlineFamily, GermVar Tertiary a GermVar Tertiary Family), aby se zobrazily varianty, které se překrývají s detekovanou CNV.

Filters

OMIM morbid only. Tento filtr je ve výchozím nastavení aktivován pro snadnou identifikaci CNV ovlivňujících klinické geny, anotované OMIM Morbid na základě jejich NCBI identifikátoru (NCBI gene_id) nebo, pokud to není možné, na základě názvu genu a synonym (gene_symbol a gene_synonyms). Stačí zrušit tento filtr a zobrazí se všechny CNV zjištěné ve vzorku.

Size. Uživatel může filtrovat CNV podle jejich očekávané velikosti (v kb).

Sample frequency. Uživatel může filtrovat CNV podle jejich velikosti a frekvence (Sample frequency). Tato frekvence odpovídá výskytu CNV mezi všemi vzorky analyzovanými do data analýzy a porovnávány se stejným modelem, včetně vzorků použitých k sestavení modelu. Frekvence vzorku je založena na 95% překryvu a zohledňuje stejný typ CNV (gain/loss/deletion). Po najetí myši na její hodnotu zobrazenou v procentech se uživateli zobrazí odpovídající podíl. Po otevření analýzy může aktualizace frekvence trvat několik minut.

Gene Panel. Stejně jako panel in silico umožňuje tento filtr uživateli vybrat a použít seznam genů pro filtrování CNV zjištěných v analýze pomocí rozbalovací nabídky (viz Správa **uživatelských parametrů**). Tento filtr není k dispozici, pokud byl na analýzu použit omezený panel ze záložky Panel bočního panelu pro správu analýzy (viz část Panel **modal view**).

Genes. Umožňuje filtrovat zjištěné CNV obsahující geny vybrané pomocí tohoto filtru.

Log Ratio. U analýz založených na datech MicroArray CGH se v dalším sloupci zobrazuje Log Ratio (mezi surovým signálem a kontrolním signálem) obsažený ve vstupních souborech Agilent. Přidružený filtr umožňuje uživateli vybrat minimální hodnoty pro přírůstky kopií (min pro přírůstek) a maximální hodnoty pro úbytky kopií (max pro úbytek). Události LOH (loss of heterozygosity) jsou zachovány i po použití tohoto filtru.

#Probes. U analýz založených na datech CGH MicroArrays sloupec #Probes udává počet sond použitých v každé oblasti (exonické, intergenické nebo intronické). Tyto hodnoty jsou převzaty ze vstupního souboru IntervalBasedReport.xls a související filtr umožňuje uživateli vybrat minimální a maximální počet použitých sond. V kombinované analýze CGH+SNP udává počet sond počet homozygotních SNP v oblasti LOH.

Třídění CNV

Seznam CNV lze také seřadit podle třídy patogenicity ACMG (ACMG), genomové pozice (Location), velikosti (Size), frekvence vzorku (Freq.), počtu kopií (Type), skóre kvality (QUAL) v GermVar Tertiary a podle LogRatio v CGH MicroArray (LogRatio).

ACMG. Třída patogenity každé varianty se vypočítává poloautomaticky podle doporučení *American College of Medical Genetics and Genomics*. Třída patogenity je výchozím způsobem řazení v tabulce CNV.

Události LOH (Loss of Heterozygosity) v CGH MicroArrays jsou vždy zařazeny do třídy 3 (VUS).

Type (Copy number). U CNV zjištěných na autozomálních chromozomech (1 až 22) se zobrazuje změna počtu kopií (např.: +2, -1, +3), protože počet kopií je ve výchozím nastavení nastaven na 2.

U CNV detekovaných na gonozomech (pohlavní chromozomy X a Y) se však zobrazuje absolutní počet kopií (např.: 0, 1, 2), nikoliv násobná změna.

Na rozdíl od analýzy genových panelů nejsou pro zahrnutí pohlavních chromozomů do volání CNV v exomu vyžadována metadata vzorku související s pohlavím.

U kombinovaných analýz CGH+SNP jsou události CN-LOH, které si zachovávají dvě kopie, ve sloupci Typ reprezentovány odznakem . 

QUAL. Skóre kvality Phred se vypočítá v CNVExome pomocí nástroje GATK caller. Pokud se CNV skládá z několika genomových intervalů, je CNV přiřazeno nejvyšší skóre QUAL ze všech intervalů. U terciárních analýz se zobrazí skóre QUAL získané ze vstupního souboru VCF (DRAGEN nebo ONT).

CNV Browser

Ikona otevře v nové záložce prohlížeč CNV Browser, který zobrazuje frekvenci alel B (BAF) jednotlivých SNP (Single Nucleotide Polymorphism) přítomných v analyzovaném vzorku. Graf je popsán v předchozí části (CNV v panelech genů , v části CNV Browser).

Podrobná stránka CNV

Shrnutí CNV. Tato úvodní karta obsahuje řadu informací a anotací o vybrané CNV, včetně:

- Podrobnosti o automatické klasifikaci ACMG
- Klinické oblasti a geny zahrnuté do intervalu ovlivněného danou variantou. Anotace klinických oblastí vychází z databází [ClinGen](#) a [OMIM](#). Ve výchozím nastavení se zobrazují pouze klinické geny s anotací OMIM Morbid. Pro zobrazení všech klinických genů může uživatel aktivovat funkci "Show all genes" v pravé části stránky.
- Překryvy mezi detekovanými CNV a databází [DGV](#).
- Malé varianty (SNV a indely) detekované v genech ovlivněných CNV. Ve výchozím nastavení je aktivován filtr SeqOne-diag. Chcete-li tento filtr deaktivovat, klikněte na

SeqOne-diag a zrušte výběr SeqOne-diag z profilů filtrů. Pro zobrazení všech variant mohou uživatelé aktivovat funkci "Show all variants" v pravé části stránky.

- Speciální oddíl pro databázi znalostí variant (VKB) pro identifikaci variant shodných z více než 95 %, které již byly ve skupině validovány, a jejich příslušná hodnocení.

ACMG Classification (Automatic) ⓘ

4

Likely Pathogenic (+0.9)

Genomic content

Genes or Genomic regions

Gene number

Literature and databases

Inheritance patterns

1A 0

2C_1 +0.9

3A 0

No selected tags, only 40 is assessed.

Not assessed

Clinical regions

Clinical genes

Show all genes

DGV overlaps

Variants

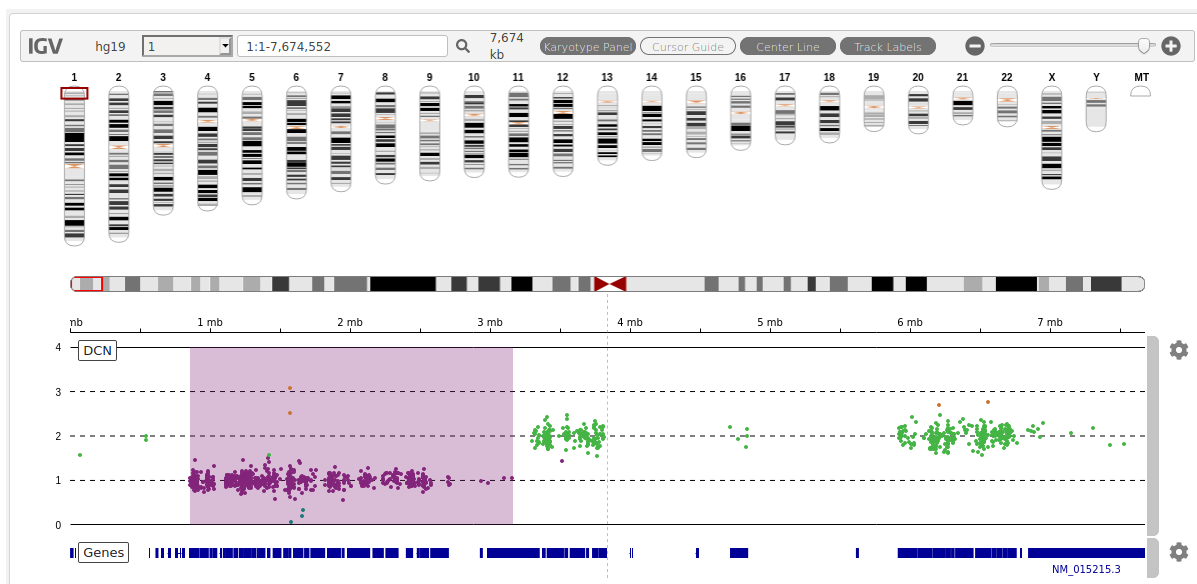
SeqOne-diag 0 / 51

Show all variants

No variant found

Obrázek 26: Overview of the CNV summary tab

- Prohlížeč genomu IGV, kde lze vizualizovat CNV (není k dispozici v analýzách GermVar Tertiary a GermVar Tertiary Family). Pro kontrolu stability signálu lze kromě analyzovaného jedince zobrazit až pět dalších pacientů z konfigurační nabídky  grafu představujícího počet kopií (DCN).



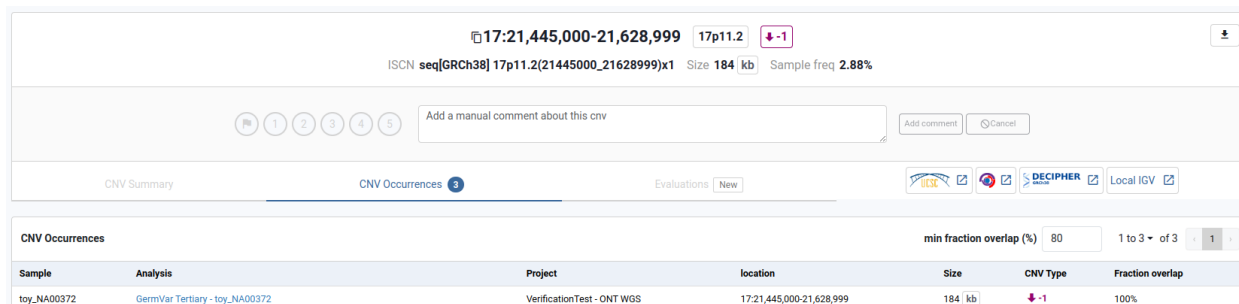
Obrázek 27: Pohled na CNV z integrovaného prohlížeče IGV

Barevný kód je pro vybraného pacienta následující: blue (0 až 0,5 kopie), fialová (0,5 až 1,5 kopie), zelená (1,5 až 2,5 kopie), oranžová (2,5 až 3,5 kopie) a red (>3,5 kopie). Ostatní vzorky jsou zobrazeny šedě.

CNV Occurrences. Tato karta slouží k identifikaci dalších vzorků obsahujících CNV nezávisle na hodnocení. Pro každý výskyt se zadají následující informace:

- název vzorku (Sample),
- název analýzy a odkaz na ni (Analýza),
- název projektu (Project),
- genomické umístění CNV (lokalizace),
- velikost CNV (Size),
- CNV-Type,
- překrytí frakcí.

Zásahy jsou seřazeny sestupně podle procenta překrytí. Zobrazené CNV je možné filtrovat podle minimální frakce překrytí vyplněním pole "min fraction overlap (%)" v horní části seznamu. Ve výchozím nastavení je tento filtr nastaven na 80 %.



Obrázek 28: Overview" záložky Occurrence v analýze CNVExome

Evaluations. Klasifikaci patogenity pro každou CNV může definovat uživatel a doplnit ji interpretací ve formě komentáře. Každé nové hodnocení také pomáhá naplnit osobní znalostní databázi laboratoře.

UCSC. Externí odkaz na [UCSC](#) Genome Browser umožňuje uživatelům zobrazit genomickou oblast CNV. Tato stránka se otevře v další záložce prohlížeče a obsahuje osobní nastavení uživatele pro daný prohlížeč, čímž nabízí personalizované zobrazení.

CNV-Hub. Externí odkaz na klasifikátor [CNV-Hub](#) umožňuje vizualizovat klasifikaci CNV s přihlédnutím k více klasifikačním nástrojům. Odkaz otevře další kartu prohlížeče a integruje osobní nastavení prohlížeče uživatele, čímž nabízí personalizovanou vizualizaci.

DECIPHER Genome Browser. Externí odkaz na [DECIPHER](#) Genome Browser umožňuje zobrazit CNV v genomu. Odkaz otevře další kartu prohlížeče a integruje osobní nastavení prohlížeče uživatele, čímž nabízí personalizovanou vizualizaci.

Local IGV. Tento odkaz přesměruje aplikaci na souřadnice právě interpretovaného CNV. Pro zobrazení CNV je nutné, abyste předtím načetli vzorek BAM do aplikace IGV Desktop. Tato možnost je kompatibilní pouze s aplikací IGV Desktop verze ≥ 17 . *IGV Desktop*²⁹

Kromě zarávnávacích souborů jsou na *kartě* Soubory v analýze CNVExome k dispozici ke stažení různé soubory:

- Výsledky analýzy CNV v podobě tabulkového souboru (cnv.tsv).
- ploidy pro každý chromozom (contig_ploidy.tsv)
- Předpokládaný počet kopií pro každý exon v manifestu (denoise.tsv)

²⁹ <https://igv.org/doc/desktop/>

Export dat CNV

Ikona  vpravo nahoře umožňuje exportovat soubor TSV, který obsahuje následující údaje pro příslušný CNV:

- jeho hodnocení VKB a komentář,
- Hodnocení ISCN,
- ACMG Classification,
- chromozomu, počáteční a koncové pozice a cytobandu,
- CNV velikost in tip,
- Copy number:
- Copy number:
- Sample Freq.
- skóre kvality (stupnice Phred)
- Log2Ratio
- počet zapojených sond (Probes)
- počet genů ovlivněných CNV.

Informace o souvisejícím genu (genech) jsou popsány v následujících sloupcích:

- název genu (Gene_Symbol),
- Fenotyp OMIM (OMIM_Phenotypes),
- pravděpodobnost genové intolerance vůči LoF (Pli_Score)
- Skóre haploinsuficience a triplosenzitivity ClinGen (HI_Score, TS_Score)
- počet exonů překrytých CNV (Nb_overlap_exon),
- celkový počet exonů v genu (Nb_total_exon).

Pokud se některá popisná pole týkají pouze určitých CNV (například v případě MicroArray CGH), mají tato pole hodnotu "NA".

13.1.6. Prohlížeč velkých variant pro CNVs a SVs (LVV)

Prohlížeč velkých variant (Large Variant Viewer, LVV) umožňuje uživateli vizualizovat a analyzovat CNVs a strukturní varianty (SVs) na jedné kartě. Rozhraní se skládá z centrální tabulky variant, panelu filtrů vlevo a postranního panelu, který se otevře vpravo při výběru varianty, podobně jako na kartě variant (část Malé varianty).

V kompatibilních analýzách se prohlížeč nachází na kartě označené "CNVs & SVs" (CNVs a SVs) nebo "SVs", v závislosti na typu dat:

- Karta CNVs & SVs zobrazuje CNVs (Ztráta/Zisk), SVs (Delece/Duplikace) a události LOH/UPD.
- Starší karta "CNV": CNVs se také zobrazují ve starším prohlížeči, aby se zachovala zpětná kompatibilita se stávajícími pracovními postupy.
- SVs jsou přístupné pouze na kartě "CNVs & SVs" nebo "SVs".

Dostupnost dat v prohlížeči velkých variant závisí na typu analýzy:

- WGS: Plně kompatibilní (CNVs, SVs, UPDs).
- WES: Plně kompatibilní (CNVs, SVs, UPDs). CNVs se zobrazí pouze v případě, že je jako součást analýzy spuštěn modul pro volání CNV.
- Panely: Částečně kompatibilní (pouze SVs).
- Terciární analýza (VCF/aCGH):
- aCGH: Zobrazují se všechny události počtu kopií.
- VCF: Jsou zobrazeny CNVs a SVs (symbolické ALT alely). SVs se specifickými sekvencemi ALT jsou zobrazeny v prohlížeči malých variant.

Tabulka výsledků CNVs & SVs

Podobně jako tabulka malých variant, hlavní tabulka zobrazuje jeden řádek pro každou událost, s následujícími sloupci poskytujícími metriky specifické pro variantu:

DiagAI. Priorizační skóre od 1 do 100, přizpůsobené pro velké varianty, které řadí varianty podle pravděpodobnosti, že způsobují onemocnění a jsou diagnosticky relevantní v kontextu daného případu (zejména pokud jsou poskytnuty termíny HPO). Skóre DiagAI kombinuje několik komponent, včetně: predikce patogenity prostřednictvím skóre UP², shody fenotypu prostřednictvím PhenoGenius (řízené HPO) a expertních pravidel/pravidel dědičnosti a kvality.

ACMG. Indikátor klasifikace ACMG založený na velkých variantách je zobrazen jako klasifikační skóre 1–5 v tabulce LVV. V zásuvce s podrobnostmi zobrazuje karta ACMG základní kritéria a živý souhrn klasifikace od Patogenní po Benigní (hlavně pro CNV Zisk/Ztráta).

Klasifikace (ikona vlajky), která umožňuje uživateli klasifikovat variantu ve VKB.

Typ CNV/SV (zisk nebo ztráta), velikost, umístění a cytopruh, Oblasti (exonové, intronové, UTR) a geny překrývající se s CNV (názvy genů HGNC), přítomnost v OMIM, podíl kódujících genů a podrobnosti o překrývajících se exonech. Varianty jsou anotovány překrývajícími se geny a geny nacházejícími se do 5 kb od zlomových bodů (5 kb přesah).

Geny spojené s variantou jsou seřazeny podle klinické relevance (např. morbidní stav v OMIM, procento překrytí, skóre omezení). Tři nejrelevantnější geny se zobrazují v hlavní tabulce. Úplný seznam je k dispozici na kartě "Genes & Regions" (Geny a regiony) v postranním panelu ("Details Drawer" (Zásuvka s podrobnostmi)).

RMSK %. Procento události překrývající se s anotacemi RepeatMasker (např. repetitivní oblasti).

ClinVar & Decipher (Klinické databáze). Oba sloupce zobrazují počet shod z příslušné klinické databáze. Nejlepší shoda je definována jako událost, kde je překrytí databáze > 50 %. Klasifikace ClinVar a Decipher jsou mapovány na Benigní (B), VUS nebo Patogenní (P).

DGV Gold & GnomAD (Populační databáze). Tabulka zobrazuje maximální populační frekvenci nalezenou pro variantu. Události jsou považovány za shodné, pokud je překrytí varianty > 80 % a typ události je ekvivalentní. Karta Frekvence v postranním panelu uvádí všechny překrývající se události s překrytím > 1 %.

Phenotypes (Fenotypy). Kliknutím na čísla zobrazená ve sloupcích odpovídajících výsledkům algoritmu Phenogenius může uživatel získat přístup k:

- procentu shody, což je nejvyšší skóre shody mezi jedním postiženým genem a termíny HPO pacienta.
- zlomku shody, což je počet termínů HPO pacienta, které se shodují s alespoň jedním postiženým genem.
- U rodinné analýzy se párování provádí pomocí termínů HPO probanda.

SampleFreq (Interní frekvence vzorku). Frekvence varianty v rámci aktuálního projektu a subjektu je zobrazena v příslušných sloupcích.

- Shody jsou založeny na > 80 % překrytí varianty a stejném typu události.
- Frekvence aktualizace: data jsou založena na snímku analýz provedených před spuštěním aktuální analýzy. V Details Drawer je k dispozici manuální tlačítko "Refresh" (Obnovit) pro aktualizaci těchto dat.

Sloupce skóre. Uživatel si může zobrazit více skóre:

- UP²: skóre predikce patogenity, které je součástí DiagAI.
- pLI, LOEUF: intolerance ke ztrátě funkce.
- pTriplo, pHAPLO: Triplosenzitivita a Haploinsuficience.
- HI ClinGen, TS ClinGen.

Sloupce vzorku. Tyto sloupce zobrazují metriky událostí specifické pro zdroj dat:

- Pro rodinné analýzy (s výjimkou aCGH) se metriky vzorku zobrazují pro všechny členy rodiny.
- Genotype: pro CNVs, kde vyvolávač (caller) neposkytuje genotyp, systém odvodí genotyp na základě absolutního počtu kopií vzhledem k ploidii:
 - Homozygot (hmz) s plným tmavým bodem
 - Heterozygot (htz) s polovičním tmavým bodem
- VAF / #Reads: frekvence alely varianty a podpůrná čtení (k dispozici pro SVs).
- #Copy: absolutní počet kopií (k dispozici pro CNVs).
- LogRatio / Probes: k dispozici pro data aCGH.
- QUAL: skóre kvality z vyvolávače variant nebo vstupního souboru.

Třídění, filtrování, prohlížení a interpretace

Filtrování. Panel filtrů umožňuje uživateli filtrovat varianty na základě dat sloupců, podobně jako systém filtrování na kartě Malé varianty (viz příslušná část).

- Genový panel: filtruje události překrývající se s geny ve specifickém seznamu genů.
- Kvalita: filtruje na základě sloupce QUAL. Ve výchozím nastavení jsou události bez skóre kvality (N/A) skryty, pokud není vybráno "Keep unknown" (Ponechat neznámé).
- Předvolby: konfigurace filtrů lze uložit jako předvolby pro budoucí použití.
- U rodinných analýz používá filtr kvality logiku "OR" (NEBO). Varianta se zobrazí, pokud alespoň jeden člen rodiny splňuje práh kvality.

Třídění. Ve výchozím nastavení je tabulka seřazena podle DiagAI Score (sestupně). Uživatel může třídit podle UP2, Size, Location nebo Sample Freq. Entity, vzestupně nebo sestupně.

Stejně jako ve starší kartě CNV může uživatel otevřít okna ploidie a prohlížeč CNV v horní části tabulky (viz část CNV: Zobrazení a interpretace). Ikona správy sloupců je také k dispozici zcela vpravo v záhlaví.

Postranní panel CNV/SV (Details Drawer)

Podobně jako na stránce variant (část Malé varianty) se výběrem varianty v Large Variant Viewer otevře na pravé straně obrazovky podrobná stránka. Tato stránka obsahuje několik karet:

Varianty. Tato souhrnná karta obsahuje nomenklaturu varianty, skóre DiagAI, seznam podezřelých složených heterozygotů (Suspected Het Compound), informace o genech a regionech a odkazy na databáze.

ACMG. Podrobná klasifikace.

Geny a regiony. úplný seznam genů a regionů spojených s variantou, s podrobnými klinickými informacemi, jako je morbidní stav OMIM, procento překrytí, skóre omezení a shoda fenotypu.

Klasifikace. Klasifikujte aktuální variantu ve vaší VKB v horní části postranního panelu a uložte ji. Aktuální klasifikaci a výskyty této varianty v jiných analýzách naleznete na této stránce.

Frekvence. Uživatel má přístup k podrobnostem o frekvencích variant v databázích DGV Gold a GnomAD, stejně jako pro subjekt a aktuální projekt.

Klinické databáze. ClinVar & Decipher.

Viewer. Uživatel si může vybrat, zda chce variantu zobrazit v SeqOne viewer nebo IGV. Na podkartě IGV může uživatel zobrazit "Start" (Začátek) nebo "End" (Konec) události, nebo "Start to End" (Od začátku do konce). Na podkartě SeqOne může uživatel přistupovat ke Genome Browser a vizualizovat událost v jedné z mnoha dostupných vrstev: CNV, SV, UPD, Gene, CN (copy number), BAF, DGV.

13.1.7. Znaký v analýze SomaCGP

Metodika výpočtu MSI skóre

Analýza MSI je individuální analýza prováděná na základě standardizované reference. Tato reference se skládá ze souboru předem vybraných a sdružených negativních vzorků. Nádorový vzorek se s touto referencí porovnává pomocí nástroje *MSIsensor*. Poté se vypočítá skóre, které je reprezentováno procentem nestabilních mikrosatelitních oblastí pokrytých panelem.

Panel normálních vzorků, který v současné době podporuje SomaCGP, odpovídá panelu Agilent SureSelect Cancer CGP. Použití jiné soupravy nebo vlastního návrhu odvozeného od CGP bude vyžadovat vytvoření speciálního panelu normálů.

Metodika výpočtu skóre TMB

TMB se vypočítá ve třech krocích:

1. Výpočet velikosti kódovacích oblastí zahrnutých do panelu
2. Identifikace somatických variant
3. Výpočet počtu somatických mutací na megabázi (mut/Mb)

Identifikace somatických variant je založena na sadě níže popsanych filtrů:

- Variant: Typ: Filter
 - Pouze SNV (MNV jsou vyloučeny)
 - Zachované efekty: zisk stop kodonu, posun čtecího rámce, sestřihová varianta (akceptor a donor), ablace transkriptu, delece/inserce ve fázi, missense varianta, synonymní varianta.

Synonymní varianty jsou z výpočtu nesynonymního skóre TMB vyloučeny.

- Variant: Filters
 - Varianty detekované nástrojem Mutect2, které prošly filtry FilterMutectCalls
 - Hloubka pokrytí (DP) $\geq 50x$
 - Skóre kvality > 30
 - Poměr alelické podpory (ASR) $> 0,1$
- Filters somatických variant
 - Frekvence variantních alel (VAF) mezi 5 a 40 %.
 - Frekvence v obecné populaci (GnomAD) $\leq 0,0001$

Rizika používání syntetických kontrolních vzorků TMB

- Distribuce nepřírodních mutací:
 - Upravené mutace nemusí odrážet genomický kontext skutečných nádorů (např. shlukování mutací, zkreslení vláken, sekvenční kontext).
- Nedostatek heterogenity nádoru:
 - Syntetické vzorky nenapodobují klonální diverzitu ani čistotu nádoru, což ovlivňuje rozložení frekvence variantních alel (VAF).
 - Žádný šum na pozadí nebo artefakty FFPE: skutečné vzorky obsahují artefakty sekvenování, poškození FFPE a Copy number variations, které se v syntetických materiálech nevyskytují.
- Rozdílná velikost fragmentů a chemické chování:
 - Syntetické konstrukty se často chovají odlišně během přípravy knihovny, hybridního zachytu a sekvenování.

Vizualizace a interpretace

Na kartě "Signatures" se zobrazují výsledky analýz MSI a TMB ve dvou sekcích:

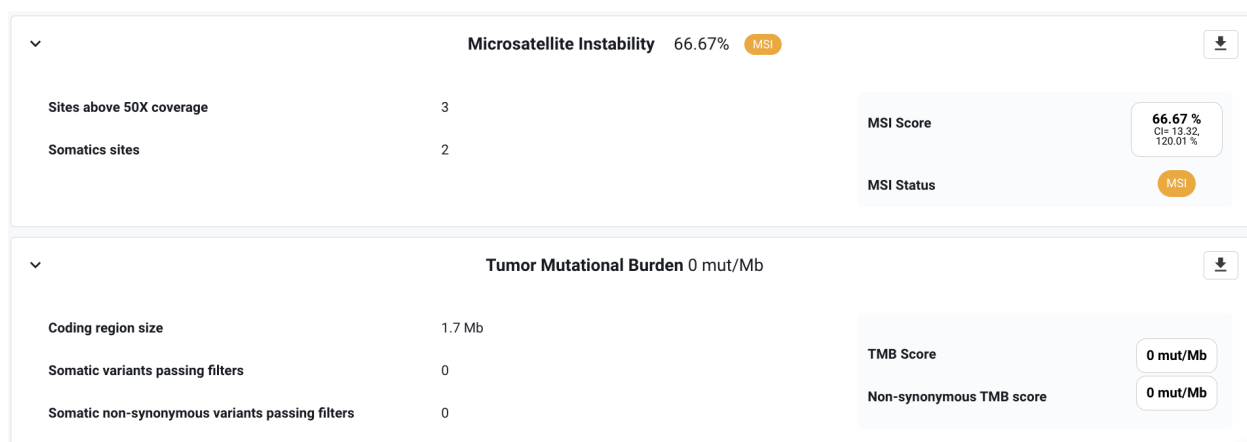
Microsatellite Instability. Toto zobrazení zobrazuje informace jako :

- počet mikrosatelitních oblastí s pokrytím větším nebo rovným 50x (Místa s pokrytím nad 50x),
- počet somatických míst,
- MSI skóre v procentech a 95% interval spolehlivosti (CI),
- Status:

Tumor Mutational Burden. Toto zobrazení zobrazuje následující informace:

- velikost kódující oblasti vyjádřená v Mb (velikost kódující oblasti),
- počet somatických a nesynonymních variant procházejících filtry (Somatické varianty procházející filtry a Somatické nesynonymní varianty procházející filtry),
- TMB skóre všech somatických variant, které prošly filtry, v mut/Mb (TMB skóre),
- TMB skóre nesynonymních variant, které projdou mut/Mb filtry.

Každý výsledek lze stáhnout kliknutím na ikonu  vpravo nahoře u každé sekce jako soubor tsv.



Obrázek 29: Zobrazení karty “Signatures” analýzy SomaCGP.

13.1.8. Krátké tandemové repetice (STRs)

Vizualizace krátkých tandemových repetic (STRs) je k dispozici pro worksety analýzy GermVar Tertiary a GermVar Tertiary Family určené k analýze dat celogenomového sekvenování (WGS), pokud jsou generována sekvenováním Oxford Nanopore Technologies.

Detekce krátkých tandemových repetic (STR) je také k dispozici v analytických pracovních sadách GermVar a GermVar Family pro WGS z dat FASTQ generovaných pomocí Illumina, Element Biosciences, MGI a Oxford Nanopore Technologies.

Metodika

Krátké tandemové repetice z FASTQ

Data krátkých tandemových repetic (STR) jsou detekována pomocí ExpansionHunter, který zpracovává soubor BAM společně s referenčním genomem a katalogem variant definujícím cílové lokusy. Výsledný VCF je poté anotován programem Stranger, který klasifikuje každý lokus jako Normální, Pre-mutaci nebo Patogenní na základě pozorovaného počtu repetic a známých prahových hodnot onemocnění.

Oba nástroje jsou integrovány přímo do pipeline GermVar a GermVar Family.

ExpansionHunter: <https://github.com/Illumina/ExpansionHunter>

Stranger: <https://github.com/moonso/stranger>

Krátké tandemové repetice z VCF

Data tandemových repetic jsou extrahována ze souboru VCF (sample.wf_str.vcf.gz) vygenerovaného z pracovního postupu wf-human-variation v2.3.0 aplikace EPI2ME (Oxford Nanopore Technologies) a dodaného uživatelem. Protože tato data byla anotována pro každý lokus nesoucí STR z katalogu variací udržovaného aplikací EPI2ME, naše pipeline anotace nemění. Premutovaný nebo patogenní status každé alely je určen podle počtu sekvenovaných opakování a známých prahových hodnot premutace a patogenity.

Zobrazení a výklad

Karta STRs umožňuje zobrazit krátké tandemové repetice ve formě tabulky, která zobrazuje následující údaje pro každý genotypovaný lokus:

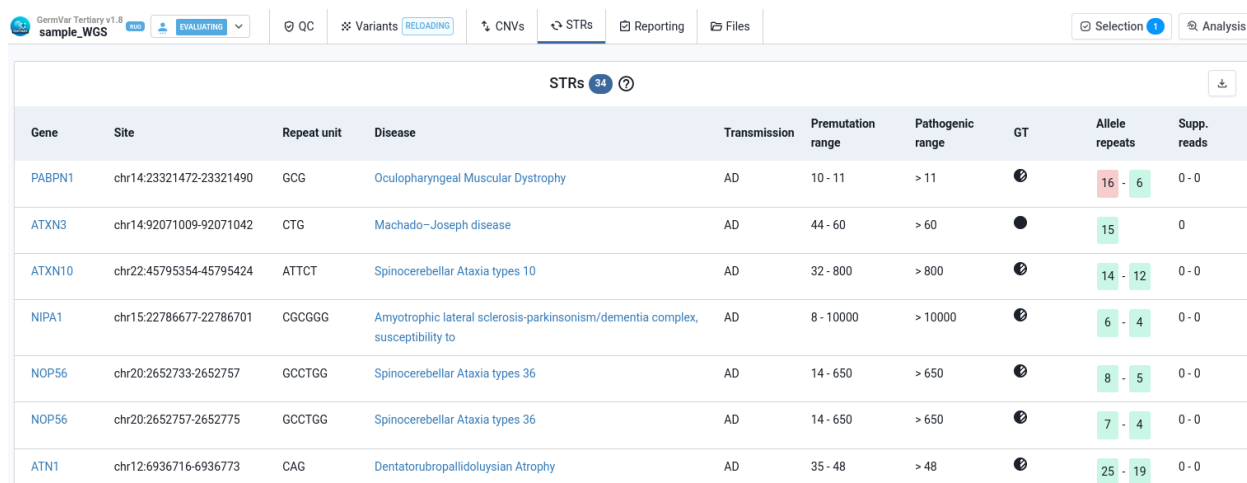
Informace o lokusech a anotace

- [Gene](#): název genu (HGNC) a odpovídající odkaz.

- **Site:** genomické souřadnice opakování.
- **Repeat Unit:** opakující se nukleotidový motiv.
- **Onemocnění:** související onemocnění podle databáze OMIM a odpovídající odkaz.
- **Transmission:** známý způsob přenosu.
- **Premutation range:** rozsah premutace odpovídající minimálnímu a maximálnímu počtu opakování, které určují status premutace.
- **Pathogenic range:** dolní hranice opakování určující patologenní status.

Podrobnosti specifické pro každý vzorek

- **GT:** homozygotní, bílelický nebo heterozygotní genotyp lokusu.
- Počet **opakování alel:** počet opakování na alelu. Homozygotní lokusy nesou jedinou hodnotu. U bílelických a heterozygotních lokusů je uveden počet opakování pro obě alely. Status každé alely je určen počtem opakování: divoký typ zeleně, premutovaný oranžově a patologenní v červeně.
- **Supp. reads:** počet podporujících čtení pro každou alelu. U homozygotních lokusů je uvedena jedna hodnota, zatímco u bílelických a heterozygotních lokusů jsou uvedeny hodnoty pro 2 alely.



Gene	Site	Repeat unit	Disease	Transmission	Premutation range	Pathogenic range	GT	Allele repeats	Supp. reads
PABPN1	chr14:23321472-23321490	GCG	Oculopharyngeal Muscular Dystrophy	AD	10 - 11	> 11	●	16 - 6	0 - 0
ATXN3	chr14:92071009-92071042	CTG	Machado-Joseph disease	AD	44 - 60	> 60	●	15	0
ATXN10	chr22:45795354-45795424	ATTCT	Spinocerebellar Ataxia types 10	AD	32 - 800	> 800	●	14 - 12	0 - 0
NIPA1	chr15:22786677-22786701	CGCGGG	Amyotrophic lateral sclerosis-parkinsonism/dementia complex, susceptibility to	AD	8 - 10000	> 10000	●	6 - 4	0 - 0
NOP56	chr20:2652733-2652757	GCCTGG	Spinocerebellar Ataxia types 36	AD	14 - 650	> 650	●	8 - 5	0 - 0
NOP56	chr20:2652757-2652775	GCCTGG	Spinocerebellar Ataxia types 36	AD	14 - 650	> 650	●	7 - 4	0 - 0
ATN1	chr12:6936716-6936773	CAG	Dentatorubropallidoluysian Atrophy	AD	35 - 48	> 48	●	25 - 19	0 - 0

Obrázek 30: Zobrazení záložky STRs Tertiary analýzy GermVar.

Tabulku lze stáhnout jako soubor repeats.tsv pomocí ikony  vpravo nahoře u tabulky. Tento soubor je rovněž k dispozici na kartě "Files" v analýze.

13.1.9. Fúzní geny

13.1.9.1. Fúze zjištěné z DNA

Metodika

Modul [GRIDSS](#) (Genome Rearrangement IDentification Software Suite), specializovaný na detekci genomových přestaveb a implementovaný v pracovních sadách SomaVar a SomaVar LF, se používá k detekci genových fúzí ze sekvenování DNA, obohacených o záchyt zaměřený na panel genů.

Zobrazení a výklad

Záložka Fúze somatické analýzy (SomaVar, SomaVar LF, SomaCGP, SomaVar Tertiary) zobrazuje seznam fúzí, které prošly různými filtry GRIDSS, ve formě tabulky s podrobnými informacemi o fúzi a každém zúčastněném partnerovi:

Informace o fúzi

- fúze se symboly genů jednotlivých partnerů,
- typ události (translokace, duplikace, inverze nebo delece),
- výskyt fúze v databázi COSMIC,
- frekvence alel (VAF),
- počet reads podporujících fúzi,
- celkové pokrytí,
- filters přiřazené službou GRIDSS.

Informace týkající se partnerů

- genomické souřadnice bodu fúze (breakend),
- příslušný transkript RefSeq,
- místo odpovídající bodu splynutí (bod zlomu): UTR, 3'UTR, Intron, Exon (nekódující bod zlomu), CDS (kódující bod zlomu) nebo Splice site,
- číslování exonu nebo intronu zapojeného do fúze vzhledem k celkovému počtu exonů nebo intronů.

Nabídka filtrů umožňuje

- vyhledat konkrétní fúzi nebo gen zapojený do fúze pomocí vyhledávacího panelu,
- zobrazit fúze, které neprošly filtry GRIDSS (LOW_QUAL), zaškrtnutím políčka "Zobrazit filtrované fúze". Toto políčko je ve výchozím nastavení nezaškrtnuté.

FUSION	BREAKEND 1	BREAKEND 2	TYPE	DATABASES	TRANSCRIPT 1	SITE 1	EXON/INTRON 1	TRANSCRIPT 2	SITE 2	EXON/INTRON 2	VAF	SUPP. READS	COVERAGE	FILTER
ALK::EML4	2:29446685 (-)	2:42523824 (+)	inversion	cosmic	NM_004304	intron	19/28	NM_019063	intron	12/22	0.329 %	14	4251	PASS
NRAS::LINC00486	1:115256634 (-)	2:33141297 (+)	translocation	N/A	NM_002524	intron	2/6	NR_027098	intron	3/5	5.629 %	128	2274	PASS
LINC00486::RET	2:33141301 (+)	10:43612238 (+)	translocation	N/A	NR_027098	intron	3/5	NM_020975	intron	11/19	3.617 %	118	3262	PASS
LINC00486::FGFR3	2:33141349 (+)	4:1803528 (+)	translocation	N/A	NR_027098	intron	3/5	NM_000142	intron	5/17	0.352 %	28	7963	PASS
PRIM2::HRAS	6:57331998 (+)	11:533875 (-)	translocation	N/A	NM_000947	intron	6/13	NM_005343	CDS	3/6	0.126 %	6	4753	LOW_QUAL

Obrázek 31: Vizualizace karty Fúze pro analýzu SomaVar nebo SomaVar LF.

Údaje v tabulce lze třídit podle genomických souřadnic (breakend), přítomnosti fúze v COSMIC (Databases), alelické frekvence (VAF) a počtu čtení podporujících fúzi (Supporting reads).

Odkaz spojený s genomickými souřadnicemi každého bodu zlomu otevře rozdělené zobrazení bodu fúze na IGV. Otisky nástrojů podrobně uvádějí počet split readů a neshodných párů ve sloupci Supporting reads (Podpurná čtení) a pokrytí v každém bodě zlomu (Cov. 1 a Cov. 2) ve sloupci Coverage (Pokrytí).

Ve sloupci Filters jsou zobrazeny 3 filtry kvality používané systémem GRIDSS, které udávají důvěryhodnost fúzního volání:

- PASS: fúze prošla filtry kvality. Skóre kvality je větší nebo rovno 500 a je považováno za volání se střední nebo vysokou důvěrou (pokud je > 1000).
- LOW_QUAL: fúze má skóre kvality nižší než 500 alespoň pro jeden ze dvou breakendů zapojených do fúze a je považována za volání s nízkou důvěrou.
- ASSEMBLY_ONLY: fúze je vyvolána na základě sestavy, nikoli na základě čtení, které ji podporuje. Sestava se týká rekonstrukce sekvencí kolem potenciálních bodů fúze, které jsou poté znovu zarovnány za účelem identifikace fúze.

Pokud je pro stejnou fúzi k dispozici více filtrů, zobrazí se ve sloupci Filters všechny filtry se znaménkem "+".

Kromě výsledků prezentovaných v rozhraní je na kartě "Files" nebo na ikoně  v pravém horním rohu stránky k dispozici ke stažení soubor fusions.tsv s podrobnými informacemi o zjištěných fúzích. Obsahuje následující informace:

- název fúze se symboly genů pro každého partnera (fúze)
- genomické souřadnice každého bodu fúze (breakend_1 a breakend_2).
- kanonický transkript pro každého partnera (transcript_canonical_1 a 2).
- místo odpovídající bodu fúze pro každého partnera (site_1 a site_2).

- číslování exonů nebo intronů podle příslušného transkriptu RefSeq (exon_1 a 2).
- 60bp doprovodných sekvencí obklopujících fúzní bod (flanking_region_1 a 2).
- filtry kvality přiřazené GRIDSS (filtr)
- typ události (type)
- počet split read (split_reads)
- počet nesouhlasných read pairs (discordant_pairs)
- pokrytí pro každého partnera (pokrytí_1 a pokrytí_2)
- podíl čtení podporujících fúzi (read_fraction)
- počet čtení podporujících fúzi (supporting_reads)
- celkové pokrytí (pokrytí)
- Identifikátor fúze COSMIC (cosmic_id)

Tento soubor obsahuje:

- Fúze zobrazené v rozhraní, které prošly filtry kvality (filtr = PASS),
- Zjištěné fúze, ale se skóre kvality nižším než 500 pro alespoň jeden ze dvou breakendů zapojených do fúze (filtr = LOW_QUAL).

Skóre kvality pro každý breakend lze nalézt v souboru gridss.raw.vcf.gz (sloupec QUAL), který lze stáhnout ze záložky Files tab.

13.1.9.2. Fúze zjištěné z RNA

SomaRNA nástroj pro analýzu dat RNA-seq z cílených přístupů, jako je zachycení a amplikon, což znamená, že uživatelé mohou identifikovat a zobrazit události genové fúze na úrovni transkriptu.

Metodika

K detekci a zobrazení fúzí pomocí SomaRNA se používá kombinace dvou nástrojů:

- [STAR](#): nástroj pro zarovnávání dat RNA-seq určený k zarovnávání nesouvislých sekvencí přímo k referenčnímu genomu.
- [Arriba](#): detekuje genové fúze z chimérických zarovnání.

Na základě chimérických zarovnání ze systému STAR používá Arriba sadu filtrů k eliminaci známých artefaktů a transkriptů pozorovaných v nepatologických kontextech. Každé potenciální události fúze přiřadí skóre důvěryhodnosti, které závisí na několika kritériích:

- Počet sekvencí podporujících fúzi
- Rovnováha mezi split read a nesouhlasnými partnery
- Vzdálenost mezi zlomovými body, jejich blízkost (intragenová nebo jiná).
- Typ události

Fúze detekované s vysokou nebo střední důvěryhodností pomocí Arriba jsou zobrazeny na SeqOne. Fúze s nízkou důvěrou jsou uvedeny v souborech "arriba_discarded" (viz níže). Důvěra Arriba odráží pravděpodobnost aberantního transkriptu, který nebyl pozorován ve zdravé tkáni, možnost základní genomové přestavby a vyloučení artefaktů.

Další podrobnosti o úrovni důvěryhodnosti najdete v dokumentaci Arriba: <https://github.com/suhrig/arriba/blob/master/documentation/08-Interpretation-of-results.md>.

Chcete-li pochopit, proč je fúze "low confidence", podívejte se do sloupce "Filters" (AA) v souboru "vyřazeno". Je v něm uveden filtr, který událost vyřadil. Podrobnosti o jednotlivých filtrech: <https://github.com/suhrig/arriba/blob/master/documentation/11-Internal-algorithm.md>

Zobrazení a výklad

Na kartě Fúze analýzy SomaRNA získá uživatel seznam předpovědí fúzí ve formě různých karet.

V horní části stránky jsou uvedeny podrobnosti o fúzi a jednotlivých zúčastněných partnerech:

Obecné informace o fúzi

- počet čtení spojovacího spoje, který se rovná součtu příslušných split readů každého z 5' a 3' genů zapojených do fúze,
- počet nesouhlasných partnerů,
- skóre důvěryhodnosti přidělené společností Arriba (vysoké, střední nebo nízké),
- typ události, která vedla k fúzi (translokace, duplikace, inverze nebo delece),
- vliv fúze na rám čtecího zařízení,
- případně výskyt fúze v databázi COSMIC.

Informace o každém fúzním partnerovi

- symbol pro každý gen,
- Transcript: Refseq nebo, pokud to není možné, ENSEMBL pro každý gen,
- číslování exonů, které se podílejí na fúzi,
- genomické souřadnice každého bodu zlomu,
- celkové pokrytí v bodě fúze (Coverage),
- read část podporující fúzi.

Split reads jsou reads zarovnaná převážně v genu 1 nebo genu 2 a překrývající fúzi. V případě discordant read pairs je jedno ze čtení v genu 1 a druhé v genu 2, aniž by se nutně překrývalo s fúzí.

Frakce čtení je součet rozdělených readů a discordant read pairs spojených s genem vydělený celkovým pokrytím. Hodnota podílu čtení může být vyšší než 100 %, pokud jsou rozdělená čtení spojená s genem také součástí nesouhlasných párů čtení. Tato reads se pak počítají dvakrát místo jednou.

DHRSX - TCEANC
E3 - E4

POLR2J2 - EFCAB13
E3 - E18

POLR2J2 - EFCAB13
E4 - E18

MEGF8 - CNFN
E3 - E4

General informations

of spanning junction reads

568

of discordant mates

246

Confidence ⓘ

high

Fusion type

translocation

Reading frame

in-frame

Database

cosmic

Partners informations

5prim

3prim

Gene

NPM1

ALK

Transcript

ENST00000517671 ⓘ

NM_004304.4

Exons

5 / 12

20 / 29

Breakpoint

5 : 170818803 (+)

2 : 29446394 (-)

Coverage

14004

8455

Read fraction

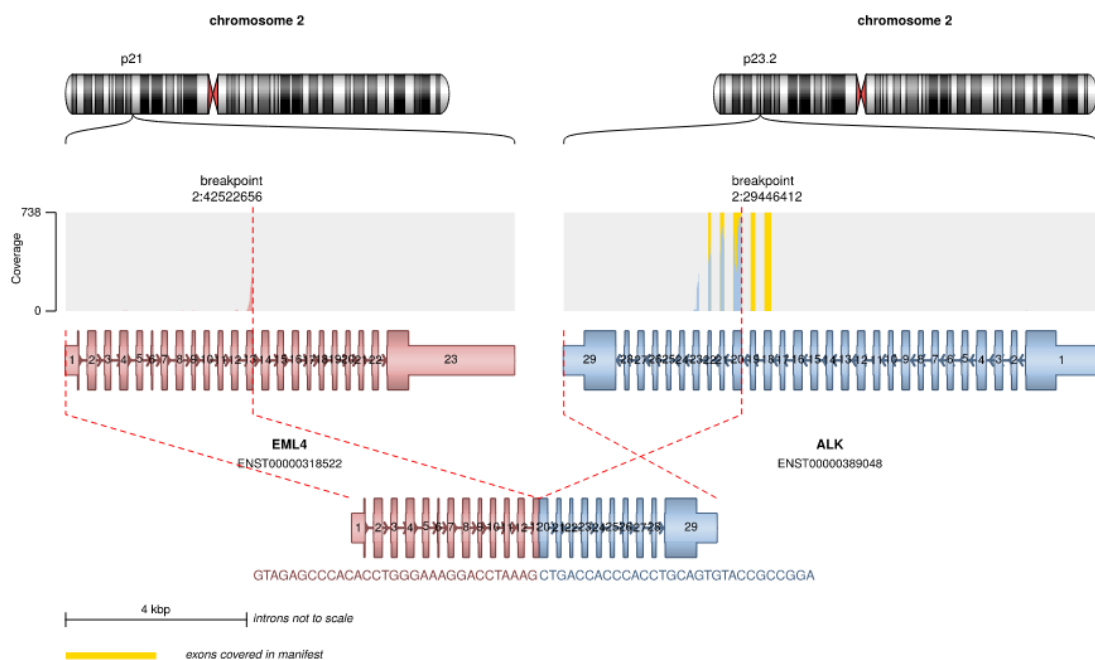
3.9%

6.08%

Obrázek 32: Vizualizace informací o fúzi z analýzy SomaRNA.

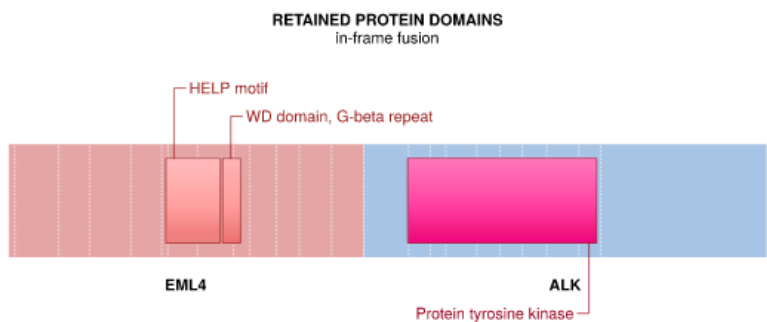
Ve spodní části stránky jsou graficky znázorněny fúze na úrovni genů a exonů a na úrovni proteinů. Tato poslední část je rozdělena do dvou záložek:

- První záložka představuje strukturu fúze a zapojené partnery, přičemž pokrytí exonů, které jsou zahrnuty do analytického manifestu, je znázorněno jako diagram na žlutém pozadí:



Obrázek 33: Pohled na fúzní událost mezi dvěma geny

- V druhém pohledu jsou zvýrazněny zachované proteinové domény po fúzi:



Obrázek 34: Pohled na polypeptid vzniklý fúzí dvou genů

Kromě výsledků prezentovaných v rozhraní jsou k dispozici různé soubory ke stažení:

- Fúze uchovávané systémem Arriba ve formátu TSV (soubor arriba.tsv),
- Fúze vyřazené filtry Arriba ve formátu TSV (soubor arriba_discarded_standard.tsv),
- fúze vyřazené filtry Arriba a uvedené v COSMIC ve formátu TSV (soubor arriba_discarded_cosmic.tsv),
- různé obrázky zobrazené v rozhraní ve formátu SVG.

13.1.10. Klonálnost

„IGHV mutational status report“ v záložce „Clonality“ analýzy SomaHemato zobrazuje výsledky mutačního stavu.

- IGHV Mutační stav a procento homologie:
 - Mutated: pokud je homologie <98% identity se zárodečnou linií
 - Unmutated: pokud je homologie ≥98% totožná se zárodečnou linií.
 - Borderline: pro klony v rozmezí 97-98 % identity.
 - Polyclonální: žádný významný klon nad 20 %
 - Non-Conclusive: pokud některý ze základních QC selže.
- Určená podmnožina (s klinickým dopadem) a spolehlivost: *2 (vysoké riziko), *8 nebo žádná
- Varovné upozornění, pokud se ve zprávě objeví varování.

V části "Clonality" jsou uvedeny identifikované klony s jejich četností a stavem produktivity (Productive, Un-productive).

Každý klon má vyhrazený oddíl, ve kterém je uveden jeho mutační stav, frekvence, odkaz na stránku IMGT_quest a další podrobnosti (homologie, úplnost) spolu s kontrolou kvality.

13.1.11. Varianty spojování

Metodika

SomaRNA workset využívá sadu nástrojů *Regtools* k detekci přeskoků exonů a identifikaci alternativních sestřihových variant ze sekvenování RNA obohaceného o zachycení zaměřené na panel genů. *Regtools* extrahuje exon-exonové spoje ze souboru BAM vygenerovaného zarovnávačem dat RNA-seq STAR a anotuje je podle známé struktury transkriptu na základě souboru GTF a anotací Ensembl.

Vizualizace a interpretace

Záložka Splice variants analýzy SomaRNA se zaměřuje na vizualizaci splice variant následujících klinicky relevantních genů a jejich hlavní akční varianty:

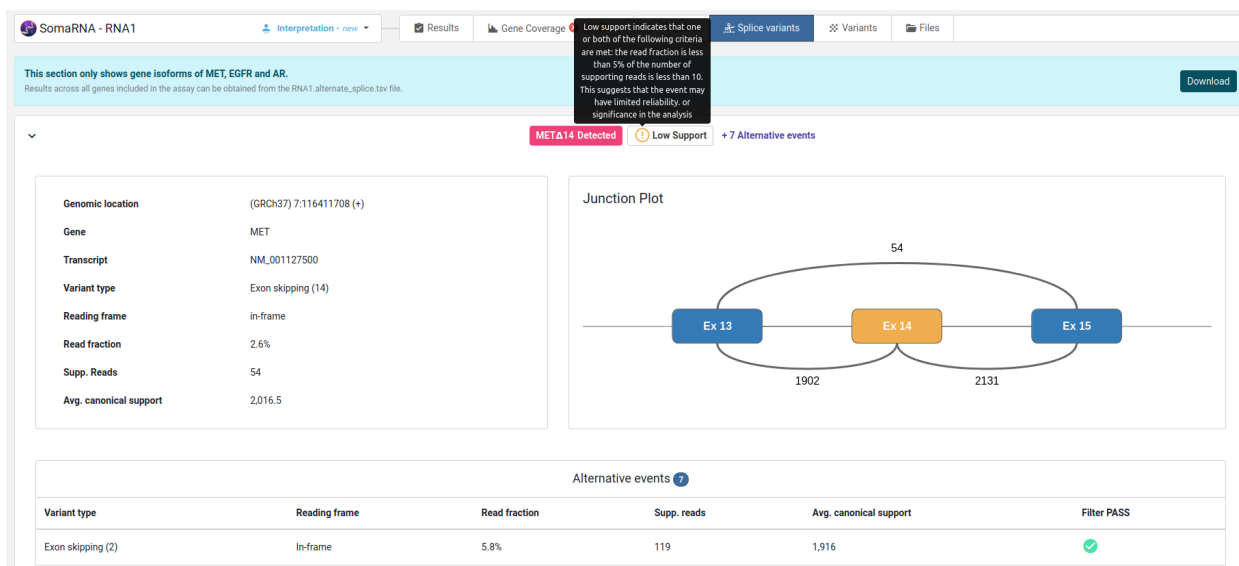
- MET (protoonkogen MET, receptorová tyrozinkináza) a akční varianta MET Δ 14
- EGFR (receptor epidermálního růstového faktoru) a jeho akční splice variant EGFRvIII
- AR (androgenní receptor) a akční varianta AR-V7

Kompletní seznam variant splicing zjištěných v analýze SomaRNA, kromě výše uvedených, lze stáhnout přímo jako soubor TSV pomocí ikony Download v rozhraní nebo na kartě Files v analýze.

U každého výše uvedeného genu se zobrazí stav detekce (detekováno nebo nedetekováno) hlavní akční varianty spolu s počtem dalších alternativních událostí detekovaných pro daný gen.

Podrobnosti o hlavních zjištěných variantách spoje si uživatelé mohou prohlédnout otevřením speciální vložky, která je rozdělena do tří částí:

- Vlevo jsou zobrazeny informace o hlavní akční variantě sestřihu: genomické souřadnice, název genu, transkript RefSeq, typ zjištěné varianty s podrobným popisem chybějících exonů, fáze čtení, podíl čtení a počet čtení podporujících variantu. Nakonec je uveden průměrný počet čtení s kanonickou podporou. To odpovídá průměru 5' a 3' reads podporujících zařazení exonu.
- vpravo je zobrazen graf sashimi (Junction Plot), který znázorňuje zjištěné spoje mezi exony a počet čtení podporujících každý spoj.
- v dolní části jsou uvedeny informace o zjištěných alternativních událostech, tj. izoformách jiných než hlavní varianta zájmu popsaná výše: typ zjištěné varianty (chybějící čísla exonů v závorkách), fáze čtení, podíl a počet čtení podporujících variantu, průměrný počet čtení podporujících kanonický transkript (Avg. canonical support), stav PASS nebo FAILED podle níže popsaných filtrů.



Obrázek 35: Vizualizace záložky *Splice variants* analýzy SomaRNA.

Detekce sestřihových variant v genech zobrazených na kartě *Splice variants* nepodléhá použití detekčních filtrů.

Ikona "Nízká podpora" spojená s detekcí MET Δ 14 a EGFRvIII však uživateli ukazuje, že metriky týkající se podporujících čtení jsou pod následujícími prahovými hodnotami: podíl čtení < 5 % nebo počet podporujících čtení < 10.

Tyto prahové hodnoty se používají i pro stav PASS nebo FAILED u ostatních alternativních událostí:

- **PASS** : podíl reads $\geq$ 5 % a počet reads $\geq$ 10.
- **Failed** : podíl reads < 5 % nebo počet reads < 10.

Soubor TSV, který lze stáhnout, obsahuje následující informace o všech zjištěných událostech spojování: RefSeq transcript, gene, strand direction (+ or -), fraction of reads and number of reads supporting the variant, number of missing exons and corresponding exon numbers, reading phase, genomic coordinates of junctions. V souboru je také v části FILTER_PASS (yes/no) uvedeno, zda varianta projde prahovými hodnotami použitými na podpůrná čtení (podíl čtení $\geq$ 5 % a počet čtení $\geq$ 10), a v části Actionable (yes/no) je uvedena akční způsobilost varianty.

13.1.12. Stav MSI/MSS

Pracovní sada analýzy SomaMSI je k dispozici při spuštění nové analýzy na datech genového panelu a slouží k určení stavu MSI nebo MSS vzorku nádoru.

Metodika

Tato neindividuální kohortová analýza je podobná analýze CNVCapture nebo SomaCNVCapture. K provedení analýzy je zapotřebí nejméně osm vzorků.

Nástroj vygeneruje referenční vzorek ze vzorků v kohortě, u nichž se odhaduje, že jsou negativní. Protože tento stav vzorků není předem znám, je tento odhad založen na výpočtu míry delecí pozorovaných v mikrosatelitních oblastech: polovina vzorků v kohortě s nejnižší mírou delecí je proto vybrána k vytvoření referenčního vzorku, přičemž každý vzorek tak přispívá částí svých sekvencí.

Každý vzorek nádoru je následně porovnán s referenčním vzorkem pomocí nástroje MSIsensor. Vypočítá se skóre, které představuje procento studovaných mikrosatelitních oblastí, které vykazují nestabilitu. Studované oblasti jsou mikrosatelitové oblasti, které pokrývá panel a které jsou detekovány pomocí MSIsensor.

Na základě tohoto skóre je stav MSI nebo MSS různých vzorků dán prahem, který se vypočítá takto: $\max(10, \text{medián} + \max(3 \cdot \text{median-abs-dev}, 0,3 \cdot \text{median}))$.

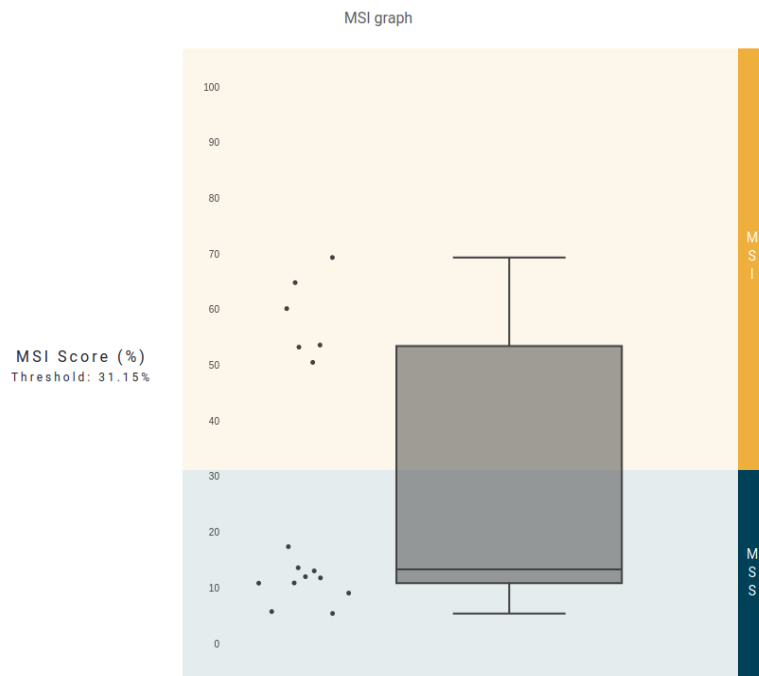
Princip tohoto hodnocení je následující:

- Při analýze vzorků MSI a MSS se práh obvykle rovná mediánu + $3 \cdot \text{median-abs-dev}$.
- Pokud však mají všechny vzorky stav MSS, pokud je jejich skóre velmi nízké a pokud je jejich rozložení velmi blízké, bude prahová hodnota nastavena na 10, aby nedocházelo ke generování falešně pozitivních výsledků.

Zobrazení a výklad

Po dokončení analýzy lze z analýzy SomaMSI zobrazit výsledky všech vzorků. Záložka Results (Výsledky) je rozdělena na dvě části:

MSI graph: grafické znázornění výsledků, kde je každý vzorek zobrazen podle svého skóre. Prahová hodnota určená podle kohorty vzorků vytváří v grafu dvě zóny představující vzorky považované za MSI, respektive MSS.



Obrázek 36: Krabicový graf znázorňující rozložení skóre MSI pro výběrovou kohortu.

Per sample details: rozdělení skóre a stát přiřazený ke každému vzorku jsou k dispozici v pravém panelu.



Obrázek 37: Rozdělení skóre a stavu všech vzorků zahrnutých do analýzy SomaMSI.

13.1.13. Stav HRD/HRP

Metodika

Detekce stavu HRD navržená společností SeqOne je založena na analýze dat sekvenování genomu s nízkým pokrytím (sWGS), ke kterým lze přidat výsledky analýzy genového panelu, pokud tento zahrnuje geny BRCA1 a BRCA2.

Princip je následující:

- Kvantifikovat genomickou nestabilitu z dat sWGS a identifikovat velké přestavby (LGA) a ztráty rodičovských kopií (LPC), které by odrážely ztrátu heterozygotnosti.
- Identifikovat změny v počtu kopií (CNV) v klíčových genech, u nichž je popsán potenciální vliv na HRD status.
- Volitelně detekovat patogenní varianty v genech BRCA1, BRCA2 a TP53, pokud jsou pro příslušného pacienta k dispozici výsledky panelu v rozhraní SeqOne. V tomto případě se berou v úvahu typy variant zjištěné pracovní sadou SomaVar: SNV, indely a SV (zjištěné pomocí Freebayes, Mutect2 a GRIDSS).

HRD status. HRD status. je definován alespoň jedním z následujících údajů (**pouze verze SomaHRD v1.2 a v1.3**): (1) alespoň jedna patogenní mutace BRCA1/2 a/nebo (2) HRD probability větší nebo rovna 0,5.

Od verze 1.4 worksetu SomaHRD je status analýzy definován pouze skóre genomické nestability (GI skóre). Uživatelé mohou vedle stavu GI vybrat stav HRD (pozitivní, negativní nebo nekonkluzivní) a zahrnout jej do zprávy místo stavu GI.

HRD probability. (**pouze verze SomaHRD v1.2 a v1.3**) Pravděpodobnost **HRD SeqOne** odpovídá pravděpodobnosti, že nádor je HRD pozitivní na základě charakteristik genomové nestability LGA a LPC stanovených z dat sWGS. Velké genomové změny (LGA) jsou definovány jako počet zlomových bodů, kde zlomový bod odpovídá změně počtu kopií mezi dvěma genomovými segmenty o délce nejméně 10 Mb a vzdálenými od sebe nejvýše 3 Mb. Losses of Parental Copy (LPC) jsou definovány jako počet haploidních segmentů dlouhých nejméně 10 Mb. HRD probability je lineární kombinací obou charakteristik genomové nestability, převedenou pomocí logistické regrese tak, aby se hodnoty umístily na stupnici pravděpodobnosti mezi 0 a 1.

GI probability (GiP) (ze **SomaHRD v1.4**) Pravděpodobnost genomové nestability (GI) představuje pravděpodobnost, že nádor vykazuje genomovou nestabilitu, na základě mělkého sekvenování celého genomu (sWGS). Výpočet této pravděpodobnosti závisí na dvou typech charakteristik genomové nestability LGA a LPC. LGA je definována jako počet zlomových bodů počtu kopií. Bod zlomu odpovídá změně počtu kopií mezi dvěma dostatečně velkými genomovými segmenty v těsné blízkosti. Skóre LGA není jediným měřítkem, ale spíše souborem několika variant. Tyto

odchylky se vypočítávají pomocí různých parametrů pro minimální velikost fragmentu a vzdálenost mezi fragmenty. Skóre LPC se naproti tomu vztahuje k počtu haploidních segmentů o délce alespoň 10 Mb. Pravděpodobnost genomové nestability SeqOne není přímou funkcí průměrných skóre LGA a LPC. Vypočítává se pomocí posilovacího modelu, který integruje několik funkcí LGA a LPC, čímž zvyšuje přesnost předpovědi. Konečná pravděpodobnost je průměrem 100 pravděpodobností (v rozmezí od 0 do 1) vypočítaných pomocí různých parametrů modelu.

GI status. (ze **SomaHRD v1.4**) **Pozitivní** stav genomové nestability (GI status) je definován pravděpodobností genomové nestability vyšší než 50 %.

BRCA Status. Pro detekci patogenních variant a určení statusu BRCA nezávisle na pravděpodobnosti HRD nebo pravděpodobnosti GI se na varianty BRCA1 a BRCA2 z analýzy SomaVar pacienta použijí následující kritéria:

- Allele frequency $\geq 15\%$ (**verze SomaHRD v1.2 a v1.3**),
- Allele frequency $\geq 10\%$ (ze **SomaHRD v1.4**)
- Coverage ≥ 35
- Průměrná kvalita každé báze nesoucí variantu (skóre Phred) ≥ 20
- Náhradní pozorování (počet sekvencí nesoucích mutací) ≥ 5
- Klasifikace 4 (*Likely Pathogenic*) nebo 5 (*Pathogenic*) podle našeho nástroje pro klasifikaci patogenity variant na základě kritérií ACMG.

TP53 variants. Na patogenní varianty TP53 z analýzy SomaVar provedené u pacienta se vztahují následující kritéria:

- Coverage ≥ 100
- Průměrná kvalita každé báze nesoucí variantu (skóre Phred) ≥ 20
- Náhradní pozorování (počet sekvencí nesoucích mutací) ≥ 5
- Klasifikace 4 (*Likely Pathogenic*) nebo 5 (*Pathogenic*) podle našeho nástroje pro klasifikaci patogenity variant na základě kritérií ACMG.

Tyto varianty se ve zprávě o analýze SomaHRD zobrazují pouze bez minimální prahové frekvence alel, ale nejsou určující pro HRD status. Ve zprávě HRD bude uvedena pouze patogenní varianta TP53 s nejvyšší frekvencí alel.

Cellularity. Metadata o buněčnosti (% nádorových buněk ve vzorku) jsou nutná ke spuštění analýzy SomaHRD. Lze je zadat při načítání vzorku nebo před spuštěním analýzy úpravou příslušných metadat vzorku (viz **Metadata** vzorku). Pokud je hodnota nižší než 20 %, výsledek bude neprůkazný. HRD status však může být pozitivní, pokud je analýza doplněna panelem zahrnujícím BRCA1 a BRCA2, který udává pozitivní stav BRCA.

Postup varování před nádorovou frakcí: "Analysis" SomaHRD obsahuje výstražný mechanismus, který indikuje, že vzorek může mít nízký obsah nádorové frakce. Ten je založen na statistickém opatření zvaném "statistika překrývání", které hodnotí podobnost mezi po sobě jdoucími

genomovými segmenty z hlediska poměru počtu kopií. Vzorky s výrazným překryvem mají často nižší podíl nádorů, což ztěžuje rozlišení genomové nestability.

Abychom stanovili spolehlivý práh pro toto varování, korelovali jsme statistiku překrývání s frekvencí variantních alel (VAF) mutací TP53 pomocí klinických údajů. Výstraha se spustí, pokud statistika překrývání překročí předem definovanou prahovou hodnotu. Toto upozornění upozorňuje na potenciální nesrovnalosti s ručně zadaným procentem obsahu nádoru, což může znamenat, že podíl nádoru je nižší než 20 %. U vzorků s podílem nádorů pod touto hranicí je zvýšené riziko vzniku falešně negativních výsledků. Takové vzorky však nejsou odmítány a upozornění slouží jako vodítko pro pečlivou interpretaci výsledků HRD.

Minimum coverage. Minimální pokrytí potřebné pro provedení testu je 0,1x, ale doporučujeme minimální pokrytí 0,5x, ideálně 1x. Při snížení pokrytí na 0,1X se chybovost klasifikace zvýší o 5 %.

Ačkoli data sWGS mohla být generována ze soupravy, která má UMI, při analýze SomaHRD se k nim nepřihlíží. V pracovní sadě je zahrnut krok ořezávání pro adaptéry i UMI.

Vizualizace a interpretace

Po dokončení analýzy jsou výsledky k dispozici v analýze SomaHRD. Kromě "QC report" a "Technical report" je na kartě HRD Report (zpráva HRD) přímo přístupná zpráva HRD (**verze SomaHRD v1.2 a v1.3**).

Od **verze v1.4 worksetu SomaHRD** se výsledky analýzy zobrazují na zvláštní kartě HRD. Každá část karty odráží část zprávy HRD popsanou níže (HRD Summary, Sample quality, Genomic instability features, BRCA1/BRCA2 & TP53 mutations, Copy number variations).

Na této kartě může uživatel tyto výsledky exportovat jako soubor CSV nebo jako zprávu HRD ve formátu PDF. V případě exportu do formátu PDF může uživatel vybrat jazyk zprávy HRD ze seznamu dostupných jazyků: Francouzština, španělština, angličtina, italianština, němčina. Výchozí volbou je angličtina.

HRD report je rozdělen do následujících oddílů.

V záhlaví jsou uvedena metadata pacienta a analýzy, typ provedeného testu a stav HRD (**SomaHRD v1.2 a v1.3**) **nebo stav GI (SomaHRD ≥ v1.4)**.

Sample quality. V této části jsou shrnuty hlavní ukazatele kvality analýzy a minimální požadované prahové hodnoty.

Kontroly QC se zobrazují jako tabulka. Pro každý parametr jsou k dispozici následující informace:

- **Název** : název parametru
- **Hodnota** : hodnota parametru, včetně :
 - Ikona označující, zda je filtr *vyhovující, varovný nebo* neúspěšný.
 - Hodnota parametru pro vzorek
- **Description** : Krátký popis parametru

SAMPLE QUALITY

NAME	VALUE	DESCRIPTION
% of tumoral cells	✓ 90%	Low value of tumoral cells is likely to generate false negative results
Coverage	⚠ 0.43X	Average sequencing coverage of the genome
% correct mapping	✓ 96.5%	Proportion of reads correctly mapped
Confidence in genomic instability	✓ 1.00	Large values indicates high confidence in genomic instability computation
Low tumor fraction	✓ NORMAL	Low tumor fraction detected from the shallow WGS

SAMPLE QUALITY

NAME	VALUE	DESCRIPTION
% of tumoral cells	⚠ 30%	A percentage of tumoral cells below 20% is likely to generate false negative results.
Coverage	✓ 0.7X	The average sequencing coverage of the genome should be above 0.1x.
GI Dispersion	✓ 0.93	A large dispersion reflects high variability in the genomic instability computation, indicating low precision in the estimated score.
Low tumor profile	⚠ WARNING	Low cellularity detected from the genomic profile.

Obrázek 38 : Shrnutí metrik QC souvisejících se vzorkem sWGS v SomaHRD v1.2 (nahore) a SomaHRD v1.5 (dole).

Kontroly kvality se provádějí u následujících parametrů podle prahových hodnot uvedených na konci zprávy HRD (**oddíl VZOROVÉ KONTROLY KVALITY**):

- procento nádorových buněk pouze na základě hodnoty zadané uživatelem.
- průměrné pokrytí sekvenování genomu,
- podíl správně zarovnaných reads vypočtený po kroku ořezání adaptérů (**pouze verze SomaHRD v1.2 a v1.3**),
- metrika důvěryhodnosti genomické nestability (LGA a LPC) na základě údajů o pokrytí (**pouze verze SomaHRD v1.2**). Tato kontrola kvality se používá k měření důvěryhodnosti vráceného stavu HRD status. Některé pozitivní a negativní výsledky jsou označeny štítkem "low confidence" (nízká důvěryhodnost), který identifikuje vzorky s nejnižší důvěryhodností stavu.
- a interval spolehlivosti HRD (*interval spolehlivosti HRD*) (**pouze verze SomaHRD v1.3**). Tato metrika bude mít stav *Warning (Varování)*, pokud je standardní chyba (*směrodatná*

odchylka) větší než 0,07; bude mít stav Failed (Nevyhovující), pokud interval spolehlivosti pravděpodobnosti HRD obsahuje hodnotu 0,5.

- interval spolehlivosti pravděpodobnosti GI (interval spolehlivosti pravděpodobnosti GI) (**pouze verze SomaHRD v1.4**). Tato metrika bude mít stav *Warning* (Varování), pokud je standardní chyba (*směrodatná odchylka*) větší než 0,07; bude mít stav Failed (Nevyhovující), pokud interval spolehlivosti pravděpodobnosti GI zahrnuje hodnotu 0,5.
- rozptyl GI (**pouze verze SomaHRD v1.5**), který ukazuje rozptyl 100 výpočtů pravděpodobnosti GI. Tato metrika zobrazí stav "Warning" s označením "Moderate" vedle hodnoty, pokud je rozptyl > 2,6, a stav "Failed" s označením "High", pokud je > 2,7. V opačném případě se vedle skóre zobrazí označení "Low".
- v případě podezření na nízký obsah nádoru (*Low tumoral fraction*) na základě sWGS se zobrazí zpráva WARNING. Pokud tomu tak není, zobrazí kontrola zprávu NORMAL (*normální*).

Pokud je některá z následujících metrik QC pod očekávanými prahovými hodnotami selhání, bude výsledek analýzy neprůkazný (NON CONCLUSIVE):

- % tumor cells
- Coverage
- % správně zarovnaných čtení (pouze **pro verze SomaHRD v1.2 a v1.3**).
- GI disperze > 2,7, **pokud také** $0,37 < GiP < 0,59$ (**pouze pro SomaHRD v1.5**)

V tomto případě se metriky důvěryhodnosti genomové nestability a nízké nádorové frakce nepočítají a v příslušných oddílech se zobrazují N/A (pro verzi **SomaHRD < v1.5**).

Pro SomaHRD v1.3, Pokud je důvěra v metriku genomické nestability Failed, tj. pod očekávaným prahem neúspěšnosti, výsledek analýzy bude:

- POSITIVE LOW CONFIDENCE" pokud je pravděpodobnost HRD větší nebo rovna 50 %.
- NEGATIVE LOW CONFIDENCE, pokud je pravděpodobnost HRD menší než 50 %.

V tomto případě je výsledek analýzy doprovázen varováním.

Nicméně pokud je analýza doprovázena panelem obsahujícím BRCA1 a BRCA2 a panel vrátí pozitivní stav BRCA, bude stav HRD pozitivní, i když je metrika kontroly kvality týkající se vzorku sWGS pod prahem selhání (pouze pro **verze SomaHRD v1.2 a v1.3**).

Copy number variations. Tento oddíl je věnován Copy number variations genů CCNE1 a RAD51 z hlediska počtu zjištěných kopií. Ty se při výpočtu pravděpodobnosti HRD a pravděpodobnosti GI neberou v úvahu.

Genomic instability features. V této části jsou uvedeny podrobnosti o vlastnostech, které byly vypočteny pro určení genomové nestability nádoru. Tyto rysy se používají k určení pravděpodobnosti SeqOne HRD (**SomaHRD verze 1.2 a 1.3**) a pravděpodobnosti GI (SomaHRD

verze ≥ 1.4). Tento oddíl je určen k pochopení způsobu výpočtu pravděpodobnosti. Pro každou z funkcí jsou uvedeny následující informace:

- **Název** : Název prvku (např. LGA Status nebo LGA Score).
- **Value** : Hodnota vypočtená pro pacienta.
- **Popis** : Stručný popis funkcí

V případě **NEZÁVĚREČNÝCH** výsledků, které se týkají metrik kontroly kvality pod očekávanými prahovými hodnotami, se sekce **AMPLIFICATIONS** a **GENOMIC INSTABILITY FEATURES** ve zprávě zobrazí šedě.

GENOMIC INSTABILITY FEATURES

The following variables are used to determine the HRD Probability.

NAME	VALUE	DESCRIPTION
LGA Status	23	Large Genomic Alteration Status
LPC Status	15	Loss of Parental Copy Status

Obrázek 39: Znaky pravděpodobnosti HRD související s genomickou nestabilitou.

BRCA1/BRCA2 & TP53 mutations. V případě kompletního HRD testu včetně analýzy SomaVar poskytujeme další část věnovanou patogenním variantám zjištěným v genech BRCA1, BRCA2 a TP53. Jsou zde uvedeny všechny patogenní varianty genů BRCA1 a BRCA2 odpovídající kritériím popsaným v části **Metodika**. Uvedena však bude pouze patogenní nebo pravděpodobně patogenní varianta TP53 s nejvyšší frekvencí alel. Ostatní patogenní varianty TP53 lze zobrazit v související analýze SomaVar. Tento oddíl je skrytý pouze pro "Report" sWGS.

Zjištěné mutace jsou uvedeny v tabulce s následujícími sloupci:

- **Gene & Transcript:** Gene & Transcript: RefSeq transcript ID
- **Variant:** HGVSC a HGVSP nomenklatura varianty
- **Evaluation:** Variant evaluation according to ACMG criteria (*Pathogenic or Likely Pathogenic*)
- **VAF:** Alelická frakce varianty

Pokud nebyly v analýze propojeného panelu identifikovány žádné patogenní mutace BRCA, zobrazí se v této části zprávy následující zpráva: "No pathogenic mutations have been identified on BRCA1 and BRCA2 genes."

Pokud nebyly v analýze propojeného panelu identifikovány žádné patogenní mutace TP53, zobrazí se v této části zprávy následující zpráva: "No pathogenic mutations have been identified on TP53 gene."

BRCA1/BRCA2 & TP53 MUTATIONS

GENE & TRANSCRIPT	VARIANT	EVALUATION	VAF
BRCA1 NM_007294.4	c.3116_3117del p.Ala1039GlufsTer6	Pathogenic	60.68%
BRCA1 NM_007294.4	c.3113_3117delinsG p.Glu1038GlyfsTer9	Pathogenic	58.18%
BRCA1 NM_007294.4	c.3113_3114del p.Glu1038GlyfsTer7	Pathogenic	60.08%
TP53 NM_000546.6	c.524G>A p.Arg175His	Pathogenic	39.29%

Obrázek 40: TP53 variants.

HRD summary. Tato poslední část zprávy je věnována výsledkům. Cílem této části je poskytnout podrobný pohled na výsledky a zobrazit následující informace:

- **HRD Probability SomaHRD v1.2 a v1.3):** HRD probability value, mezi 0 a 1. Status je kladný, pokud je hodnota větší nebo rovna 0,5.
- **GI probability (SomaHRD verze $\geq$ v1.4):** Stav je pozitivní, pokud je tato hodnota větší nebo rovna 0,5.
- **BRCA Status:** status určený detekcí alespoň jedné patogenní mutace BRCA: MUTATED: DIVOKÝ TYP nebo N/A v případě, že není provedena panelová analýza.
- **HRD Status or GI Status :** POSITIVE, NEGATIVE, NEGATIVE LOW CONFIDENCE, POSITIVE LOW CONFIDENCE or NON CONCLUSIVE as displayed in the report header
- **Metoda :** Metoda (nebo způsob), která byla použita k vytvoření skóre HRD:
 - Úplný HRD $\Rightarrow$ 'WGS SHALLOW + BRCA1/2 GENE PANEL'
 - Pouze mělké $\Rightarrow$ "SHALLOW WGS ONLY"

V případě výsledků *NEGATIVE LOW CONFIDENCE*, *POSITIVE LOW CONFIDENCE* nebo *NON CONCLUSIVE* jsou tyto výsledky zobrazeny světle šedou barvou.

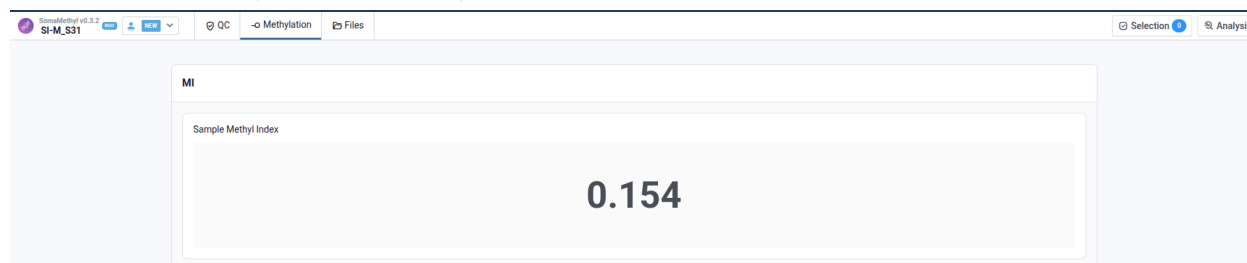
Podrobnosti o jednotlivých hodnotách zobrazených ve zprávě a o celkové metodice testu jsou uvedeny v oddílech "**Footnotes**" a "**Methodology**", které následují za výsledkem.

Kromě zprávy HRD jsou na kartě Soubory analýzy SomaHRD k dispozici ke stažení grafy genomické nestability spojené s analýzou HRD. Tyto grafy představují normalizované pokrytí všech chromozomů z dat sWGS.

V případě **NEZÁVĚŘEČNÝCH** výsledků, které se týkají metrik kontroly kvality pod očekávanými prahovými hodnotami, se grafy genomové nestability negenerují.

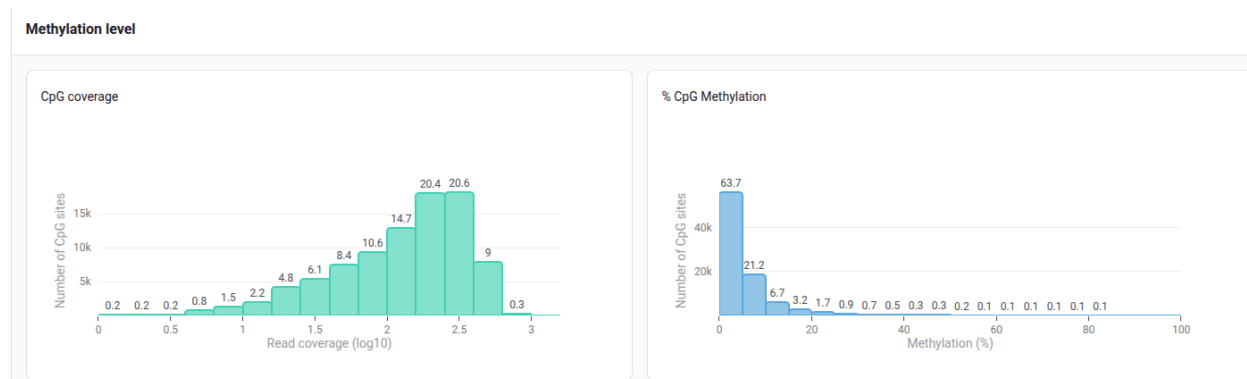
13.1.14. Metylační data pro SomaMethyl

Methylation Index (Index metylace)



Index metylace (MI) je jediné číselné skóre, které kvantifikuje celkový stav metylace biologického vzorku a shrnuje signál metylace spojený s nádorem. Dokáže citlivě detekovat i nízké hladiny DNA pocházející z nádoru na pozadí normální DNA. V SomaMethyl se MI počítá ve standardním režimu pomocí proprietárního algoritmu společnosti Agilent (MethylAlgo). Pokud je k dispozici referenční kohorta, lze MI vzorku porovnat s hodnotami MI kohorty a interpretovat pomocí prahové hodnoty definované uživatelem. Pro vyhodnocení MI poskytla společnost Agilent titrační vzorky s očekávanými hladinami metylace. Hodnoty MI vypočítané modulem SomaMethyl ve standardním režimu se porovnají s očekávanými hodnotami MI poskytnutými společností Agilent a výkon se shrne pomocí vážené metriky shody definované jako $1 - wMAPE$ (Vážená průměrná absolutní procentuální chyba).

Methylation level (Úroveň metylace)



CpG coverage (histogram) (Pokrytí CpG (histogram))

Distribuce hloubky čtení napříč místy CpG v cílových oblastech (manifest).

Osa X (pokrytí čtení log10): pokrytí na místo CpG (počet čtení podporujících volání metylace na daném CpG, tj. počet metylovaných + nemetylovaných).

Osa Y (# míst CpG): počet míst CpG v manifestu, která mají toto pokrytí.

Můžete zkontrolovat, zda jsou cílové CpG dostatečně pokryty a zda je pokrytí dostatečně rovnoměrné pro spolehlivou interpretaci procent metylace.

% CpG methylation (histogram) (% metylace CpG (histogram))

Distribuce úrovní metylace napříč místy CpG v cílových oblastech (manifest).

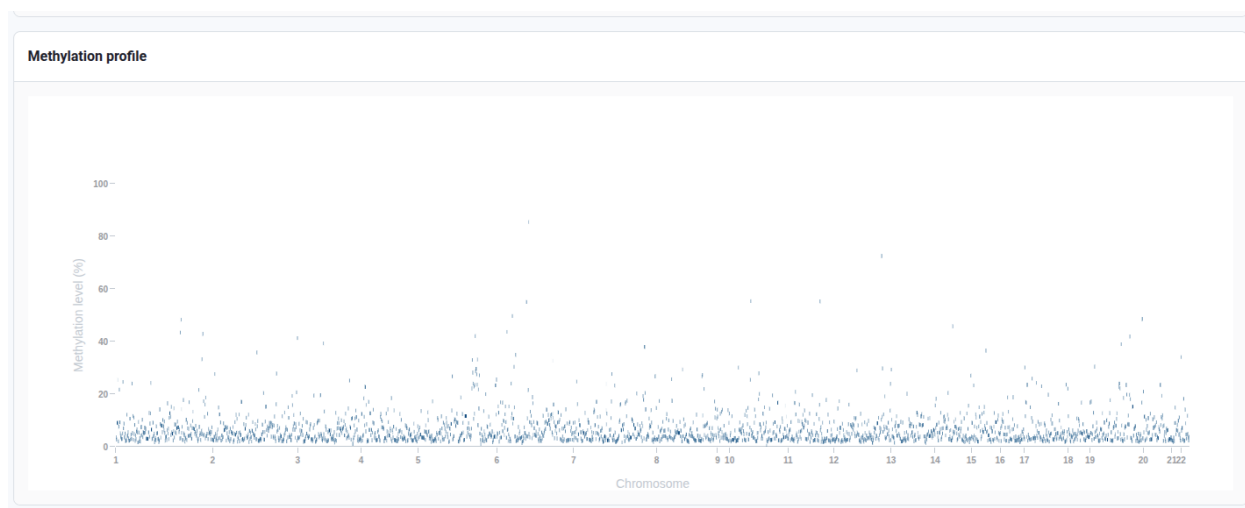
Osa X (% metylace): pro každé místo CpG zlomku metylovaných čtení ze všech čtení pokrývajících toto CpG.

Osa Y (# míst CpG): počet míst CpG v manifestu vykazujících tuto úroveň metylace.

Poskytuje pohled na úrovní vzorku na stavy metylace CpG napříč cílovými oblastmi (často ukazuje vrcholy v blízkosti nízké a vysoké metylace v závislosti na zachycených oblastech a biologii vzorku).

Místa CpG s velmi nízkým pokrytím poskytují méně spolehlivá procenta metylace, protože malý počet čtení může % vychýlit nahoru nebo dolů. Z tohoto důvodu by měl být histogram % metylace CpG vždy interpretován společně s histogramem pokrytí CpG (a dalšími metrikami kontroly kvality).

Methylation profile (Profil metylace)



Graf profilu metylace ukazuje signál metylace CpG vzorku napříč genomem, zobrazený podle chromozomu. Každý bod odpovídá průměrnému procentu metylace CpG vypočítanému v po sobě jdoucích genomických oknech (jako ve výstupu `.mcpG.track.bedGraph`). Tento pohled pomáhá vizualizovat široké vzorce metylace a identifikovat oblasti s vyšší nebo nižší metylací ve srovnání se zbytkem genomu.

Methylation Table (Tabulka metylace)

Table

Search genes

Locus

▼

FT

REGION	METH. LEVEL	CpG NBR.	CpG REGIONS	GENOMIC FEATURE	REFSEQ CDS
C1orf94 chr1:34176744-34177348	11	53	Island +1	silencer	
chr1:35577191-35577675	3	42			
chr1:37033668-37034149	5	51			
chr1:37034542-37035023	3	61			
EPHA10 chr1:37735070-37735507	24	26		enhancer	NM_001099439.2
chr1:38559361-38559798	5	22			
chr1:39080823-39081222	12	28	Island +1		
chr1:39672120-39672554	3	28			
HPCAL4 chr1:39683853-39684273	11	47	Island		
chr1:40884160-40884512	5	43			

31 to 40 of 2217

<

2

3

4

5

6

>

Tabulka metylace uvádí genomické oblasti (chromozom:začátek-konec), ve kterých SomaMethyl shrnuje metylaci CpG (pevná okna nebo sloučené segmenty, v závislosti na kroku). Pro každou oblast METH. LEVEL hlásí agregovanou úroveň metylace vypočítanou z volání metylace CpG v tomto intervalu a CpG NBR. udává, kolik míst CpG přispělo k výpočtu. Každý řádek je poté anotován pomocí CpG REGIONS (ostrov, břeh, šelf, kde je to vhodné) a funkčními anotacemi, včetně překrývajících se regulačních prvků GENOMIC FEATURE (promotor, enhancer, silencer) a informací REFSEQ CDS (kódující sekvence).

14. Zpráva

14.1.1. Výstupní soubory

Na kartě *Soubory* jsou uvedeny soubory vytvořené v různých fázích worksetu. Každý soubor lze stáhnout kliknutím na ikonu . 

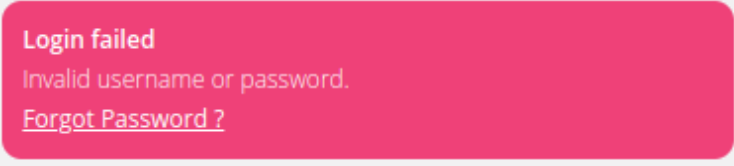
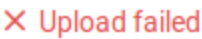
Některé soubory neobsahují název vzorku jako předponu. Chcete-li je stáhnout s tímto prefixem, musíte aktivovat možnost "Prefix with sample name" v pravém horním rohu stránky.

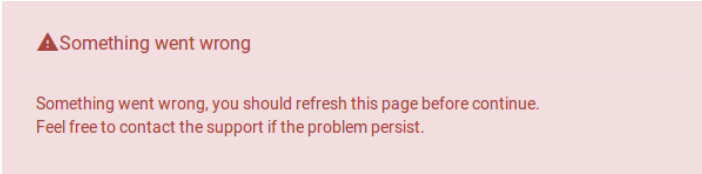
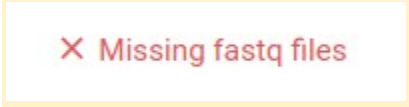
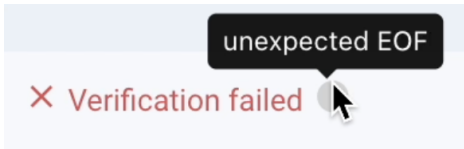
Soubory jsou k dispozici ke stažení v různých sekcích rozhraní:

- Soubory FastQ jsou k dispozici na *kartě* Samples projektu *po* výběru příslušného vzorku a poté ikony . 
- Manifesty (soubory BED) jsou k dispozici v nastavení uživatele prostřednictvím *ikony* v pravém horním rohu ovládacího panelu a poté na kartě Manifesty.
- Zpráva(y) o analýze ve formátu PDF je(jsou) k dispozici na kartě Reporting *po* výběru analýzy.
- Soubory vygenerované během analýzy (BAM, VCF atd.) a různé zprávy HTML (kvalita, technické atd.) jsou k dispozici na kartě Soubory pro každou analýzu.

15. Řešení problémů a chybové zprávy

K upozornění uživatelů na chybu se používá několik zpráv.

Chybová zpráva	Description
 Login failed Invalid username or password. Forgot Password ?	Heslo a/nebo přihlašovací údaje jsou nesprávné. Uživatelé jsou vyzváni, aby znovu zadali své přihlašovací jméno a heslo nebo klikli na "Forgot password?" a obnovili své heslo.
 ... Upload pending	Nahrávání vzorku probíhá. Uživatelům doporučujeme vyčkat, nezavírat okno prohlížeče a nepřerušovat připojení k internetu.
 × Upload failed	Vzorek se nepodařilo nahrát. <i>K této chybě může dojít, pokud bylo okno prohlížeče zavřeno nebo pokud během nahrávání došlo k výpadku internetového připojení.</i> Uživatelé jsou vyzváni, aby se pokusili soubor(y) nahrát znovu.
	Ukazuje průběh aktuálního nahrávání Uživatelům doporučujeme vyčkat, nezavírat okno prohlížeče a nepřerušovat připojení k internetu.
	Načtení stránky. Uživatelé jsou vyzváni, aby počkali.

... Bioinformatics - <i>pending</i>	Na analýzu se čeká. Uživatelé jsou vyzváni, aby počkali.
× Failed - <i>error</i>	Analysis failed: Uživatelé jsou vyzváni, aby zkontrolovali nastavení a data použitá ke spuštění analýzy a poté ji znovu spustili.
	Došlo k chybě. Uživatel je vyzván, aby kliknul na klávesu "Esc". Pokud problém přetrvává, kontaktujte podporu na adrese support@seqone.com.
	U tohoto vzorku chybí související soubory FastQ, což znamená, že pro tento vzorek nelze spustit žádné nové analýzy.
	Zpráva "Unexpected EOF" (End Of File) odpovídá zkrácenému souboru. Uživatel je vyzván, aby zkontroloval data a znovu načetl příslušné soubory.

Uživatelé mohou kontaktovat tým podpory kliknutím na logo  v pravém horním rohu uživatelského panelu a poté na ikonu  Support

16. Záruka

Viz záruční podmínky ve vaší smlouvě o údržbě.

17. Právní upozornění

17.1. Podmínky použití

Informace o uživatelské licenci získáte kliknutím na logo uživatelského účtu, po kterém následují *"Terms of use"*. Uživatelé si mohou informace zobrazit a stáhnout ve formátu PDF z vyskakovacího okna kliknutím na *"OPEN PDF"*.

17.2. Omezení odpovědnosti

Společnost SeqOne se zříká jakékoli odpovědnosti za jakékoli zneužití nebo neodborné použití platformy SeqOne.

SeqOne je nástroj pro podporu klinického rozhodování a není zodpovědný za diagnózy odvozené ze získaných výsledků.

Uživatelé proto musí při používání softwaru a jeho výsledků postupovat s náležitou opatrností. Diagnózy by měly být potvrzeny různými klinickými a lékařskými vyšetřeními.

17.3. Autorská práva a ochranné známky

Názvy společností a produktů uvedené v této příručce mohou být ochrannými známkami příslušných vlastníků.