



Platforma SeqOne

Návod na použitie

Platné od verzie platformy SeqOne v2.1

1. Úvod	7
2. Kontaktujte výrobcu a zástupcu spoločnosti Swiss	8
3. Vigilancia zdravotníckych pomôcok	8
4. Určené použitie platformy	8
4.1. Zamýšľané použitie	8
4.2. Indikácia pre použitie	9
4.3. Popis	9
4.4. Populácia pacientov	10
4.5. Určené pre používateľov	10
4.6. Charakteristiky	10
4.6.1. Analytická výkonnosť	12
4.6.2. Bioinformatické moduly	13
4.6.2.1.1. Moduly anotácie	21
4.6.3. Klinická výkonnosť	23
5. Bezpečnostné opatrenia	24
5.1. Správne používanie	24
5.2. Upozornenia	24
5.3. Obmedzenia	25
5.4. Poruchy	26
6. Minimálne požiadavky	27
6.1. Hardvér	27
6.2. Sieť	27
7. Overenie	28
7.1. Prihlásenie	28
7.2. Odhlásenie	29
7.3. Domovská stránka	29
7.4. Správa používateľského účtu	30
7.4.1. Správa "My Account"	31
7.4.2. Správa nastavení používateľa – definície	31
Entita	31
Users	31
Worksety	31
Manifest	31
Panels - Transkript	31
Panels - Gény	32
Marker	32
Protocols	32
Reporting - Global settings	32
Reporting - Template list	33

Tags	33
API Keys	33
7.4.3. Správa administrátora – rozhranie Prezentácia	34
7.4.4. Prepínanie medzi režimami Main a Preview	37
8. Popis operačných krokov	38
9. Vytvorenie projektu	39
9.1. Vytvorenie projektu	39
9.2. Správa projektov	40
9.3. Vymazanie projektu	40
9.4. DNA údaje (sekvenovanie DNA, sekvenovanie ctDNA)	40
9.5. Údaje o RNA (sekvenovanie RNA)	42
9.6. Nastavenia projektu	43
9.6.1. Nastaviť predvolené metriky QC projektu	43
9.6.2. Nastavte predvolený profil filtra	43
10. Ukážka nahratia súboru	44
10.1. Nahratie vzoriek	44
Z počítača	44
Prostredníctvom SDS (SeqOne Data Sync)	44
Stav nahratia vzorky	44
Kapacita nahratia	45
10.2. Formát súboru na zadávanie vzoriek	45
10.2.1. Súbory primárnej analýzy	45
10.2.2. Súbory na sekundárnu analýzu	46
10.3. Vzorové metadáta	48
10.4. Úprava vzorky	49
10.5. Vymazanie vzorky	49
11. Analýza	50
11.1. Spustenie novej analýzy vzorky	50
11.1.1. Výber worksetu analýzy	51
Analytické worksety dostupné z FASTQ súborov	51
GermlineVar	52
GermVar	52
GermlineVarLR	52
GermlineFamily	53
GermVar Family	53
CNVCapture	54
CNVExome	54
SomaVar	54
SomaCNVCapture	55
SomaRNA	55

SomaMSI	55
SomaHRD	56
SomaVar LF	56
SomaCGP	57
SomaHemato	57
SomaLBx	58
SomaMethyl	58
Analytické pracovné sady dostupné zo sekundárnych analytických súborov	58
GermVar Tertiary	59
GermVar Tertiary Family	59
SomaVar Tertiary	59
11.1.2. Výber režimu spracovania UMI	60
Prezentácia režimov spracovania UMI	60
Proces spracovania UMI	61
Kompatibilita s analytickými worksetmi	61
11.2. Stav analýzy	62
11.3. Kvalita vzorky	62
12. Kontrola kvality údajov	64
12.1. Metriky QC a správa	64
12.1.1. Metriky QC a kontroly kvality QC	64
12.1.2. Správa QC	64
12.1.3. Správa o monitorovaní identity „ID-Check“	68
12.1.4. Technický report	69
12.1.5. Správa UPD	70
Skóre UPD	70
Graf frekvencie B-alely (BAF) zo správy UPD	71
13. Interpretácia výsledkov analýzy	72
13.1. Výsledky analýzy	72
13.1.1. Správa interpretácie analýzy Pracovný postup	72
13.1.1.1. Priradovanie analýz na interpretáciu v rámci programu	72
13.1.1.2. Zamknutie analýzy	73
13.1.1.3. Bočný panel pre správu analýz	74
Modálny pohľad na pracovný tok	75
Modálne zobrazenie pacienta	76
Modálny pohľad HPO	77
Modálny panel Zobrazenie	79
13.1.1.4. Bočný panel na správu hodnotení	80
13.1.2. Prehľad výsledkov analýzy na karte Dashboard	81
13.1.3. Podrobné pokrytie podľa exónu	82
Metodika	82

Zobrazenie a interpretácia	83
13.1.4. Malé varianty	86
Metodika	86
Vizualizácia a interpretácia	90
Triedenie variantov	103
Konfigurácia stĺpcov variantov v tabuľk	105
Úprava vzorových metadát z panela analýzy worksetu	106
Filtrovanie variantov	107
Filtry génov a variantov	107
Sekcia Vzorky a kvalita	107
Sekcia Frekvencia	109
Sekcia genetika	110
Sekcia Databáza	111
Sekcia prenosu a fenotypov	112
Sekcia klasifikácie	115
Sekcia Skóre	116
Vyberte indikáciu	118
Použitie preddefinovaného profilu filtra	119
Úprava a uloženie profilov filtrov	119
Oblíbené profily filtrov	119
Kombinovanie profilov filtrov	120
Filtrovanie profilov SeqOne	120
Označiť ako prehliadnuté Funkcia	121
Výber variantov	121
Hodnotenie variantov	121
Stránka variantov	124
Vytvorenie klinickej správy	132
13.1.5. CNV	133
13.1.5.1. CNV v paneloch génov	133
Metodika	133
Zobrazenie a interpretácia	134
13.1.5.2. CNV v SomaCGP a SomaLBx Analýza	139
Metodika	139
Vizualizácia a interpretácia	139
13.1.5.3. CNV v exónoch a genónoch	140
Metodika pre exóm	140
Metodika pre genóm	141
Zobrazenie a interpretácia	141
Tabuľka výsledkov CNV	141

Podrobná stránka CNV	145
13.1.6. Prehliadač veľkých variantov pre CNVs a SVs (LVV)	150
Tabuľka výsledkov CNVs & SVs	150
Triedenie, filtrovanie, prezeranie a interpretácia	152
Bočný panel CNV/SV (Details Drawer)	152
13.1.7. Signatúry v analýze SomaCGP	154
Metodika výpočtu skóre MSI	154
Metodika výpočtu skóre TMB	154
Vizualizácia a interpretácia	155
13.1.8. Krátke tandemové repetície (STR)	157
Metodika	157
Zobrazenie a interpretácia	157
13.1.9. Fúzne gény	159
13.1.9.1. Fúzie zistené z DNA	159
Metodika	159
Zobrazenie a interpretácia	159
13.1.9.2. Fúzie detegované z RNA	161
Metodika	161
Zobrazenie a interpretácia	162
13.1.10. Klonálnosť	165
13.1.11. Varianty splicingu	166
Metodika	166
Vizualizácia a interpretácia	166
13.1.12. Stav MSI/MSS	168
Metodika	168
Zobrazenie a interpretácia	168
13.1.13. Stav HRD/HRP	170
Metodika	170
Vizualizácia a interpretácia	172
13.1.14. Metylačné dáta pre SomaMethyl	177
14. Správa	180
14.1.1. Výstupné súbory	180
15. Riešenie problémov a správy o chybách	181
16. Záruka	183
17. Právne upozornenie	183
17.1. Podmienky používania	183
17.2. Obmedzenie zodpovednosti	183
17.3. Autorské práva a ochranné známky spoločnosti	183

1. Úvod

Pred použitím softvéru si pozorne prečítajte pokyny. Papierovú verziu týchto pokynov vám na požiadanie poskytne tím podpory SeqOne (support@seqone.com). Aby ste dosiahli najlepšie výsledky pri používaní platformy, pozorne si prečítajte všetky bezpečnostné opatrenia označené

symbolom . V tejto príručke a v softvéri sa používajú rôzne regulačné symboly a iné symboly, ktoré sú uvedené nižšie:

Symbol	Význam
	Výrobca
	Označuje autorizovaného zástupcu so sídlom vo Švajčiarsku.
	Dátum výroby (dátum vydania softvéru)
	In vitro diagnostický zdravotnícky prostriedok.
	Jedinečný identifikátor zariadenia
	Referenčné číslo výrobcu v katalógu
	Súlad s požiadavkami európskeho nariadenia (EÚ) 2017/746 Označenie CE s identifikačným číslom notifikovaného orgánu
	Pozrite si návod na použitie
	Upozornenie Tento bezpečnostný symbol označuje, že pri používaní zariadenia je potrebná osobitná opatrnosť.

Štítok (vrátane informácií o verzii softvéru SeqOne a kontaktných údajov) a tento dokument opisujúci návod na použitie (v rôznych jazykoch) sú k dispozícii v časti „About“ (O nás).

2. Kontaktujte výrobcu a zástupcu spoločnosti Swiss



SeqOne S.A.S

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

22 rue Durand
34000 Montpellier
FRANCÚZSKO

Podpora používateľov:

support@seqone.com



Arazy Group Swiss GmbH
Bruderholzallee 53
4059 Bazilej, Švajčiarsko
swiss.ar@arazygroup.com

3. Vigilancia zdravotníckych pomôcok

Ak pri používaní zariadenia dôjde k vážnemu incidentu alebo ak je pravdepodobné, že k nemu dôjde, používatelia sú povinní bez zbytočného odkladu nahlásiť incident výrobcovi a príslušnému orgánu v krajine, kde k incidentu došlo alebo kde je pravdepodobné, že k nemu dôjde.

Príslušné nariadenia a zoznam príslušných orgánov v EÚ sú k dispozícii na tejto adrese:

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



4. Určené použitie platformy

4.1. Zamýšľané použitie

Platforma SeqOne je softvérové diagnostické zdravotnícke zariadenie typu *in vitro* určené na podporu zdravotníckych pracovníkov pri analýze a interpretácii surových genomických a sekvenčných údajov od pacientov podstupujúcich analýzu NGS alebo CGH array. Poskytovaním kvalitatívnych a využiteľných informácií funguje platforma ako systém na podporu rozhodovania, ktorý pomáha zdravotníckym pracovníkom pri klinickom rozhodovaní.

4.2. Indikácia pre použitie

Platforma SeqOne podporuje sekundárnu a terciárnu analýzu surových sekvenčných údajov, ako aj samotnú terciárnu analýzu. Jej analytické schopnosti nezávislé od patológie majú uplatnenie v rôznych medicínskych odboroch, vrátane onkológie a zriedkavých ochorení.

4.3. Popis

Platforma SeqOne je softvérové *in vitro* diagnostické zdravotnícke zariadenie, ktoré integruje súbor aplikácií pracujúcich spoločne, aby poskytli odborným používateľom výkonný nástroj na interpretáciu komplexných genomických a sekvenčných údajov.

Platforma SeqOne integruje overené informácie z širokej škály **externých zdrojov**, vrátane verejných databáz, vedeckej literatúry a proprietárnych znalostných báz. Tieto externé zdroje údajov poskytujú kontext a význam identifikovaným genomickým variantom.

Vstupné súbory pre rôzne aplikácie v rámci platformy sa spracúvajú pomocou prispôbenej kombinácie bioinformatických a anotačných modulov, čím sa zabezpečuje, že každá analýza je optimalizovaná pre svoje špecifické požiadavky.

Bioinformatické moduly vykonávajú sekundárnu analýzu pomocou výpočtových metód, ako je zarovnanie a identifikácia variantov, s cieľom identifikovať genetické varianty v údajoch sekvenovania DNA a RNA porovnaním s referenčným genómom (hg19 alebo hg38). Údaje, ktoré možno použiť na analýzu zárodočných buniek aj somatických buniek, môžu pochádzať zo sekvenovania celého genómu, sekvenovania exómu alebo cielených génových panelov.

Anotačné moduly vykonávajú terciárnu analýzu s cieľom anotovať varianty s podrobnosťami, ako je gén a jeho známa funkcia, umiestnenie variantu v géne a akékoľvek publikácie alebo klinické správy, ktoré spomínajú gén alebo variant. Táto anotácia určí, ktoré klasifikačné pravidlá ACMG a AMP/CAP/ASCO sa vzťahujú na každý variant, a na základe toho je možné dospieť k verdiktu o patogenite. Anotácia sa vykonáva aj prostredníctvom DiagAI, súboru nástrojov poháňaných umelou inteligenciou, ktoré pomáhajú analyzovať genetické varianty s cieľom rýchlo identifikovať potenciálne kauzálne varianty.

Táto modulárna architektúra umožňuje SeqOne vysokú prispôbivosť rôznym typom analýz a výskumným otázkam. Vstupné súbory pre rôzne aplikácie v rámci platformy sú spracovávané pomocou prispôbenej kombinácie týchto bioinformatických a anotačných modulov, čím sa zabezpečuje, že každá analýza je optimalizovaná pre svoje špecifické požiadavky. Integráciou týchto rôznorodých zdrojov údajov a analytických nástrojov umožňuje SeqOne odborným používateľom získať hlboký pohľad do komplexnej oblasti genomických údajov a ich vplyvu na ľudské zdravie a choroby.

Platforma funguje ako cloudové riešenie SaaS, pričom všetky funkcie sú dostupné prostredníctvom webového prehliadača. Zahŕňa SeqOne Flow, voliteľný nástroj, ktorý je možné nasadiť buď v cloude AWS, alebo na serveroch zákazníka, aby sa uľahčil prenos údajov a integrácia pracovných postupov.

Z bezpečnostných dôvodov sú všetky údaje v pokoji šifrované pomocou šifrovania AES-256 a prevádzka medzi používateľmi a platformou je zabezpečená pomocou certifikátov SSL.

4.4. Populácia pacientov

Riešenie je vhodné pre všetkých pacientov, ktorí potrebujú genomickú analýzu. Cieľovými pacientmi môžu byť pacienti s rakovinou, zriedkavými ochoreniami, dedičnými ochoreniami, genetickou predispozíciou na rakovinu alebo zdraví ľudia.

4.5. Určené pre používateľov

Platforma SeqOne je určená pre zdravotníckych pracovníkov zodpovedných za rutinnú klinickú genomovú analýzu. Zahŕňa analýzu a interpretáciu surových súborov generovaných sekvencerom zo sekvenovaných biologických vzoriek. Cieľovými používateľmi platformy sú genetici, bioinformatici, biológovia, inžinieri a technici platformy NGS.

4.6. Charakteristiky

Každá aplikácia (workset) pozostáva z viacerých **bioinformatických a anotačných modulov**, z ktorých každý zodpovedá za špecifické analytické úlohy. V tejto kapitole sa výkonnosť hodnotí na úrovni modulov, aby sa zabezpečilo podrobné posúdenie presnosti a spoľahlivosti každej zložky.

Analytické látky detekované SeqOne Platform spadajú do dvoch kategórií genómových zmien:

- Genetické varianty, ako sú SNV, indely, MNV, CNV, UPD, génové fúzie, štrukturálne varianty, šplhanie varianty, Alu inzercie a FLT3-ITD (Internal Tandem Duplication).
- Signatures alebo komplexné biomarkery, ako napríklad stav IGHV, nestabilita mikrosatelitových miest (MSI), Tumor Mutational Burden (TMB), metylácia alebo defekty opravy zlomenín dvojitej špirály prostredníctvom homologickej rekombinácie (HRD, Homologous Recombination Deficiency).

Modul bioinformatiky	Workset	Typ analytu
Detekcia CNV Exóm Germline	CNVExome, GermVar, GermVar Family	Genetická varianta
Detekcia 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 inzercie	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 somatický LF	SomaVar LF	Genetická varianta
CNV Somatické veľké panely	SomaCGP	Genetická varianta
Index metylácie	SomaMethyl	Signatúra
Fúzia somatickej DNA	SomaCGP, SomaVar, SomaVar LF, SomaHemato	Genetická varianta
Fúzia somatickej LBx DNA	SomaLBx	Genetická varianta
TMB	SomaCGP	Signatúra
MSI	SomaCGP, SomaMSI	Signature
Pravdepodobnosť GI (GiP)	SomaHRD	Signature
Sentieon SNV/Indel Germline	GermVar, GermVar Family	Genetická varianta
Mozaikové varianty Germline	GermVar, GermVar Family	Genetická varianta
CNV WGS	GermVar, GermVar Family	Genetická varianta
Sentieon SNV/Indel Somatic	SomaLBX	Genetická varianta
Fúzia RNA	SomaRNA	Genetická varianta
Splicingová RNA	SomaRNA	Genetická varianta
CNV Malý somatický panel	SomaCNVCapture	Genetická varianta
Stav IGHV	SomaHemato	Signatúra
Podskupiny IGHV (subsets)	SomaHemato	Signatúra
CFH	GermlineVarLR	Genetická varianta
FLT3	SomaVar, SomaHemato	Genetická varianta

4.6.1. Analytická výkonnosť

Výkonnosť uvedená v týchto pokynoch na použitie zodpovedá výkonnosti najnovších verzií worksetov pre in vitro diagnostiku, ktoré sú uvedené v tabuľke nižšie:

Názov worksetu	Verzia
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

- **Detekcia CNV Exómová germline a panelová germline**

Typ variantu	Verzia genómu	Pozitívne kontroly	Skutočné pozitívy	Falošné negatívy	Kritériá akceptácie	Citlivosť
Detekcia CNV Exóm Germline (GermlineVar)	GRCh37	8	8	0	100%	100%
Detekcia CNV Exóm Germline (GermVar)	GRCh37	8	8	0	100%	100%
Detekcia CNV Panel Germline	GRCh37	15	15	0	100%	100%
Detekcia CNV Panel Germline	GRCh38	11	11	0	100%	100%

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

Verzia genómu	Typ variantu	Kritériá akceptácie	Citlivosť pred filtrom	Citlivosť po filtrovaní
GRCh37	SNV	100	98 %	100 %
GRCh38	SNV	100 %	98 %	100 %

Verzia genómu	Typ variantu	Kritériá prijateľnosti	Presnosť pred filtrom	Presnosť po filtrovaní
GRCh37	SNV	100	90 %	100 %
GRCh38	SNV	100 %	90 %	100 %

- **SNV / indel Germline – údaje o exóme**

Verzia genómu	Type of variant	Kritériá akceptácie*	Citlivosť pred filtrom	Citlivosť po filtrovaní
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

* pre detekciu variantov nachádzajúcich sa v oblastiach s vysokou spoľahlivosťou NA12878 a NA24385.

Verzia genómu	Type of variant	Kritériá akceptácie*	Presnosť pred filtrom	Presnosť po filtrovaní
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

* pre detekciu variantov nachádzajúcich sa v oblastiach s vysokou spoľahlivosťou NA12878 a NA24385.

- **SNV / Indel Germline – Opakovateľnosť a reprodukovateľnosť**

- **Opakovateľnosť**

Type of variant	Citlivosť	Presnosť
SNV	98,59 % ± 0 %	95,92 % ± 0 %
indel	87,75 % ± 0 %	81,63 % ± 0 %
všetko	98,38 % ± 0 %	95,63 ± 0 %

- **Reprodukovateľnosť**

Type variantu	Citlivosť	Presnosť
SNV	98,57 % ± 0,03 %	95,93 % ± 0,08 %
indel	86,58 % ± 0,42 %	78,97 % ± 0,96 %
všetko	98,34 % ± 0,03 %	95,58 ± 0,09 %

- **Mozaikové varianty Germline**

- **Citlivosť SNV**

Verzia genómu	VAF biny	Počet skutočných variantov	Kritériá akceptácie	Citlivosť 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 %

- **Citlivosť Indels**

Verzia genómu	VAF biny	Počet skutočných variantov	Kritériá akceptácie	Citlivosť 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 variantu	Verzia genómu	Pozitívne kontroly	Skutočné pozitívne	Falošné negatívy	Kritériá prijateľnosti	Citlivosť
SV Germline	GRCh37	2	2	0	100 %	100 %

- Vloženia Alu

Verzia genómu	Pozitívne kontroly	Skutočné pozitívne výsledky	Falošné negatívy	Kritériá akceptácie	Citlivosť
GRCh37	1	1	0	100 %	100 %
GRCh38	1	1	0	100 %	100 %

- MNV

Verzia genómu	Type of variant	Kritériá akceptácie	Presnosť exómu	Presnosť genómu
GRCh37	MNV	> 85 %	87,0 %	86,9 %
GRCh38	MNV	> 85 %	93,0 %	91,5 %

- SNV / Indel Somatický

- o Citlivosť

Verzia genómu	Type of variant	Pozitívne kontroly	Skutočné pozitívny	Falošné negatívy	Kritériá akceptácie*	Citlivosť
GRCh37	SNV ($\geq 5\%$)	121	121	0	100 %	100
	Indel ($\geq 5\%$)	21	21	0	100 %	100 %
GRCh38	SNV ($\geq 5\%$)	121	121	0	100 %	100
	Indel ($\geq 5\%$)	21	21	0	100 %	100

*Kritérium akceptácie pre detekciu variantov je definované ako 100 % citlivosť pre SNV a indely s očakávanou VAF nad 5 %.

o Správnosť

			Správnosť			
			GRCh37		GRCh38	
Gene	Variant	Očakávaná frekvencia	Priemerná frekvencia*	Odchýlka	Priemerná frekvencia*	Odchýlka
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

*Priemerná hodnota frekvencií bola získaná na základe 12 opakovaných meraní vzorky HD200.

• SNV/Indel Somatic LF

Genóm verzia	Pozitívne kontroly	Skutočne pozitívne	Falošné negatívne	Kritériá prijateľnosti*	Citlivosť
GRCh37	42	42	0	100 %	100 %
GRCh38	42	42	0	100 %	100 %

*Kritérium akceptácie pre detekciu variantov je definované ako 100 % citlivosť pre SNV a indely s očakávanou hodnotou VAF nad 0,1 %.

• Veľké panely somatických CNV

Verzia genómu	Type of variant	Pozitívne kontroly	Skutočné pozitívny	Falošné negatívny	Kritériá akceptácie	Citlivosť
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 metylácie

Verzia genómu	Kritériá akceptácie	Vážená zhoda
GRCh37	>99 %	99,88

- Fúzia somatickej DNA

Genóm verzia	Pozitívne kontroly	Skutočné pozitívne	Falošné negatívne	Prijatie Kritériá	Citlivosť
GRCh37	9	9	0	100 %	100 %
GRCh38	9	9	0	100 %	100 %

- Somatic LBx DNA fúzia

Verzia genómu	Podiel nádoru	Pozitívne kontroly	Skutočné pozitívne	Falošné negatívny	Kritériá akceptácie	Citlivosť
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

Verzia genómu	Kritériá prijatia	Citlivosť	Presnosť
GRCh37	>80 %	83 %	85 %

- MSI

Verzia genómu	Pozitívne kontroly	Negatívne kontroly	Skutočné pozitívne	Falošné negatívny	Kritériá prijateľnosti	Citlivosť	Presnosť
GRCh37	22	24	22	0	100 %	100 %	100 %
GRCh38	6	17	6	0	100 %	100 %	100 %

- GiP

Genóm verzia	Pozitívne kontroly	Skutočne pozitívne	Falošné negatívne	Prijatie Kritériá	Citlivosť
GRCh37	171	165	6	90 %	96,5 %

Genóm verzia	Pozitívne kontroly	Skutočne pozitívne	Falošné pozitívne	Prijatie Kritériá	Presnosť
GRCh37	171	165	7	90 %	95,9 %

- Sentieon SNV / Indel Germline

Verzia genómu	Type of variant	Kritériá akceptácie *	Citlivosť Exome.	Citlivosť Genóm
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

* pre detekciu variantov nachádzajúcich sa v oblastiach s vysokou spoľahlivosťou NA12878 a NA24385

Verzia genómu	Type of variant	Kritériá akceptácie *	Presnosť Exome.	Presnosť genómu
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 %

* pre detekciu variantov nachádzajúcich sa v oblastiach s vysokou spoľahlivosťou NA12878 a NA24385.

- CNV WGS

Verzia genómu	Pozitívne kontroly	Skutočné pozitívny	Falošné pozitívny	Kritériá akceptácie	Citlivosť
GRCh38	Duplikácia	1 kb – 5 kb	8	10	37,5
	Duplikácia	≥ 5 kb	3	10	100
	Delécia	1 kb – 5 kb	321	80	61,7 %
	Delécia	≥ 5 kb	132	80	92,4

Hoci citlivosť na malé delécie (1–5 kb) nedosahuje akceptovateľné kritériá, toto zníženie je prijateľným kompromisom s cieľom maximalizovať detekciu významnejších veľkých udalostí a duplikácií.

Verzia genómu	Pozitívne kontroly	Skutočné pozitív	Falošné pozitívny	Kritériá akceptácie	Presnosť
GRCh38	Duplikácia	1 kb – 5 kb	8	"NA"	30
	Duplikácia	≥ 5 kb	3	"NA"	75 %
	Delécia	1 kb – 5 kb	321	"NA"	94,3 %
	Delécia	≥ 5 kb	132	"NA"	92,4 %

- **Sentieon SNV/Indel Somatický**

Verzia genómu	Typ variantu	Pozitívne kontroly	Skutočné pozitívy	Falošné pozitívy	Kritériá akceptácie	Citlivosť
GRCh38	SNV	39	37	2	85 %	95 %
	Indel	21	16	5	75 %	76 %

- **Fúzna RNA**

Genóm Verzia	Pozitívne kontroly	Skutočne falošné	Falošné negatívne	Prijatie Kritériá	Citlivosť
GRCh37	20	20	0	100 %	100 %
GRCh38	20	19	1	95 %	95 %

- **Splicing RNA**

Genóm Verzia	Pozitívne kontroly	Skutočne falošné	Falošné negatívne	Prijatie Kritériá	Citlivosť
GRCh37	3	3	0	100 %	100 %
GRCh38	3	3	0	100 %	100 %

- **CNV Malý panel somatických**

Genóm Verzia	Pozitívne kontroly	Skutočne falošné	Falošné negatívne	Prijatie Kritériá	Citlivosť
GRCh37	4	4	0	100 %	100 %
GRCh38	4	4	0	100 %	100 %

- **Status IGHV IMGT**

Genóm verzia	Pozitívne kontroly	Skutočne pozitívne	Falošné negatívne	Prijatie Kritériá	Citlivosť
GRCh37	122	116	6	95 %	95,9 %

- **Status IGHV igblast**

Verzia genómu	Pozitívne kontroly	Skutočné pozitívne	Falošné negatívny	Kritériá akceptácie	Citlivosť
GRCh37	122	117	5	95 %	95,1 %

- **Podskupiny IGHV**

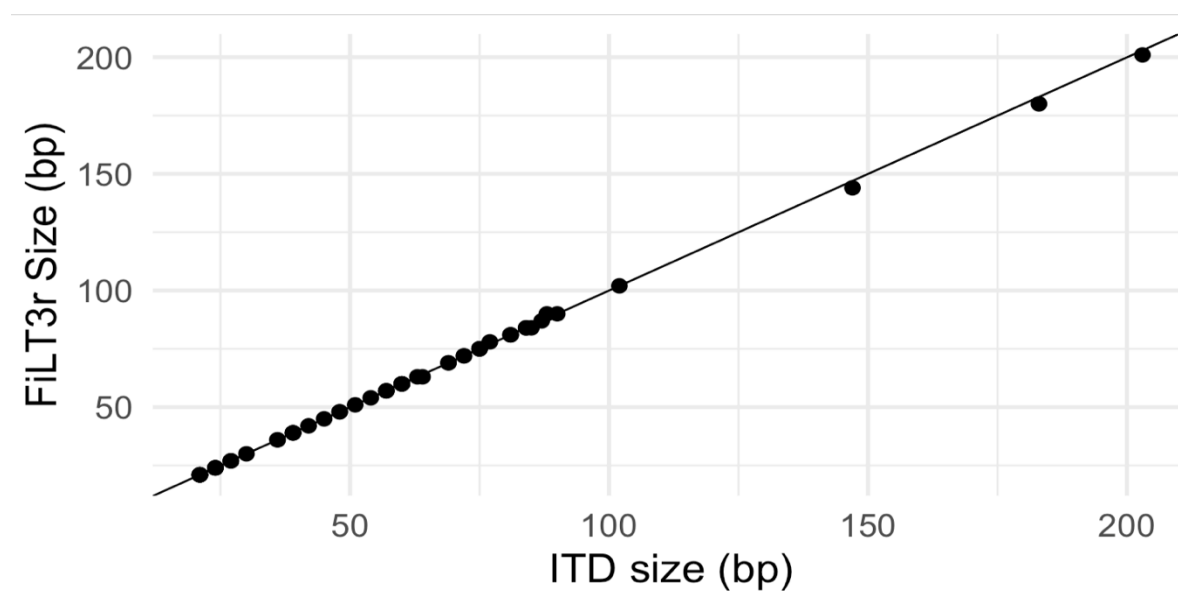
Genóm verzia	Pozitívne kontroly	Skutočne pozitívne	Falošné negatívne	Prijatie Kritériá	Citlivosť
GRCh37	61	10	0	95 %	100 %

Genóm verzia	Pozitívne kontroly	Skutočne pozitívne	Falošné negatívne	Prijatie Kritériá	Presnosť
GRCh37	61	10	0	95 %	100 %

- CFH

Genóm verzia	Pozitívne kontroly	Skutočne falošné	Falošné negatívne	Prijatie Kritériá	Citlivosť
GRCh38	4	4	0	100 %	100 %

- FLT3



Pozorovaná vs. teoretická veľkosť udalosti FLT3-ITD. **Pearsonov koeficient $R^2 > 0,99$**

Výsledky ukazujú 100 % citlivosť pri detekcii udalostí FLT3-ITD a takmer dokonalú koreláciu medzi očakávanými a pozorovanými veľkosťami udalostí z analýzy.

- UPD

- Citlivosť

Type of analysis	Verzia genómu	Pozitívne kontroly	Negatívne kontroly	Skutočné pozitívne výsledky	Falošné negatívy	Kritériá akceptácie	Citlivosť
Singleton	GRCh37/38	7	23	7	0	80 %	100 %
Trio	GRCh37/38	7	2	6	1	80 %	86 %

- Presnosť

Type of analysis	Verzia genómu	Pozitívne kontroly	Negatívne kontroly	Skutočné pozitívne výsledky	Falošné pozitívy	Kritériá akceptácie	Presnosť
Singleton	GRCh37/38	7	23	7	0	80 %	100 %
Trio	GRCh37/38	7	2	6	0	80 %	100 %

4.6.2.1.1. Moduly anotácie

- Skóre DiagAI

Súbor údajov	Analýza typu	Metrické vyhodnotený	Kritériá Kritériá	Metrika hodnota
Exómy pozitívne na zriedkavé ochorenia	'Germline'	Stredná hodnota	= 1	1
Exome pozitívne na zriedkavé ochorenia	'Germline'	top_10	> 90 %	95 %
Exome CNV vzácných ochorení	"Germline"	Stredná hodnota	= 1	1
Exómy s pozitívnym nálezom CNV pri zriedkavých ochoreniach	'Germline'	top_1	> 80 %	83,3

- Somatické skóre DiagAI

Súbor údajov	Analýza typu	Metrika vyhodnotené	Kritériá prijateľnosti	Hodnota metriky
Rakovina pozitívna TS0500	"Somatic"	top_10	> 90 %	100 %

- **DiagAI užší výber**

Dataset	Analýza typu	Metrika hodnotená	Prijateľnosť Kritériá	Hodnota metriky
Exómy pozitívne na zriedkavé ochorenia	'Germline'	citlivosť	> 90	95
Exome pozitívne na zriedkavé ochorenia	"Germline"	Priemerná dĺžka	< 25	18,475
Exome s pozitívnym nálezom zriedkavých ochorení	"Germline"	Stredná dĺžka	< 25	17

- **DiagAI SmartPicks**

Súbor údajov	Analýza typu	Metrika hodnotená	Prijateľnosť Kritériá	Hodnota metriky
Exómy pozitívne na zriedkavé ochorenia	"Germline"	citlivosť	> 40 %	45
Exome pozitívne na zriedkavé ochorenia	"Germline"	presnosť	> 80 %	78 %

- **ACMG CNV**

Súbor údajov	Analýza typu	Metrika vyhodnotená	Kritériá prijateľnosti	Hodnota metriky
Súbor údajov CNV s poznámkami odborníkov ACMG	"Germline"	citlivosť	> 70 %	76

- **ACMG SNV**

Súbor údajov	Analýza typu	Metrika vyhodnotená	Kritériá prijateľnosti	Hodnota metriky
Datový súbor s poznámkami odborníkov ACMG	"Germline"	presnosť	> 50 %	56 %

- **DiagAI UP2**

Súbor údajov	Analýza typu	Metrika vyhodnotená	Kritériá prijateľnosti	Metrika hodnota
Varianty spôsobujúce ochorenia	"Germline"	citlivosť	90 %	92

- **DiagAI UP2 SV**

Súbor údajov	Analýza typu	Metrika vyhodnotená	Kritériá prijateľnosti	Metrika hodnota
Varianty spôsobujúce ochorenie	"Germline"	citlivosť	90 %	98

- **DiagAI Phenogenius**

Súbor údajov	Analýza typu	Metrika vyhodnotené	Kritériá prijateľnosti	Metrika hodnota
Asociácie génov a symptómov	'Germline'	top_100	50 %	49,8
Súvislosti medzi génmi a príznakmi	"Germline"	top_200	60 %	60,7
Súvislosti medzi génmi a príznakmi	"Germline"	top_500	75 %	77,0 %
Súvislosti medzi génmi a príznakmi	"Germline"	Stredná hodnota v top 100	<20	15
Súvislosti medzi génmi a príznakmi	"Germline"	Priemer v top 100	<30	25

- **Compermed**

Súbor údajov	Analýza typu	Metrika vyhodnotený	Kritériá prijateľnosti	Metrika hodnota
Rakovina pozitívna TSO500	'Somatic'	citlivosť	90 %	91

4.6.3. Klinická výkonnosť

SeqOne Platform nie je určená na posudzovanie konkrétneho klinického stavu alebo fyziologického stavu. Hoci poskytuje informácie o fyziologickom stave pacienta, ako je identifikácia určitých variantov, nie je určená na posudzovanie konkrétnych klinických stavov alebo fyziologických stavov. SeqOne Platform preto sama o sebe nevyžaduje preukázanie klinickej výkonnosti.

5. Bezpečnostné opatrenia

5.1. Správne používanie

(Pozri časť **Určení používateľa**)

Akéhokoľvek nevhodné používanie platformy je zakázané.

Používatelia sú zodpovední za kontrolu, či je platforma vždy funkčná, a za zabezpečenie podmienok potrebných na jej správne fungovanie.

5.2. Upozornenia

⚠ Poskytnuté výsledky by nemali byť jediným základom pre zmenu starostlivosti o pacienta. Používatelia musia preto pri používaní softvéru a jeho výsledkov postupovať s náležitou opatrnosťou. Diagnózy by mali byť potvrdené prostredníctvom rôznych klinických a lekárskeho vyšetrení.

Vyšetrovanie ochorení pomocou NGS môže odhaliť sekundárne alebo náhodné nálezy, ako je predispozícia na rakovinu, z génov, ktoré nesúvisia s primárnym ochorením (Kuhlen a Borkhardt 2015). Napriek politickým vyhláseniam agentúr ako AMP a ACMG (Green et al. 2013; Allyse a Michie 2013; Predstavenstvo ACMG 2015; Hegde et al. 2015) sa národné smernice týkajúce sa hlásenia náhodných zistení líšia. Preto je kľúčové zabezpečiť dokument „informovaného súhlasu“ (Anderson et al. 2016). Ošetrojúci lekár je zodpovedný za to, aby pacient alebo jeho zákonný zástupca pochopil potenciálne výsledky a riziká pred objednaním akéhokoľvek testu NGS.

Tento produkt nezaručuje detekciu všetkých genetických variantov prítomných vo vzorke. Pravdepodobnosť detekcie hľadaného variantu je vo svojej podstate závislá od kvality a charakteristík vstupnej vzorky (napr. integrita vzorky, príprava knižnice, hĺbka a uniformita sekvenovania, bioinformatické filtrovanie a typ alebo kontext variantu). Negatívny výsledok definitívne nevylučuje prítomnosť patogénnej varianty, ak je jej frekvencia (VAF) nižšia ako detekčné prahy použitého analytického modulu. Spoľahlivá detekcia rozsiahlych genómových prestavieb (štrukturálnych variantov, vrátane CNV a delécií celých chromozómov) nie je zaručená. Je to spôsobené vnútorným obmedzením údajov sekvenovania novej generácie (NGS) s krátkymi čítaniami, čo sťažuje identifikáciu týchto udalostí. Keďže čítania NGS sú krátke (približne 150–200 bp), toto obmedzenie je charakteristické pre samotnú sekvenčnú technológiu a nie je obmedzením SeqOne Platform. (Pozri Tan R. et al., Hum Mutat. 2014 Jul;35(7):899–907. doi: 10.1002/humu.22537 a odkazy v ňom). Používatelia musia vyhodnotiť vhodnosť zvoleného worksetu a kvalitu sekvenovania vo vzťahu k očakávaným variantom, aby sa uistili, že bol pokrytý príslušný klinický kontext.

Klinické rozhodnutia založené na výsledkoch skríningu novorodencov alebo Carrier screening musia byť prijímané s mimoriadnou opatrnosťou a nikdy sa nesmú spoliehať výlučne na výstupy tohto softvéru. V prípade skríningu novorodencov si pozitívny nález vyžaduje okamžitú klinickú koreláciu so symptómami dieťaťa a urgentné potvrdzujúce testovanie pred začatím liečby ovplyvňujúcej život alebo diétnych obmedzení. V prípade carrier screeningu negatívny výsledok nevylučuje riziko, že daná osoba je nositeľom, ale skôr ho znižuje na úroveň reziduálneho rizika, keďže určité genómové varianty môžu klesnúť pod detekčný limit softvéru. Keďže výsledky z tohto zariadenia priamo ovplyvňujú nezvratné reprodukčné rozhodnutia a kritickú neonatálnu starostlivosť, všetky nálezy, na základe ktorých je možné konať, musia byť overené pomocou alternatívnej validovanej metódy a interpretované kvalifikovaným odborníkom v kontexte úplnej klinickej anamnézy pacienta a relevantných biochemických alebo fenotypových údajov.

Novozistené varianty, ktoré ešte nie sú dôkladne zdokumentované vo vedeckej literatúre, môžu byť nesprávne klasifikované.

Výsledky sú presné v čase analýzy a odrážajú súčasné vedecké poznatky. Opätovné spustenie analýzy na novej verzii platformy však môže priniesť odlišné výsledky.

Výsledky klasifikácie z SeqOne Platform sa snažia zosúladiť so všeobecným konsenzom. Uvedomte si, že akékoľvek výsledky vygenerované po zmene nastavení klasifikácie sú mimo rozsahu CE-IVD a preto nie sú validované.

Tento softvér obsahuje komponenty umelej inteligencie (AI). Používatelia si musia byť vedomí nasledujúcich obmedzení, ktoré sú vlastné týmto nástrojom založeným na AI, a zohľadniť ich pri interpretácii výsledkov:

- Trénovacie údaje pochádzajú prevažne z európskych/kaukazských populácií predkov vzhľadom na zloženie súčasných klinických referenčných databáz;
- Výkonnosť sa môže líšiť v prípade zriedkavých variantov s obmedzeným zastúpením v trénovacích údajoch;
- Skóre spoľahlivosti odrážajú neistotu modelu, nie klinickú istotu; musia sa interpretovať v spojitosti s klinickým kontextom;
- Výkonnosť sa môže zhoršiť v dôsledku posunu údajov z nových sekvenčných platforiem alebo protokolov prípravy knižníc, ktoré nie sú zastúpené v trénovacích údajoch;
- Výkonnosť nástroja DiagAI Phenogenius pri prioritizácii génov sa hodnotí na základe asociácií génov a symptómov a závisí od kvality a úplnosti anotácie termínov HPO zo strany objednáväjúceho klinického lekára.

5.3. Obmedzenia

Carrier screening for genetic conditions is subject to specific national laws and ethical regulations. The laboratory and the healthcare provider are responsible for ensuring that all legal requirements –

including mandatory genetic counseling, physician-led ordering, and informed consent – are met within their specific jurisdiction.

5.4. Poruchy

V prípade poruchy:

- Okamžite prestaňte používať softvér.
- Skúste identifikovať alebo odstrániť príčiny podľa postupu riešenia problémov opísaného v tejto príručke (pozri časť **Riešenie problémov a chybové správy**).
- Ak pomocou tohto návodu nedokázate identifikovať alebo odstrániť príčinu, pošlite e-mail tímu podpory na support@seqone.com.

6. Minimálne požiadavky

6.1. Hardvér

Softvérové moduly platformy SeqOne sú inštalované ako súčasť klientsko-serverového usporiadania. Poskytovateľ služieb musí byť serverom Linux, aby bola zabezpečená kompatibilita s väčšinou nástrojov bežne používaných v bioinformatike. Každá systémová položka je zapuzdrená v prostredí runtime, ako je kontajner Docker®, ktorý možno nasadiť na akejkolvek kompatibilnej platforme Docker®, bez ohľadu na základnú hardvérovú architektúru. Používatelia musia mať počítač s pripojením na internet a webovým prehliadačom (pozri nižšie).

Minimálne systémové požiadavky pre klientsky počítač sú nasledovné:

- 4 GB RAM
- Dvojjadrový procesor
- Prístup k internetu
- Prehliadač Chrome (verzia 36.0.1985 a novšie) alebo Firefox (verzia 31 a novšie) alebo Microsoft Edge (všetky verzie založené na Chromium).

6.2. Sieť

Minimálna odporúčaná rýchlosť odosielania/sťahovania pre používanie platformy cez internet je 100 Mbps. Na pripojenie k platforme sa používa protokol HTTPS. Aby ste mali správny prístup k platforme, odporúčame požiadať vaše IT oddelenie o autorizáciu nasledujúcich adries:

*.eu.seqone.com

*.common.seqone.com

7. Overenie

7.1. Prihlásenie

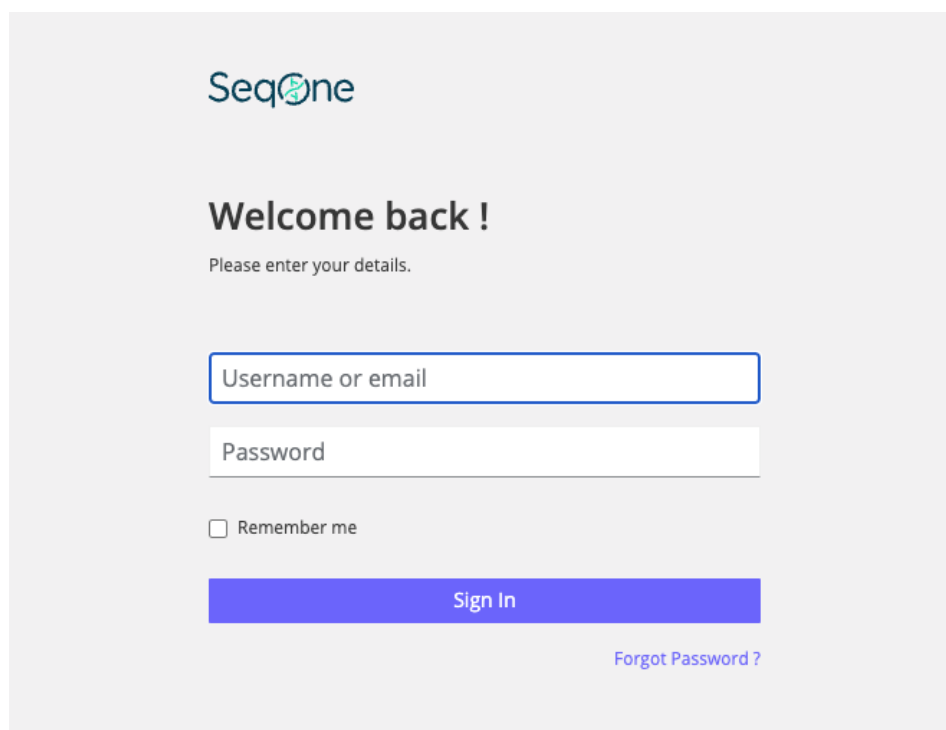
Z bezpečnostných dôvodov majú používatelia prístup k platforme až po prihlásení sa na domovskej obrazovke.

Domovská obrazovka je dostupná na nasledujúcej adrese URL:

- Pre Európu a Afriku: <https://auth.eu.seqone.com/>
- Pre Švajčiarsko: <https://auth.ch.seqone.com/>
- Pre Áziu a Oceániu: <https://auth.au.seqone.com/>
- Pre Ameriku: <https://auth.us.seqone.com/>

Pri prihlasovaní na domovskej obrazovke musíte zadať:

- Prihlasovacie meno alebo e-mailovú adresu do poľa „Username“ (Prihlasovacie meno) alebo „email“ (E-mail)
- Heslo do poľa „Heslo“



Obrázok 1: Prihlasovacia stránka

Ak je pre váš účet aktivovaná dvojfaktorová autentifikácia, budete musieť zadať 6-miestny overovací kód zaslaný e-mailom na adresu spojenú s vaším účtom SeqOne. Tento kód je platný 30 minút.

Ak máte prístup k viacerým pracovným priestorom alebo entitám, budete požiadaní, aby ste vybrali entitu, ku ktorej sa chcete pripojiť, prostredníctvom roletového menu v časti „Choose a workspace“.

Heslo môžete zmeniť kliknutím na „Forgot Password“. Postupujte podľa pokynov a zadajte e-mailovú adresu spojenú s vaším účtom SeqOne. Následne dostanete e-mail, ktorý vám umožní resetovať heslo.

Zásady pre heslá sa riadia nasledujúcimi pravidlami:

- Minimálna dĺžka: 12 znakov.
- Zložitosť: heslá musia obsahovať aspoň jedno veľké písmeno, jedno malé písmeno, jednu číslicu a jeden špeciálny znak. Čím väčšia je sada znakov, tým silnejšie je heslo.
- História hesiel: nesmie sa opakovať použitie predchádzajúcich hesiel.
- Žiadna maximálna dĺžka: aby bolo možné používať tajné frázy.

Užívatelia nemôžu zmeniť svoje vlastné užívateľské meno.

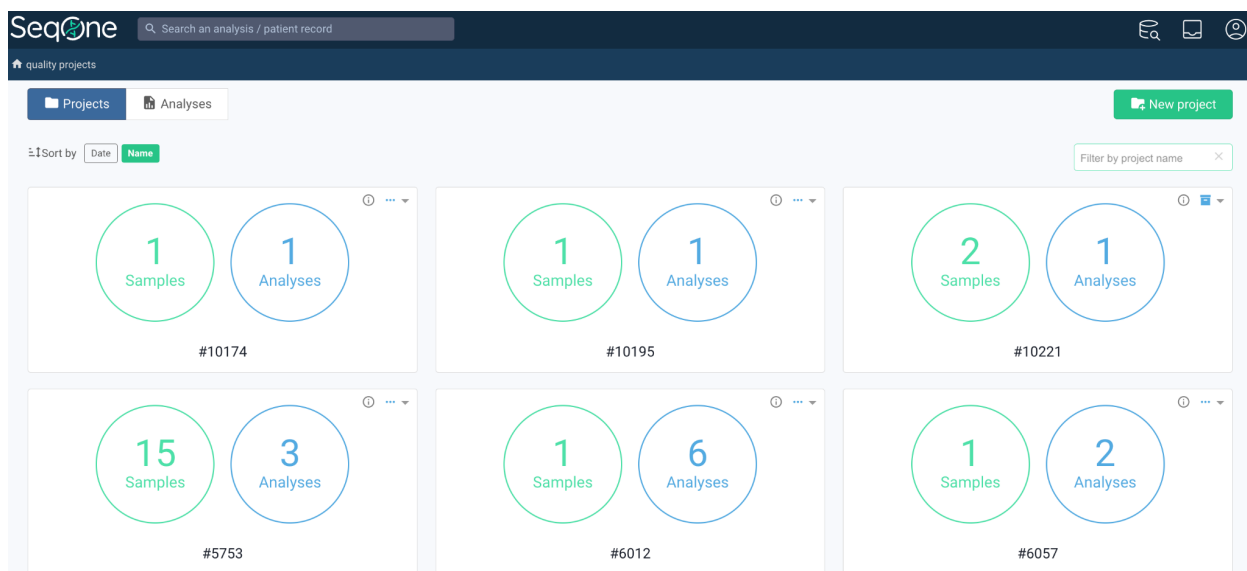
7.2. Odhlásenie

Používatelia sa môžu odhlásiť kliknutím na logo používateľského účtu v pravom hornom rohu stránky a potom na „Odhlásiť sa“.

7.3. Domovská stránka

Po overení sa používateľ dostane na stránku riadiaceho panela, ktorá je rozdelená na dve časti:

- modrý vyhľadávací a akčný banner, dostupný počas celej navigácie, obsahujúci zľava doprava:
 - ikonu SeqOne pre návrat na hlavnú obrazovku kedykoľvek,
 - vyhľadávací riadok pre priamy prístup k danej analýze,
 - ikonu vyhľadávania VKB  na vyhľadávanie variantov vyhodnotených v databáze VKB (pozri časť **Vyhodnocovanie variantov**),
 - ikonu  , ktorá otvára „zoznam úloh“ používateľa, z ktorého má prístup k analýzam, ktoré mu boli pridelené (pozri časť **Pridelenie analýz na interpretáciu**),
 - paneli ponuky, dostupnom cez ikonu  , ktorý poskytuje prístup k parametrom používateľa, panelu správcu, používateľskej príručke, podmienkam používania, e-mailovej adrese podpory a ktorý umožňuje používateľovi odhlásiť sa (pozri časť **Správa používateľského účtu**).
- hlavná obrazovka na správu projektov a analýz.



Obrázok 2: Používateľský panel

K používateľskému paneli sa môžete kedykoľvek dostať kliknutím na logo  .

7.4. Správa používateľského účtu

Svoje používateľské konto môžete spravovať kliknutím na logo v pravom hornom rohu panela úloh.

V tejto časti má používateľ možnosť:

- prístup k nastaveniam účtu kliknutím na "My account"
- prístup k používateľským parametrom kliknutím na "Entity Settings"
- zmena entity, ku ktorej je používateľ pripojený, kliknutím na "Entity" a výberom požadovanej entity v časti "Choose your workspace",
- prístup k panelu administrátora kliknutím na "Admin", ktorý umožňuje prístup k zoznamu používateľov s prístupom k entite a vytváranie, úpravu alebo mazanie používateľských účtov.
- kliknutím na "About" získate prístup k používateľskej príručke v rôznych jazykoch, informáciám o verzii softvéru SeqOne a kontaktným údajom,
- prístup k podmienkam vašej licencie SeqOne kliknutím na "Terms of Use",
- kliknutím na "Support" získate prístup k e-mailovej adrese podpory pre akékoľvek otázky alebo problémy zistené na platforme,
- prístup ku skratke kliknutím na "Keyboard Shortcuts"

- prepínanie medzi štandardnou verziou a verziou preview aplikácie kliknutím na "SeqOne Mode" a výberom buď "SeqOne" (standard mode) alebo "SeqOne Preview" (preview mode). Aktuálne aktívny režim je označený zaškrtnutím.
- odhlásenie kliknutím na "Logout".

7.4.1. Správa "My Account"

V sekcii "My account" môže používateľ:

- nahrať svoj podpis na jeho zahrnutie do klinickej správy,
- zmeniť svoje heslo kliknutím na "Change your password" (otvorí sa na novej karte) v sekcii "Manage your credentials".

7.4.2. Správa nastavení používateľa – definície

V používateľských nastaveniach môže používateľ spravovať nastavenia, súbory a konfigurácie popísané nižšie.

Entita

Entita sa vzťahuje na zariadenie alebo tím, ktorý si zakúpil licenciu na platformu SeqOne. Platforma je navrhnutá tak, aby ju mohlo používať viacero členov v rámci tej istej skupiny. V nastaveniach používateľa sa karta *Účet* používa na správu informácií týkajúcich sa entity.

Users

Pri tejto karte sa prosím obráťte na **Správa administrátora – rozhranie Prezentácia**.

Worksety

Táto karta obsahuje zoznam rôznych worksetov analýzy, ktoré sú k dispozícii v entite, ku ktorej je používateľ pripojený. Keďže na platforme SeqOne môže existovať viacero verzií určitých pipeline, používateľ môže v tomto menu definovať, ktorá verzia by mala byť predvolená pri spustení analýzy.

Zmeny vykonané v tomto zozname sa sledujú, takže sa automaticky zobrazí používateľ, ktorý definoval verziu, ktorá sa má spustiť ako predvolená verzia pracovnej sady, ako aj dátum zmeny.

Manifest

Manifest zobrazuje genómové oblasti záujmu (ROI) ako súbor BED, ktorý môže byť doplnený o príslušné anotácie. Nový manifest môžete nahrať na karte *Manifest* v nastaveniach používateľa kliknutím na „Pridať manifest“ a výberom súboru manifestu, ktorý chcete nahrať. Po nahratí bude tento manifest k dispozícii všetkým používateľom pri vytváraní nového projektu a pri použití in silico panelu v analýze.

Panels - Transkript

Používatelia môžu nahrať nový zoznam referenčných transkriptov (RefSeq) cez kartu *Transkript*.

Panels - Gény

Táto karta vám umožňuje importovať alebo generovať genové panely na základe NCBI GeneID kliknutím na „Pridať zoznam génov“.

Očakávaný formát súboru je tabuľkový súbor (napr. súbor .tsv) bez záhlaví stĺpcov, čo je podrobne uvedené na platforme. Prvý stĺpec očakávaného súboru je vyhradený pre identifikátory génov (NCBI Entrez gene ID), druhý stĺpec (voliteľný) obsahuje transkript záujmu (RefSeq ID) a tretí stĺpec (voliteľný) obsahuje symbol HGNC génu. Tento súbor je možné vytvoriť aj vložením požadovaného zoznamu génov, oddelených čiarkou, do poľa „Gene symbols“ (Symboly génov).

Marker

Táto záložka slúži na import zoznamu genetických identifikačných markerov kliknutím na „Add marker config“ (Pridať konfiguráciu markerov). Očakávaný formát súboru je tabuľkový súbor (.tsv alebo .txt) s hlavičkou stĺpca obsahujúcou 5 stĺpcov s podrobnými informáciami: chromozóm (CHR), identifikátor variantu alebo rsID (ID), genomická pozícia (POS), referenčná alela (REF) a alternatívna alela (ALT). Súbor nesmie obsahovať viac ako 100 riadkov vrátane záhlavia.

Protocols

Analýzny protokol je vopred definovaná kombinácia úloh navrhnutá na štandardizáciu interpretácie analýzy. Príslušná karta umožňuje používateľom zobrazíť zoznam protokolov dostupných v entite, ich verzie a vytvárať ich.

Ak chcete vytvoriť protokol, kliknite na "New protocol", potom pridajte a zordťte úlohy, ktoré zodpovedajú SOP vášho laboratória. Ak chcete upraviť existujúci protokol, kliknite na ikonu ceruzky vedľa neho a uložte novú verziu protokolu.

Vyberte protokol pri spustení analýzy alebo ho pripojte neskôr z karty Workflow v paneli analýzy. Označte úlohy ako dokončené priamo v paneli analýzy. Každá dokončená úloha automaticky zaznamená používateľa a dátum pre úplnú sledovateľnosť.

Reporting - Global settings

Sekcia Global settings na karte Reporting umožňuje používateľom povoliť alebo zakázať automatické uzamknutie analýzy, keď sa stav zmení na "Reporting Validating", a prispôbiť sekcie klasifikácie a reportovania. Ak to chcete urobiť, v časti Germline / Somatic kliknite na "New section", zadajte názov sekcie a vyberte tag, ktorý chcete priradiť k požadovanej klasifikácii. Varianty, ktoré zodpovedajú jednému z vybraných tagov alebo klasifikácií, sa do tejto sekcie pridávajú automaticky. Ak chcete túto funkciu používať, musíte najprve vytvoriť tagy na príslušnej karte.

Reporting - Template list

Karta Reporting pod "Template list" umožňuje používateľom konfigurovať šablóny používané pri exporte výsledkov analýzy. Na tejto karte môžete prispôsobiť informácie obsiahnuté v generovaných reportoch.

Tags

Karta "Tags" umožňuje používateľom spravovať vlastné štítky na organizovanie a filtrovanie variantov v rámci VKB a reportu. Používatelia môžu vytvárať tagy kliknutím na "Add tag", alebo premenovávať a mazať tagy, ktoré sú dostupné všetkým používateľom v rámci entity, pomocou zodpovedajúcich ikon pera a koša.

API Keys

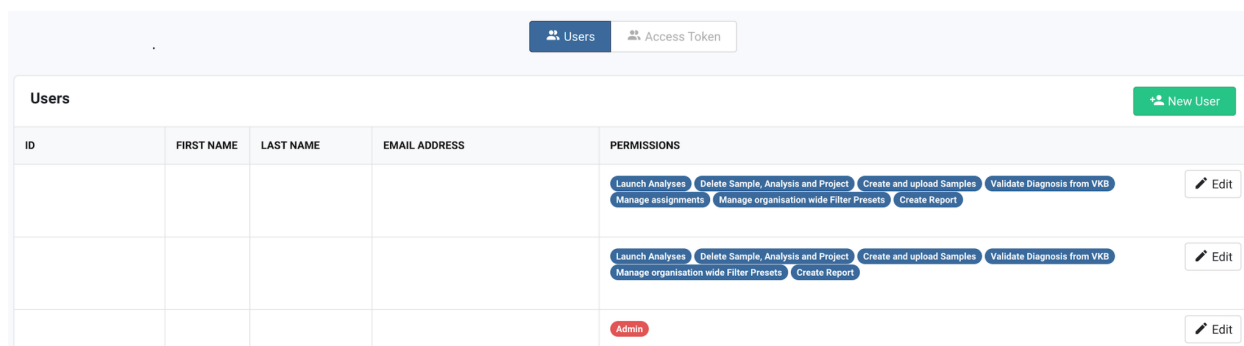
Kľúč API je jedinečný autentifikačný token, ktorý umožňuje externým aplikáciám alebo skriptom programovo komunikovať s platformou SeqOne. Na karte API Keys si môžete zobrazíť všetky existujúce kľúče spolu s prislúchajúcimi informáciami.

Nový kľúč vytvoríte kliknutím na "+ New API Key". Expirované kľúče môžete zobrazíť alebo skryť pomocou prepínača "Show expired key". Každý kľúč je možné odstrániť jednotlivo. Kľúče API sa musia uchovávať ako dôverné, pretože udeľujú prístup k platforme s prislúchajúcimi oprávneniami.

7.4.3. Správa administrátora – rozhranie Prezentácia

Toto rozhranie je určené správcom účtov a umožňuje im:

- prístup k zoznamu používateľov v organizácii,
- vytvárať nových používateľov,
- upravovať práva používateľov: vytvárať projekty a nahrávať vzorky, spúšťať analýzy, mazať vzorky, analýzy alebo projekty, hodnotiť varianty vo VKB, vytvárať správy, spravovať profily filtrov, spravovať úlohy.
- aktivovať dvojité overovanie,
- deaktivovať používateľský 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

Obrázok 3: Panel správcu

ID zodpovedá používateľskému menu používateľa. Musí byť v malých písmenách a nesmie obsahovať diakritiku ani medzery; špeciálne znaky, ako bodka a podčiarknutie, sú povolené. Pri vytváraní nového používateľa sa odporúča dodržiavať nomenklatúru meno.priezvisko.

Pokiaľ ide o e-mailovú adresu používateľa, z bezpečnostných dôvodov dôrazne neodporúčame používať generickú e-mailovú adresu.

Keďže povolenia sú štandardne deaktivované, ich aktivácia je na správcovi:

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

Obrázok 4: Oprávnenia používateľa

Bezpečnosť môžete zvýšiť aj aktivovaním dvojitého overovania. Táto možnosť odošle používateľovi jedinečný kód zakaždým, keď sa pripojí k platforme.

Úloha „Spravovať prednastavenia filtrov v rámci organizácie“ umožňuje používateľovi upravovať verejné profily filtrov. Ak nemáte túto úlohu, môžete upravovať len súkromné profily filtrov.

Oprávnenia spravuje administrátor entity z panelu Admin. Pri vytvorení používateľa sú všetky oprávnenia v predvolenom nastavení zakázané a administrátor musí manuálne povoliť práva vyžadované pre daného používateľa. Práva pridelené používateľovi sa vzťahujú na všetky entity (pracovné priestory), ku ktorým má používateľ prístup na platforme SeqOne (napríklad ak má používateľ prístup k viacerým entitám, platí pre ne rovnaká sada oprávnení).

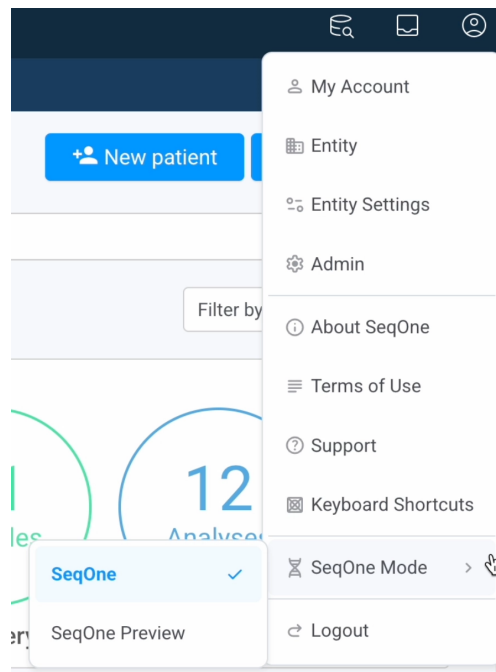
Ak sa vám zobrazí chybová správa, kontaktujte zákaznícku podporu na adrese support@seqone.com.

7.4.4. Prepínanie medzi režimami Main a Preview

Pripojenie k SeqOne prostredníctvom predvolenej prihlasovacej adresy URL presmeruje používateľov do režimu "Main" ich subjektu. Tento predvolený režim hostí pracovné sady s označením CE a ďalšie miestne certifikácie.

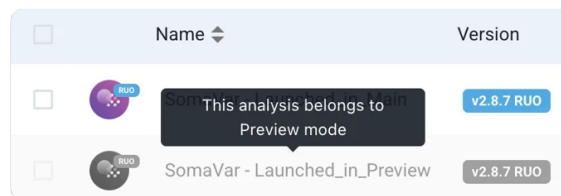
Pre skorší prístup k novým funkciám a verziám pracovných sád prepnite do režimu "Preview" (RUO: Research Use Only) tak, že prejdete do spodnej časti ponuky nastavení a v časti SeqOne One vyberiete "Preview".

Režimy "Main" a "Preview" zdieľajú rovnaké základné údaje (projekty, vzorky, analýzy), povolenia, zdroje (súbory, šablóny, manifesty, zoznamy génov), znalosti (VKB) a frekvencie (údaje o frekvencii vzoriek, výskyty).



Pri navigácii v režime "Preview" sa zobrazí logo "SeqOne Preview RUO" ako pripomienka, že tento režim obsahuje funkcie a pracovné sady RUO.

V režime "Main" sa analýzy spustené v "Preview" zobrazujú sivo. Ak kliknete na jednu z týchto sivých analýz, platforma vás vyzve, aby ste na jej zobrazenie prepli do režimu Preview.



8. Popis operačných krokov

Krok	Popis
1	Vytvorenie projektu Analýzy vykonávané na platforme SeqOne sú organizované v rámci projektov, ktoré definujú charakteristiky analyzovaných vzoriek prostredníctvom metadát zadaných používateľom (typ sekvenovania, metóda prípravy, gény záujmu).
2	Nahratie vzorky Vzorky sa nahrávajú ako surové sekvenčné súbory (FASTQ), ktoré by mali byť umiestnené v určenom projekte. Po nahratí súborov sa v rámci overovacieho kroku skontroluje integrita odoslaných súborov pred akoukoľvek analýzou.
3	Spustenie analýzy Po nahratí údajov na servery SeqOne si používateľ môže vybrať z ponuky aplikácií alebo „pracovných súborov“ v závislosti od typu analyzovanej vzorky a typov hľadaných genomických zmien (SNV, CNV, génové fúzie, varianty spliceovania atď.). Tieto aplikácie vykonávajú bioinformatickú analýzu.
4	Kontrola kvality údajov Aplikácie dostupné na platforme SeqOne obsahujú niekoľko modulov určených na analýzu kvality údajov, ktoré využívajú rôzne metriky sekvenovania, zarovnanie a hĺbku pokrytia.
5	Interpretácia výsledkov analýzy Genómové variácie zistené vo vzorke sa zobrazujú v rôznych kartách alebo „prehliadačoch“, kde môže používateľ vyhľadávať diagnostickú variantu zodpovednú za patológiu pacienta.
6	Získanie terapeutických odporúčaní a klinických štúdií súvisiacich s patológiou pacienta .
7	Vydanie správy Výsledky analýzy je možné exportovať ako správu vo formáte PDF, ktorú je možné následne stiahnuť.

9. Vytvorenie projektu

Analýzy vykonávané na platforme SeqOne sú organizované v rámci projektov, ktoré definujú charakteristiky analyzovaných vzoriek prostredníctvom metadát zadanych používateľom (typ sekvenovania, metóda prípravy, gény záujmu).

9.1. Vytvorenie projektu

Po kliknutí na „Nový projekt“ sa používateľom zobrazí výzva na zadanie hodnôt pre niekoľko typov informácií alebo metadát, ktoré sa používajú na popis projektu a súvisiacich vzoriek.

- **Popis projektu**
- **Sekvenovanie:**
 - Typ testu: typ analýzy (germline alebo somatický),
 - Projekty Germline (typ testu *Germline*) obsahujú začiarňovacie políčko UMI Trimming, ktoré umožňuje spracovanie vzoriek obsahujúcich UMI. Výberom tejto možnosti sa UMI odstránia zo sekvencií FASTQ, ako v režime s vypnutým UMI pre somatické pracovné súbory (pozri časť **Výber režimu spracovania UMI**).
 - Sekvenčná platforma: sekvenčná platforma
 - Typ knižnice: typ vykonaného sekvenovania čítania (paired-end alebo single-end),
 - Typ údajov: DNA, RNA, ctDNA, alebo Methylation Sequencing
 - Ak je ako typ údajov zvolená DNA, je potrebné špecifikovať metódu sekvenovania pomocou roletového menu „DNA Sequencing method“ a metódu amplifikácie (Capture alebo Amplicon) pomocou „Amplification method“.
 - Typ vstupného súboru: *súbory Fastq*, *súbory VCF*
 - Genóm: verzia referenčného genómu (GRCh37 alebo GRCh38),
 - Oblasť mapovania nízkej kvality: volanie variantov v oblastiach s nízkou kvalitou zarovnaním s ohľadom na čítania MAPQ0,
 - Metóda sekvenovania: génový panel, sekvenovanie celého exómu, komplexné genómové profilovanie, plytké WGS, sekvenovanie celého genómu, CGH mikročipy.
 - Poskytovateľ sekundárnej analýzy: AmoyDX, Pillar, Thermofisher,
 - Manifest: manifest v prípade génového panelu,
 - ID marker (voliteľné): súbor identifikačných markerov.
- **Vzorky:**
 - (Voliteľné) Nové metadáta vzorky je možné pridať kliknutím na „*You can also create your own metadata*“ („*Môžete tiež vytvoriť svoje vlastné metadáta*“). Po zobrazení výzvy zadajte názov a hodnotu metadát.

Po zadaní rôznych metadát pre projekt sú používatelia vyzvaní, aby vytvorili svoj projekt kliknutím na tlačidlo „Vytvoriť projekt“.

9.2. Správa projektov

Prostredníctvom riadiaceho panela môžu používatelia spravovať rôzne projekty, ktoré boli vytvorené na platforme. Konkrétne môžu:

- Prístup k projektom cez kartu *Projekty*, ako aj k vzorkám a analýzám pre každý projekt kliknutím na logá „*Samples*“ („Vzorky“) a „*Analysis*“ („Analýza“).
- Prístup k analýzám priamo cez kartu *Analýzy*.
- Vytvoriť nový projekt kliknutím na „*New project*“ („Nový projekt“, pozri časť s názvom **Vytvorenie projektu**).

9.3. Vymazanie projektu

Na vymazanie projektu musia používatelia:

- Vybrať projekt
- Kliknite na kartu „*Settings*“ („Nastavenia“) a potom na „*Delete project*“ („Odstrániť projekt“).
- Potvrdiť zaškrtnutím „*Yes*“ v zobrazenom okne. Táto akcia je nevratná.

9.4. DNA údaje (sekvenovanie DNA, sekvenovanie ctDNA)

Oblasť s nízkou kvalitou mapovania. Pri sekvenovaní génových panelov, exómov alebo genómov môže používateľ pri volaní variantov zahrnúť oblasti s nízkou kvalitou zarovnania. Tieto oblasti zodpovedajú takzvaným MAPQ0 čítaniam, ktorých kvalita zarovnania je rovná 0, ako napríklad pseudogény a homologické oblasti genómu.

Táto možnosť je použiteľná na špecifické cieľové gény alebo oblasti. Ak chce používateľ túto možnosť využiť, je vyzvaný, aby načítal manifest zameraný na požadované oblasti alebo aby ho vybral z roletového menu.

Táto možnosť môže znížiť kvalitu a presnosť výsledkov. Môže spôsobiť šum v frekvencii alel a mať negatívny vplyv na prioritizáciu variantov. Môže tiež zvýšiť čas analýzy v dôsledku vyššej výpočtovej záťaže.

Exóm. V prípade sekvenovania exómu používatelia nemusia špecifikovať manifest, ktorý sa má použiť; môžu vybrať výrobcu použitej súpravy kliknutím na jednu z ikon v kategórii Dodávateľa súprav pre exóm. Z roletového menu si potom používateľ môže vybrať požadovaný manifest. Ak požadovaný manifest nie je na platforme k dispozícii, používateľ je vyzvaný, aby kontaktoval podporu používateľov (support@seqone.com).

Génový panel. V prípade génového panelu sú používatelia vyzvaní, aby v roletovom menu vybrali *amplifikačnú metódu*, ako aj manifest (vo formáte BED) zodpovedajúci génom, ktoré ich

zaujímajú. Ak požadovaný manifest nie je k dispozícii na platforme, je možné ho nahrať výberom možnosti „Nahrať nový manifest“ a postupovaním podľa pokynov na obrazovke.

Povrchové WGS, CGH mikročipy, sekvenovanie celého genómu. V prípade sekvenovania genómu nemusí používateľ špecifikovať, ktorý manifest sa má použiť.

Zachytávanie. Výberom metódy amplifikácie zachytávania je k dispozícii ikona UMI (*Unique Molecular Identifiers, jedinečné molekulárne identifikátory*). Jedinečné molekulárne identifikátory sú indexy, ktoré sa pridávajú pred amplifikáciou, aby pomohli eliminovať duplikáty PCR a znížiť šum pozadia po analýze.

Amplicon. Pri analýze amplikónov je potrebné venovať osobitnú pozornosť amplifikačným primerom. Pri vytváraní projektu pre údaje o amplikónoch sa používateľom odporúča vybrať manifest, v ktorom regióny zodpovedajúce cieľom neobsahujú sekvencie primerov. Ak takýto BED nie je k dispozícii, je možné vybrať možnosť *odrezať* pevnú dĺžku na konci sekvencií.

Monitorovanie identity. Pri sekvenčných metódach zodpovedajúcich génovému panelu, exómu, genomickému profilovaniu (CGP) alebo genómu (povrchné WGS) zo súborov FASTQ je používateľ vyzvaný, aby z roletového menu vybral súbor markerov na monitorovanie identity. Ak požadovaný súbor nie je na platforme k dispozícii, je možné ho importovať výberom možnosti „Nahrať nové markery“ a postupovaním podľa pokynov na obrazovke (pozri tiež odsek **Markery** v časti **Správa používateľských nastavení**).

9.5. Údaje o RNA (sekvenovanie RNA)

Ak je pre typ údajov vybraná RNA, používatelia musia prostredníctvom roletového menu Stranded protocol (Protokol reťazca) určiť, či ich protokol zachováva pôvodné informácie o reťazci. Ak je protokol špecifický pre reťazec, poradie syntézy reťazcov (pre forward-reverse alebo reverse-forward) musí byť určené pomocou možnosti Read orientation (Orientácia čítania). Nakoniec je potrebné určiť metódu sekvenovania (genový panel).

Orientácia čítania. Na špecifikovanie orientácie sekvencie sú k dispozícii dve možnosti:

- **Forward-Reverse** : prvé čítanie (R1) pre každý pár zodpovedá sekvencii transkribovanej mediátorovej RNA (sense reťazec).
- **Reverse-Forward** : prvé čítanie (R1) pre každý pár zodpovedá komplementárnej sekvencii transkribovanej mediátorovej RNA (antisense reťazec).

Vo všeobecnosti je orientácia čítaní pre nižšie uvedené súpravy nasledovná:

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

V prípade pochybností kontaktujte podporu (support@seqone.com).

Po špecifikovaní typu údajov a správnom nakonfigurovaní projektu sa používatelia vyzývajú na potvrdenie kliknutím na tlačidlo „*Ďalší krok*“.

9.6. Nastavenia projektu

Na karte „Nastavenia“ projektu môžete:

- zmeniť názov a popis projektu v poliach „Názov projektu“ a „Popis projektu“,
- zmeniť profil kontroly kvality (QC),
- zmeniť predvolený profil filtra v tabuľke variantov,
- trvalo odstrániť projekt.

9.6.1. Nastaviť predvolené metriky QC projektu

Používateľ si môže zvoliť použitie predvolených metrík QC alebo definovať vlastné prahové hodnoty v záložke „Settings“ („Nastavenia“) projektu. Používateľ je vyzvaný, aby kontaktoval používateľskú podporu (support@seqone.com) a definoval požadované hodnoty.

Budú použité vlastné metriky QC:

- *Kontroly kvality v zobrazení Analýzy projektu,*
- *Kontroly kvality v sekcii automatického overovania kvality na karte analýzy kvality,*
- na karte GeneCov analýzy.

9.6.2. Nastavte predvolený profil filtra

Používateľ si môže zvoliť nastavenie preddefinovaného profilu filtra, ktorý sa uplatňuje na predvolenú tabuľku variantov, na karte „Nastavenia“ projektu (pozri tiež časť **Používanie preddefinovaného profilu filtra**).

Profil filtra je možné vybrať pre každú pracovnú sadu prostredníctvom roletového menu, ak je verejný. Profil filtra sa potom automaticky použije pri zobrazení tabuľky variantov pre všetky analýzy projektu pre všetkých používateľov v organizácii.

Ak tento parameter nie je používateľom upravený, pri vytvorení projektu bude predvolenou hodnotou profil definovaný spoločnosťou SeqOne.

Predvolené profily filtrov je možné uzamknúť kliknutím na prepínač "Lock Variants Filters".

10. Ukážka nahratia súboru

10.1. Nahratie vzoriek

Používatelia môžu importovať svoje vzorky prostredníctvom ikony „Pridať vzorky“ v každom projekte. Vzorky je možné importovať do platformy SeqOne dvoma spôsobmi:

Z počítača

Používatelia môžu nahráť svoje súbory FastQ priamo zo zložky vo svojom počítači kliknutím na kartu „Z vášho počítača“ v ikone „Pridať súbory“. Používatelia môžu tiež presunúť súbory zo svojho počítača do vyhradenej oblasti v rozhraní pomocou funkcie drag and drop.

Prostredníctvom SDS (SeqOne Data Sync)

SDS je voliteľný prémiový modul na automatické vytváranie projektov, prenos údajov, spúšťanie analýz a obnovovanie štruktúrovaných údajov (TSV) na integráciu do laboratórneho informačného systému (LIMS).

Tento modul kompatibilný s Linuxom, MacOS a Windows je spustiteľný súbor, ktorý používateľom umožňuje autentizáciu a interakciu s rozhraniami API SeqOne v preddefinovanom a vysoko zabezpečenom prostredí.

Minimálne požiadavky:

- Linux 1.19 minimálne
- MacOS 10.15 s procesorom Intel minimálne
- MacOS s procesorom Apple Silicon (čip M1) minimálne
- Windows 10 64-bitový minimálne
- 4 GB RAM (8 GB pre Windows)
- Štvorjadrový procesor
- Prístup k internetu a platforme SeqOne (app.eu.seqone.com)
- Prístup k systému ukladania dát (priamo alebo cez sieť)

Ak máte záujem o toto riešenie alebo potrebujete ďalšie informácie, kontaktujte náš tím podpory (support@seqone.com).

Stav nahratia vzorky

Platforma SeqOne používa rôzne stavy, aby používatelia mohli sledovať priebeh nahrávania svojich vzoriek:

- *Upload pending* (Nahratie čaká na spracovanie): import prebieha,
- *Verifying sample...* (Overovanie vzorky...): overovanie zhody vzorky,
- *Mapping...* (Mapovanie...): krok zarovnania vzorky,
- *Verified sample* (Overená vzorka): overená vzorka,
- *Upload failed* (Nahratie sa nepodarilo): import sa nepodaril.

Kapacita nahratia

Kapacita nahrávania súborov (najmä súborov fastq.gz) závisí priamo od internetového pripojenia používateľa. Pri nahrávaní vzoriek je preto potrebné zohľadniť rýchlosť internetového pripojenia. Pre jasnejšiu predstavu o odhadovaných časoch nahrávania pozrite nasledujúcu tabuľku:

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
100Mo	1min 20s	1min 45s	8s	9s
Poids du fichier fastq				
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

Obrázok 5: Odhadovaná doba nahrania podľa veľkosti súboru a rýchlosti internetového pripojenia.

10.2. Formát súboru na zadávanie vzoriek

10.2.1. Súbory primárnej analýzy

Formát súboru, ktorý sa má nahráť pre každú vzorku, závisí od použitej sekvenčnej platformy. Názvy súborov musia byť v súlade s konvenciou názvov súborov každého sekvenéra.

Súbory musia byť komprimované pomocou príponou .gz (gzip), aby mohli byť importované do platformy. V prípade párového sekvenovania sa očakávajú dva súbory, ktoré zodpovedajú čítaniu z oboch koncov každej molekuly.

Keď platforma správne rozpozna vzorku v čase importu, zobrazí sa celé meno súboru. Potom je možné súbor premenovať kliknutím naň. Súbory typu paired-end sa automaticky zlúčia do jednej vzorky.

Platforma podporuje nasledujúce sekvencery:

- **Illumina**: formát fastq.gz, iba párové konce,
- **Element Biosciences**: formát fastq.gz, iba párové konce,
- **MGI**: formát fastq.gz, iba párové konce. Očakávané názvy súborov sú: sample_1.fq.gz a sample_2.fq.gz **alebo** flowcellID_L01_sample_1.fq.gz a flowcellID_L01_sample_2.fq.gz

10.2.2. Súbory na sekundárnu analýzu

Formát súboru, ktorý sa má nahrať pre každú vzorku, závisí od sekvenčnej platformy a aplikácie použitej na sekundárnu analýzu. Súbory musia byť pred importom do platformy komprimované s príponou .gz (gzip), pokiaľ nie je uvedené inak.

Formát importu súborov sekundárnej analýzy generovaných pomocou aplikácií Illumina, Oxford Nanopore Technologies, Agilent, AmoyDx, Pillar, Thermofisher závisí od použitého worksetu (**pozri časť Worksety dostupné zo súborov sekundárnej analýzy**).

Keď platforma správne rozpozná vzorku počas importu, zobrazí kompletný názov súboru. Súbor môžete premenovať kliknutím naň. Súbory, ktoré majú rovnaký základný názov, sa automaticky zlúčia do jednej vzorky.

Súbory DRAGEN. Pri vytváraní projektu, ak sú zvolené súbory VCF (údaje sekundárnej analýzy) a metadáta Illumina, očakávané súbory a formáty pre každú pracovnú sadu sú nasledovné:

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

Súbory Oxford Nanopore Technologies. Pri vytváraní projektu môžu používatelia importovať súbory VCF (údaje sekundárnej analýzy) a metadáta Oxford Nanopore Technologies. Očakávané súbory a formáty umožňujú importovať samostatný súbor variantov (SNP) alebo ho kombinovať s doplnkovými súbormi pre analýzu CNV, STR alebo 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

Súbory Thermofisher/Affymetrix. Importný formát pre súbory sekundárnej analýzy CGH je formát .txt generovaný aplikáciou Affymetrix spoločnosti Thermofisher. Tento formát je podporovaný len vtedy, ak sú pri vytváraní projektu vybrané metadáta Thermofisher/Affymetrix aj CGH MicroArrays. Súbory a formáty očakávané pre každú pracovnú sadu sú nasledovné:

GermVar Tertiary ARRAY_SampleID.txt

Súbory CGH MicroArrays. Formát importu pre súbory sekundárnej analýzy CGH je formát .xls generovaný aplikáciou Cytogenomics spoločnosti Agilent. Tento formát je podporovaný len vtedy, ak sú pri vytváraní projektu vybrané metadáta Agilent aj CGH MicroArrays. Súbory a formáty očakávané pre každú pracovnú sadu sú nasledovné:

GermVar Tertiary IntervalBasedReport_sample.xls

Súbory Alissa Reporter. Importný formát pre sekundárne analytické súbory Alissa Reporter je formát .xls generovaný aplikáciou Alissa Reporter spoločnosti Agilent. Tento formát je podporovaný iba v prípade, ak sú pri vytváraní projektu vybrané metadáta Agilent aj CGH MicroArrays. Súbory a formáty očakávané pre každú pracovnú sadu sú nasledovné:

GermVar Tertiary IntervalBasedReport_sample.xls
ABERRATION_&_LOH_INTERVAL_sample.xls

PacBio-HiFi files. Ak sú pri vytváraní projektu vybraté súbory VCF (Secondary Analysis data) a metadáta PacBio-HiFi, očakávané súbory a formáty pre každý workset sú nasledovné:

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

Somatické génové panely VCF súbory. Importný formát pre SNV, CNV, fúzie sekundárne analytické súbory je formát .vcf.gz generovaný aplikáciami Illumina, AmoyDx, Pillar, Thermofisher. Pôvodné názvy súborov musia byť konvertované tak, aby zodpovedali požadovanému formátu názvov opísanému nižšie.

Pri výbere súborov VCF – sekundárnych analýz a metadát génového panelu počas vytvárania somatického projektu sa očakávajú nasledujúce súbory a formáty. Používatelia môžu importovať buď samotný súbor variantov (SNV), alebo ho kombinovať s doplnkovými súbormi pre analýzu CNV alebo fúzií:

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

Pri výbere „VCF súbory – údaje sekundárnej analýzy“ a „Komplexné genómové profilovanie“ metadáta počas vytvárania somatického projektu sa očakávajú nasledujúce súbory a formáty.

Používatelia môžu importovať multi-variantný VCF súbor, napríklad VCF súbor „Všetky varianty“ z Oncomine™ Comprehensive Assay V3/Plus.

SomaVar Tertiary

Oncomine_<LibPrepID>_<analysisID>.vcf

10.3. Vzorové metadáta

Pre každý nahratý vzor môžu používatelia priradiť hodnoty k metadátam, ktoré vybrali pri vytváraní projektu, alebo definovať nové hodnoty kliknutím na ikonu v pravom hornom rohu zoznamu vzoriek  .

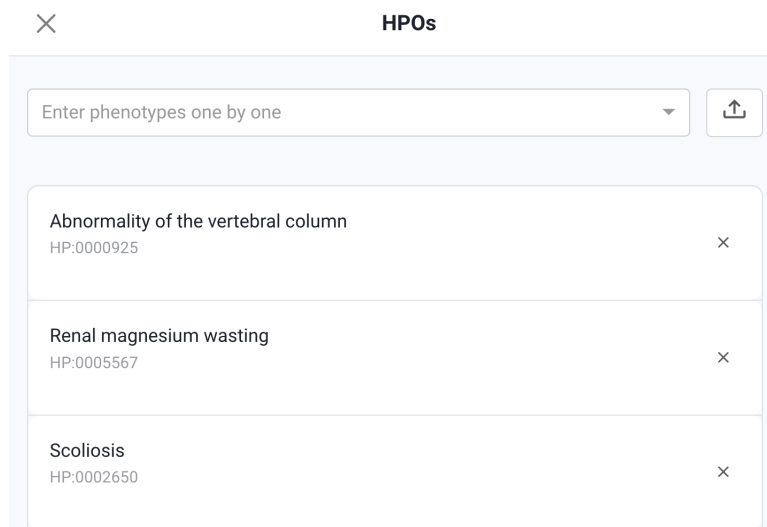
Niektoré metadáta vzoriek je možné nahrať priamo ako súbor s hodnotami oddelenými tabulátorom výberom ikony  .

Tieto kritériá zohrávajú dôležitú úlohu pri zabezpečovaní sledovateľnosti údajov a pri krížových odkazoch na rôzne analýzy vykonané na platforme. Niektoré metadáta, ako napríklad pohlavie, choroba a identifikátory HPO, sa podieľajú na procese identifikácie variantov.

Ak je pole „Assay Type“ (Typ testu) vyplnené hodnotou „Germline“, metadáta pridružená k vzorke sú štandardne „HPO“, „Sex“ („Pohlavie“) a skupina „GnomAD“. Ak je pole „Assay Type“ (Typ testu) vyplnené hodnotou „Somatic“ (Somatický), metadáta pridružená k vzorke sú štandardne „Disease“ (Choroba) a „Sex“ (Pohlavie).

Ak chcete pridať alebo upraviť HPO vo vzorke, po výbere vzorky môžete kliknúť na „+ Pridať“ v stĺpci HPO. Otvorí sa vyskakovacie okno, v ktorom môžete pridávať HPO jeden po druhom vyhľadávaním fenotypu alebo kódu HPO v vyhľadávacom riadku prepojenom s rozhraním API hpo.jax.org. Je tiež možné importovať textový súbor obsahujúci kódy HPO kliknutím na ikonu

 . Potom je k dispozícii zhrnutie zadaných termínov s možnosťou ich vymazania, ak je to potrebné.



Obrázok 6: Vyskakovacie okno na import HPO

10.4. Úprava vzorky

Názov, ako aj metadáta spojené so vzorkou môže používateľ kedykoľvek upraviť výberom vzorky v zozname „Vzorky“ a kliknutím na ikonu . Úpravy musia byť následne potvrdené, aby nadobudli platnosť.

10.5. Vymazanie vzorky

Používatelia môžu odstrániť súbory FastQ vzorky z platformy kliknutím na ikonu „“ po výbere požadovanej vzorky. Tieto súbory uvoľnia úložný priestor a výsledky budú naďalej dostupné na SeqOne.

Ak chcete vzorku odstrániť, kliknite na ikonu „“ (Odstrániť vzorku) v záložke „Samples“ (Vzorky) po výbere požadovanej vzorky. Vzorku je možné trvalo odstrániť až po vymazaní všetkých analýz, ktoré sa k nej viažu.

11. Analýza

11.1. Spustenie novej analýzy vzorky

Používatelia môžu spustiť novú analýzu kliknutím na „New analysis“ (Nová analýza) na obrazovke „Overview“ (Prehľad) pre každý projekt.

Spustenie analýzy sa vykonáva v troch samostatných krokoch, ktoré sú opísané nižšie.

Krok „1. Workset“ vám umožňuje vybrať vhodný workset analýzy (pozri časť **Výber worksetu analýzy**).

Krok „2. Select Samples“ (Výber vzoriek) je rozdelený na dve časti:

- 2.1. Nastaviť možnosti:
 - opisuje vybranú pracovnú sadu a referenčný genóm
 - umožňuje pridať vyplňanie prostredníctvom funkcie „Extra calling padding“ (ak je to vhodné) (pozri časť **Výber pracovnej sady analýzy**)
 - umožňuje výber režimu spracovania UMI (pozri časť **Výber režimu spracovania UMI**), ak je to vhodné.
 - umožňuje výber základnej hodnoty CNV, základnej hodnoty MSI alebo úpravu minimálnej prahovej hodnoty VAF, ak je to vhodné.
 - Umožňuje vybrať protokol.
- 2.2. Výber vzoriek: umožňuje výber vzoriek na analýzu. Používatelia môžu konkrétne vybrať nové vzorky, ktoré sú označené ako „New“ (Nové), pomocou funkcie „select all new“ (vybrať všetky nové).

Krok „3. Potvrdiť analýzu“ sumarizuje vybrané vzorky. Pre každú analýzu, ktorá sa má spustiť, je možné použiť obmedzený panel výberom jedného alebo viacerých génových panelov prostredníctvom možnosti „Obmedzený panel“ (pozri tiež časť **Modálne zobrazenie panelu**).



GermlineVar - Patient	Sample PATIENT	Restricted panel ⓘ		
-----------------------	--------------------------	--------------------	------------------------	---

Možnosť „Obmedzený panel“ uplatní na analýzu filter vo forme jedného alebo viacerých génových panelov. Táto možnosť nemá vplyv na volanie variantov, ktoré je určené manifestom projektu a pridaným vyplňaním. Výber obmedzených panelov je reverzibilný; používateľ môže tieto panely po analýze upraviť na karte *Panel* v bočnom paneli správy analýzy (pozri časť **Modálne zobrazenie panelu**).

11.1.1. Výber worksetu analýzy

Na zodpovedanie rôznych biologických otázok je k dispozícii niekoľko worksetov analýzy. Worksety navrhované pri vytváraní novej analýzy závisia od metadát, ktoré boli definované pri vytváraní projektu, a preto sa môžu líšiť v závislosti od údajov.

Analytické worksety dostupné z FASTQ súborov

Nasledujúce worksety sú k dispozícii, ak sa pri vytváraní projektu vyberie možnosť metadát Fastq súbory – Primárne analytické údaje.

S výnimkou projektov, kde je metódou amplifikácie amplikón, niektoré worksety (tj. GermlineVar, GermlineFamily, SomaVar, SomaVar LF (verzie ≥ 2.0), SomaCGP (verzia ≥ 0.6), GermVar, GermVar Family, SomaLBx, SomaMethyl) umožňujú rozšíriť oblasť vyhľadávania pre identifikáciu variantov a pridať k pôvodnému manifestu (padding) v rozsahu 0 až 100 bp.

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

Táto možnosť je k dispozícii v kroku 2 (Vybrať vzorku) spustenia analýzy prostredníctvom funkcie „Extra calling padding“ (Dodatočné vyplňanie volania). Predvolená hodnota je nastavená na 0 a musí byť prirodzeným celým číslom.

Niektoré pracovné sady (napr. SomaVar (verzie ≥ 2.6) a SomaHemato (verzie ≥ 0.2)) ponúkajú možnosť nastaviť minimálnu prahovú hodnotu VAF pre detekciu variantov. Táto prahová hodnota sa môže nastaviť v rozmedzí od 0,01 % do 1 %, pričom 1 % je predvolená hodnota.

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)

Ďalšie informácie o tejto možnosti získate od používateľskej podpory na adrese support@seqone.com.

GermlineVar



GermlineVar sa používa na detekciu mutácií v zárodočných bunkách na základe sekvenovania génového panelu alebo exómu pacienta.

Pri spustení analýzy GermlineVar sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorý chcú spustiť, a kliknúť na „Vybrať“.
- Vybrať vzorky, ktoré sa majú analyzovať, a potom kliknúť na „Ďalší krok“.
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

GermVar



GermVar umožňuje detekciu mutácií v zárodočných bunkách na základe sekvenovania génového panelu, exómu alebo genómu pacienta.

Pri spustení analýzy GermVar sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorý chcú vykonať, a kliknúť na „Vybrať“.
- Vyberte model, ktorý sa má použiť,
- Vybrať vzorku (vzorky) na analýzu a kliknúť na „Ďalší krok“.
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

GermlineVarLR



GermlineVarLR sa používa na detekciu mutácií v germline v údajoch sekvenovania s dlhým čítaním zo sekvenovania génového panelu.

Pri spustení analýzy GermlineVar LR sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorý chcú vykonať, a kliknúť na „Vybrať“.
- Vybrať model, ktorý sa má použiť, a vzorku (vzorky), ktoré sa majú analyzovať, a potom kliknúť na „Ďalší krok“,
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

GermlineFamily



GermlineFamily sa používa na detekciu dedičných mutácií na základe sekvenovania génového panelu, exómu alebo genómu pacienta a členov jeho rodiny. Táto pracovná sada sa môže použiť aj na skríning nositeľov.

Pri spustení analýzy GermlineFamily sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorý chcú spustiť, a kliknúť na „Vybrať“.
- Vybrať aspoň 2 vzorky.
- V časti „Nastaviť podmienky“ uveďte, či vzorka patrí pacientovi alebo členovi rodiny, a to tak, že v časti „Rodinná väzba“ definujete rodinnú väzbu. Pacient je automaticky označený ako „postihnutý“. V prípade skríningu nositeľov uveďte pre každú vzorku rodinný vzťah ako „Matka“ alebo „Otec“ bez zaškrtnutia políčka „Je postihnutý?“.
- Skontrolujte, či sa parametre zhodujú, a kliknite na „Ďalší krok“.
- Spustíte analýzu kliknutím na „Začať analýzu“.

GermVar Family



GermVar Family umožňuje detekciu mutácií v zárodočných bunkách na základe sekvenovania génového panelu, exómu alebo genómu pacienta.

Pri spustení analýzy GermVar Family sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorý sa má vykonať, a kliknúť na „Vybrať“,
- Vyberte model, ktorý sa má použiť,
- Vybrať aspoň 2 vzorky.
- V časti „Nastaviť podmienky“ uveďte, či vzorka patrí pacientovi alebo členovi rodiny, a to tak, že v časti „Rodinná väzba“ definujete rodinnú väzbu. Pacient je automaticky označený ako „postihnutý“. V prípade skríningu nositeľov uveďte pre každú vzorku rodinný vzťah ako „Matka“ alebo „Otec“ bez zaškrtnutia políčka „Je postihnutý?“.
- Skontrolujte, či sa parametre zhodujú, a potom kliknite na „Ďalší krok“.
- Spustíte analýzu kliknutím na „Začať analýzu“.

CNVCapture



CNVCapture sa používa na detekciu variácií počtu kópií germline (CNV) z údajov zachytených génovým panelom.

Pri spustení analýzy CNVCapture sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorý chcú spustiť, a kliknúť na „Vybrať“.
- Vybrať minimálny počet 8 vzoriek na analýzu a kliknúť na „Ďalší krok“.
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

CNVExome

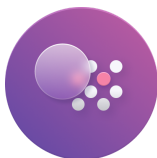


CNVExome sa používa na detekciu variácií počtu kópií (CNV) z údajov zachytených exómom.

Pri spustení analýzy CNVExome sú používatelia vyzvaní, aby:

- Vyberte verziu worksetu, ktorý chcete vykonať, a kliknite na „Vybrať“.
- Vybrať model, ktorý sa má použiť, a vzorku (vzorky), ktoré sa majú analyzovať, a potom kliknúť na „Ďalší krok“,
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

SomaVar



SomaVar detekuje somatické mutácie z nádorových vzoriek. Táto workset sa dá použiť s údajmi o zachytení a amplikónoch.

Pri spustení analýzy SomaVar sú používatelia vyzvaní, aby:

- Vyberte verziu worksetu, ktorý chcete spustiť, a kliknite na „Vybrať“.
- V možnostiach vyberte príslušný režim spracovania UMI z režimov, ktoré sú vhodné pre analýzu (pozri časť **Výber režimu spracovania UMI**).
- Vyberte vzorku (vzorky), ktoré chcete analyzovať, a kliknite na „Ďalší krok“,
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

SomaCNVCapture



SomaCNVCapture sa používa na detekciu somatických variácií počtu kópií (CNV) z údajov zachytených génovým panelom.

Pri spustení analýzy SomaCNVCapture sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorú chcú spustiť, a kliknúť na „Vybrať“.
- V možnostiach vybrať príslušný režim spracovania UMI z režimov, ktoré sú vhodné pre analýzu (pozri časť **Výber režimu spracovania UMI**),
- Vybrať minimálny počet ôsmich vzoriek, ktoré sa majú analyzovať, a potom kliknúť na „Ďalší krok“,
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

SomaRNA



SomaRNA sa dá použiť na detekciu génových fúzií z údajov zachytených pomocou RNA-seq.

Pri spustení analýzy SomaRNA sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorú chcú spustiť, a kliknúť na „Vybrať“.
- V možnostiach vyberte príslušný režim spracovania UMI z režimov, ktoré sú vhodné pre analýzu (pozri časť **Výber režimu spracovania UMI**).
- vybrať vzorku (vzorky), ktoré sa majú analyzovať, a kliknúť na „Ďalší krok“,
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

SomaMSI



SomaMSI detekuje stav mikrosatelitných oblastí (MSS alebo MSI).

Pri spustení analýzy SomaMSI sú používatelia vyzvaní, aby:

- Vybrať minimálne osem vzoriek, ktoré majú byť analyzované, a potom kliknúť na „Ďalší krok“.
- V možnostiach vyberte príslušný režim spracovania UMI z režimov, ktoré sú vhodné pre analýzu (pozri časť **Výber režimu spracovania UMI**).
- vybrať minimálny počet 8 vzoriek na analýzu a kliknúť na „Ďalší krok“,
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

SomaHRD

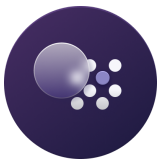


SomaHRD umožňuje detekciu stavu HRD len na základe povrchových údajov WGS alebo v kombinácii s panelom génov vrátane BRCA1 a BRCA2.

Pri spustení analýzy SomaHRD sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorú chcete spustiť, a kliknúť na „Vybrať“,
- Zadaťte metadáta o bunkovosti (% nádorových buniek) pre každú vzorku (pozri časť **Metadáta vzorky**),
- vybrať vzorku (vzorky), ktoré sa majú analyzovať, a kliknúť na „Ďalší krok“,
- Vybrať vhodnú analýzu SomaVar alebo SomaCGP (GRCh37 alebo GRCh38) pre každú vzorku pomocou možnosti „Vybrať panelovú analýzu (voliteľné)“,
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

SomaVar LF

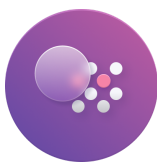


Detekcia nízko-frekvenčných krátkych variantov a štrukturálnych variantov, ako aj fúzií génov z nádorových vzoriek (tuhé alebo tekuté biopsie, cirkulujúca nádorová DNA) obohatených zachytením na paneli génov.

Pri spustení analýzy SomaVar LF sú používatelia vyzvaní, aby:

- Vyberte verziu worksetu, ktorú chcete spustiť, a kliknite na „Vybrať“.
- V možnostiach vybrať príslušný režim spracovania UMI z režimov, ktoré sú vhodné pre analýzu (pozri časť **Výber režimu spracovania UMI**),
- Vyberte vzorku (vzorky), ktoré chcete analyzovať, a kliknite na „Ďalší krok“.
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

SomaCGP



SomaCGP umožňuje detekciu somatických mutácií, štrukturálnych variantov, génových fúzií, ako aj stavu MSI a TMB z nádorových vzoriek. Tento workset je určený na analýzu FFPE DNA pripravenej pomocou panelu Agilent SureSelect Cancer CGP Assay a je k dispozícii len v prípade, ak sú pri vytváraní projektu vybrané metadáta komplexného genómového profilovania. Tento workset sa môže používať s capture dátami.

Pri spustení analýzy SomaCGP sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorú chcú spustiť, a kliknúť na „Vybrať“.
- Ak je to požadované, vybrať „CNV baseline“ a/alebo „MSI baseline“ (ak nie je vybraná žiadna základná hodnota, analýza bude prebiehať bez detekcie CNV/MSI),
- V možnostiach vyberte príslušný režim spracovania UMI z režimov, ktoré sú vhodné pre analýzu (pozri časť **Výber režimu spracovania UMI**),
- Vyberte vzorku (vzorky), ktoré chcete analyzovať, a kliknite na „Ďalší krok“.
- Spustíte analýzu kliknutím na „Spustiť analýzu“.

SomaHemato

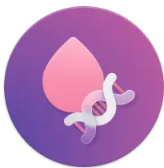


SomaHemato umožňuje detekciu somatických mutácií zo vzoriek s hematologickými špecifikami. Tento workset sa môže používať s údajmi o zachytení.

Pri spustení analýzy SomaHemato sú používatelia vyzvaní, aby:

- Vyberte verziu worksetu, ktorý chcete spustiť, a kliknite na „Vybrať“.
- V možnostiach vyberte príslušný režim spracovania UMI z režimov, ktoré sú vhodné pre analýzu (pozri časť **Výber režimu spracovania UMI**).
- Vyberte vzorku (vzorky), ktoré chcete analyzovať, a kliknite na „Ďalší krok“.
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

SomaLBx



SomaLBx umožňuje detekciu nízko-frekvenčných somatických mutácií, štruktúrálnych variantov a génových fúzií z nádorových vzoriek. Táto workset je určená na analýzu ctDNA a môže sa používať s údajmi zachytenými zachytením.

Pri spustení analýzy SomaLBx sú používatelia vyzvaní, aby:

- Vyberte verziu worksetu, ktorú chcete spustiť, a kliknite na „Vybrať“.
- Vyberte „CNV baseline“ (ak nie je vybraná žiadna základná línia, analýza prebehne bez detekcie CNV).
- V možnostiach vyberte príslušný režim spracovania UMI z režimov, ktoré sú vhodné pre analýzu (pozri časť **Výber režimu spracovania UMI**).
- Vyberte vzorku (vzorky), ktoré sa majú analyzovať, a kliknite na „Ďalší krok“.
- Spustíte analýzu kliknutím na „Spustiť analýzu“.

SomaMethyl



SomaMethyl umožňuje detekciu metylovaných CpG ostrovčekov z vzoriek cirkulujúcej tumorovej DNA. Táto pracovná sada je určená na analýzu metylácie v ctDNA a môže sa používať s údajmi zachytenými bisulfátom.

Pri spustení analýzy aSomaMethyl sú používatelia vyzvaní, aby:

- Vyberte verziu worksetu, ktorú chcete spustiť, a kliknite na „Vybrať“.
- Vyberte „Methyl baseline“ zodpovedajúcu kontrolnej kohorte (ak nie je vybraná žiadna kohorta, analýza sa spustí bez detekcie prahu a diferenciálnej metylácie).
- Vyberte vzorku (vzorky), ktoré sa majú analyzovať, a kliknite na „Ďalší krok“.
- Spustíte analýzu kliknutím na „Spustiť analýzu“.

Analytické pracovné sady dostupné zo sekundárnych analytických súborov

Nasledujúce pracovné sady sú k dispozícii, ak sú pri vytváraní projektu vybrané *súbory VCF – údaje sekundárnej analýzy* alebo ak sú vybrané metadáta Agilent aj CGH MicroArrays (pozri časť **Súbory sekundárnej analýzy**). V tomto prípade platforma prijíma výstupné údaje z aplikácie vybraného poskytovateľa, pokračuje v anotovaní a vkladaní variantov do príslušných vizualizačných rozhraní.

GermVar Tertiary



Pipeline analýzy jedného vzorku variantov detegovaných aplikáciou Illumina Germline alebo aplikáciou EPI2ME od Oxford Nanopore Technologies, z údajov WGS a WES alebo variantov detegovaných aplikáciou Cytogenomics od Agilent z údajov MicroArray CGH.

Pri spustení terciárnej analýzy GermVar sú používatelia vyzvaní, aby:

- Vyberte verziu worksetu, ktorú chcete spustiť, a kliknite na „Vybrať“.
- Vybrať vzorku (vzorky), ktoré sa majú analyzovať, a kliknúť na „Ďalší krok“.
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

GermVar Tertiary Family

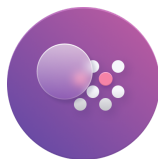


Pipeline analýzy viacerých vzoriek variantov detegovaných aplikáciou Illumina Germline alebo aplikáciou EPI2ME od Oxford Nanopore Technologies, z údajov WGS a WES.

Pri spustení analýzy GermVar Tertiary Family sú používatelia vyzvaní, aby:

- Vyberte verziu worksetu, ktorú chcete spustiť, a kliknite na „Vybrať“.
- Vyberte aspoň 2 vzorky.
- V časti „Nastaviť podmienky“ uveďte, či ide o vzorku pacienta alebo člena rodiny, a to tak, že v časti „Rodinné prepojenie“ definujete rodinné prepojenie. Pacient je automaticky označený ako „postihnutý“.
- Skontrolujte, či nastavenia zodpovedajú.
- Spustíte analýzu kliknutím na „Spustiť analýzu“.

SomaVar Tertiary



SomaVar Tertiary je terciárny analytický nástroj založený na VCF, ktorý spracováva súbory Illumina WES VCF aj súbory génového panelu VCF od rôznych dodávateľov súprav (pozri časť **Formát vstupného súboru**).

Pri spustení terciárnej analýzy SomaVar sú používatelia vyzvaní, aby:

- Vybrať verziu worksetu, ktorú chcú spustiť, a kliknúť na „Vybrať“.
- Vybrať vzorku (vzorky), ktoré sa majú analyzovať, a kliknúť na „Ďalší krok“.
- Spustiť analýzu kliknutím na „Spustiť analýzu“.

11.1.2. Výber režimu spracovania UMI

Prezentácia režimov spracovania UMI

V praxi platforma navrhuje jednu alebo viacero metód spracovania UMI pri spustení analýzy:

1. **Štandardný režim** (odporúčaný): konsenzus sa generuje z PCR duplikátov s rovnakým UMI, ak je ich počet väčší alebo rovný 2. UMI reprezentované jedinou sekvenciou (singletony) sa tiež zachovávajú. Minimálna alelová frekvencia hlásených variantov a minimálny počet čítaní podporujúcich mutáciu (AO, Alternate Observation) sa líšia v závislosti od použitej pracovnej sady:
 - V SomaVar, SomaHemato, SomaRNA, SomaCNVCapture, SomaCGP a SomaMSI je minimálna alelová frekvencia 1 % a AO je minimálne 3 čítania.
 - V SomaVar LF a SomaLBx je minimálna alelová frekvencia 0,1 % a minimálne AO je 2 čítania. Táto pracovná sada umožňuje detekciu variantov s alelovou frekvenciou menšou ako 1 %. Odporúča sa, ak je hĺbka sekvenovania väčšia ako 5000x, a pre aplikácie ako sekvenovanie ctDNA.
2. **UMI deaktivované** (k dispozícii iba v SomaVar, SomaRNA, SomaHemato, SomaLBx, SomaCNVCapture, SomaCGP a SomaMSI): UMI sa nepoužívajú na deduplikáciu a pred analýzou sa odstránia z konca sekvencií. Preto sú údaje generované týmto analytickým postupom surové (nededuplikované). Minimálna alelová frekvencia hlásených variantov je 1 % a minimálny počet čítaní podporujúcich mutáciu (AO) je 3.

Rozdiely v nastaveniach medzi týmito režimami sú zhrnuté v nasledujúcej tabuľke:

	UMI vypnuté	UMI štandard
Minimálna mediánová/priemerná kvalita mutovanej bázy (Phred skóre)*	30	30
Minimálny počet čítaní podľa konsenzu	N/A	2
Minimálna kvalita každej bázy v konsenze (Phred skóre)	N/A	30
Čítania bez filtrovaných konsenzov	N/A	nie

Tabuľka 1: Nastavenia rôznych konfigurácií spracovania UMI

* Tieto filtre sa vzťahujú len na varianty, ktoré nie sú vybrané modelom SVM po volaní variantov (potrubia SomaVar a SomaRNA).

Proces spracovania UMI

Štruktúra UMI závisí od súpravy každého dodávateľa a na platforme SeqOne sú v súčasnosti k dispozícii nasledujúce súpravy:

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

 Warning: Konfigurácia UMI "QIAseq Targeted DNA Pro" by sa mala používať iba pre SomaVar s dĺžkou čítania aspoň 150 bp. Pre viac informácií kontaktujte podporu.

Spracovanie UMI je založené na použití súboru nástrojov fgbio, ktorý dokáže spracovať UMI a čítania označené UMI. Náš proces analýzy UMI pozostáva z nasledujúcich krokov:

1. V prípade potreby odstráňte UMI z konca sekvencií a vygenerujte neusporiadaný BAM z dvojice súborov FASTQ.
2. Vyrovnajte sekvencie bez UMI pomocou BWA-MEM2 (DNA/ctDNA) alebo STAR (RNA)
3. Zlúčte súbory BAM obsahujúce UMI so zarovnanými sekvenciami pomocou ZipperBam,
4. Skupina číta spoločne, ak má rovnakú pozíciu a zdieľa rovnaké UMI,
5. Vygenerujte konsenzuálnu sekvenciu pre každú skupinu UMI.
6. Filtrujte každú bázu podľa jej kvality: spoľahlivosť konsenzuálnej sekvencie sa zvyšuje s počtom čítaní na skupinu.
7. Filtrujte konsenzuálne sekvencie: ak je filtrovaných viac ako 20 % báz v konsenzuálnej sekvencii, konsenzuálna sekvencia sa nezachová.
8. Premena týchto filtrovaných konsenzuálnych sekvencií na FASTQ a konečné opätovné zarovnanie na referenčnom genóme pomocou BWA-MEM2 (DNA/ctDNA) alebo STAR (RNA).

Kompatibilita s analytickými worksetmi

Sady SomaVar, SomaHemato, SomaCGP, SomaLBx, SomaMethyl, SomaVar LF, SomaCNVCapture, SomaMSI a SomaRNA sú kompatibilné s údajmi UMI. Možnosti spracovania, ktoré sú k dispozícii pri spustení rôznych worksetov, sú uvedené v tabuľke nižšie:

	UMI deaktivované	UMI štandard
SomaVar	Áno	Áno
SomaCNVCapture	Áno	Áno
SomaMSI	Áno	Áno
SomaRNA	Áno	Áno
SomaVar LF	Nie	Áno

SomaCGP	Áno	Áno
SomaLBx	Áno	Áno
SomaHemato	Áno	Áno
SomaMethyl	Nie	Áno

Tabuľka 2: Možnosti spracovania UMI dostupné na stránke pri spustení kompatibilných worksetov.

11.2. Stav analýzy

Po spustení analýzy platforma SeqOne používa rôzne stavy, aby používatelia mohli sledovať priebeh svojej analýzy:

- **Bioinformatika Čaká:** platforma čaká na ukončenie inej analýzy
- **Bioinformatika prebieha:** analýza prebieha; tento stav tiež ukazuje, ktorá fáza analýzy sa práve vykonáva
- **Bioinformatika Dokončená:** analýza bola dokončená
- **Analýza zlyhala:** počas analýzy došlo k chybe; v tomto prípade sú používatelia vyzvaní, aby kontaktovali tím podpory používateľov
- **Interpretácia – nová:** analýza je dokončená a pripravená na interpretáciu, zatiaľ ju neotvoril žiadny používateľ
- **Analýza zlyhala – chyba:** Pri analýze došlo k chybe. V tomto prípade je používateľ vyzvaný, aby kontaktoval podporu používateľov.
- **Analýza zlyhala – zrušená:** analýza sa nespustila, pretože pri načítaní alebo overovaní vzorky, ktorá bola k nej pripojená, došlo k chybe.

11.3. Kvalita vzorky

Karta QC projektu je rozdelená na 2 podkarty: „Všeobecné štatistiky“ a „Súhrn kontroly ID“.

Všeobecné štatistiky. Táto karta zobrazuje tabuľku, ktorá predstavuje celkovú kvalitu zarovnania a pokrytia pre každú vzorku projektu. Používateľ môže tabuľku exportovať ako súbor TSV kliknutím na  .

Výpočet je založený na nástroji SeQC, programe vyvinutom spoločnosťou SeqOne na výpočet štatistík zarovnania a pokrytia zo všetkých čítaní zarovnaných s genómom.

Používateľ si môže zvoliť použitie predvolených metrík QC alebo definovať vlastné prahové hodnoty na karte „Nastavenia“ projektu (pozri časť **Konfigurácia predvolených metrík QC projektu**).

Navrhované metriky sú nasledovné:

- celkový počet čítaní

- priemerná dĺžka čítaní
- medián veľkosti fragmentov DNA v knižnici (inserty)
- percento kvalitných readov >30
- percento readov zosúladených s referenčným genómom
- percento readov na target
- percento duplicitných čítaní
- pomer tranzícií k transverziám
- Priemerné pokrytie
- Podiel targetov pokrytých pri rôznych prahových hodnotách v závislosti od kontextu (amplikón, capture).

Ciele panelu alebo exómu sa určujú na základe súboru BED vybraného pri vytvorení projektu.

Súhrn kontroly ID. Táto karta podrobne opisuje genotypy všetkých markerov zahrnutých v analýzach pre každú vzorku (pozri odsek **Marker** v časti **Správa parametrov používateľa**). Každý riadok zodpovedá jednej vzorke a každý stĺpec predstavuje marker s príslušným ID. Používateľ môže exportovať túto tabuľku ako súbor TSV kliknutím na  .

12. Kontrola kvality údajov

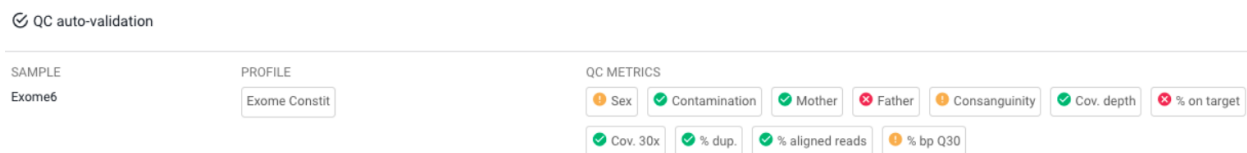
12.1. Metriky QC a správa

12.1.1. Metriky QC a kontroly kvality QC

Na karte QC má používateľ prístup k kontrolám kvality uplatneným na analýzu vo forme podkaret popísaných v nižšie uvedených častiach: *Správa QC*, *Kontrola ID* (ak je to relevantné) a *Technická správa*.

V prípade analýzy MicroArray CGH nie je karta QC k dispozícii. Technická správa je však k dispozícii na karte *Súbory analýzy* vo formáte *report.html*.

V hornej časti karty QC umožňuje sekcia „QC auto-validation“ (Automatické overenie QC) používateľovi farebne označiť metriky QC ako „*Passed*“ (Úspešné), „*Warning*“ (Upozornenie) alebo „*Failed*“ (Neúspešné). Kliknutím na príslušnú ikonu sa používateľ dostane k uplatneným prahovým hodnotám a podrobnostiam QC: pohlavie, kontaminácia, rodičovstvo a príbuzenské kríženie, hĺbka pokrytia, % na cieľi, 30X alebo 100X pokrytie, % duplikátov, % čítaní zarovnaných s referenčným genómom alebo %bp Q30.



12.1.2. Správa QC

Globálna *správa QC (kontrola kvality)* obsahuje rôzne metriky opisujúce kvalitu a zarovnanie sekvencií, pokrytie regiónov v manifeste a niektoré štatistiky súvisiace s variantmi. Táto správa sa skladá z nasledujúcich častí:

Alignment details (Podrobnosti zarovnania) (SomaMethyl). Celkový počet spracovaných čítaní, unikátne a neunikátne zarovnané čítania a miery mismatchov.

Deduplication (Deduplikácia) (SomaMethyl). Metriky duplikácie pre zarovnané čítania, ktoré upozorňujú na potenciálne problémy s prípravou knižnice.

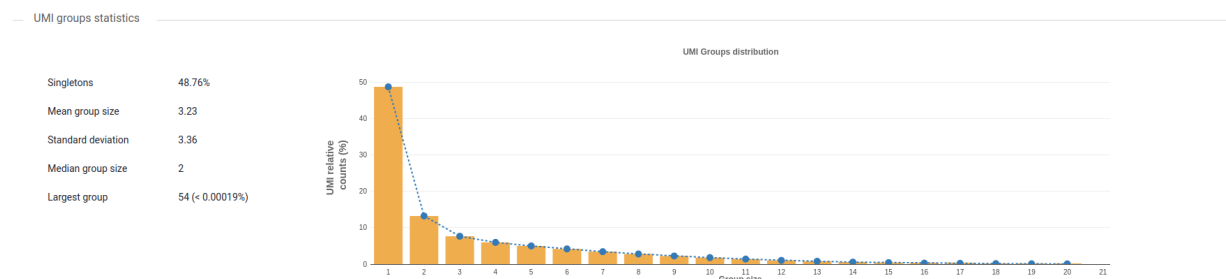
Cytosine coverage & methylation (Pokrytie cytosínu a metylácia) (SomaMethyl). Miery konverzie cytosínu na uracil pre každý kontext (CpG, CHG, CHH).

Štatistiky skupín UMI (Ilen SomaVar, SomaVar LF, SomaMethyl a SomaRNA UMI). Pozrite si odsek **Výber režimu spracovania UMI**.

Správa o kontrole kvality analýz SomaVar, SomaVar LF, SomaMethyl a SomaRNA ponúka podrobnejší pohľad na zloženie vzorky, najmä na rozloženie UMI podľa počtu sekvencií, ktoré ich obsahujú.

Štatistiky opisujú najmä:

- Podiel singletonov (UMI nesúcich jedinečnú sekvenciu)
- Priemernú a mediánovú veľkosť skupín UMI, ako aj štandardnú odchýlku
- Veľkosť najväčšej skupiny a podiel celkového počtu sekvencií, ktoré obsahuje.



Obrázok 7: Štatistiky týkajúce sa distribúcie UMI z hľadiska počtu sekvencií vo vzorke.

Globálne metriky QC pre projekt sú tiež podrobnejšie a rozdelené do troch častí:

- Metriky QC po spracovaní UMI, ktoré indikujú trend v každej metrike (ako percento) na základe surových údajov.
- Surové metriky QC vypočítané zo všetkých sekvencií (pred spracovaním UMI).
- Distribúcia UMI z hľadiska počtu sekvencií v každej vzorke.

Odhad krížovej kontaminácie (len SomaVar LF). Krížová kontaminácia sa odhaduje individuálne pomocou nástroja GATK CalculateContamination. Ten využíva signál referenčných alií detegovaných v homozygotných lokusoch na výpočet podielu sekvencií v jednotlivcovi pochádzajúcich z krížovej kontaminácie.

Integrity a identita vzorky (len analýzy germline). V tejto časti sú uvedené metriky súvisiace s integrity a identitou vzorky:

- **Rodinná väzba:** metadáta definované používateľom
- **Pôvodné pohlavie:** metadáta súvisiace s pohlavím zadané používateľom (muž/žena/neznáme)
- **Zistené pohlavie:** zistené pohlavie (muž/žena/neznáme)
- **Koeficient príbuznosti:** koeficient príbuznosti vypočítaný pomocou nástroja Somalier pre každú vzorku v analýze rodiny, tento ukazovateľ zobrazuje N/A pre samotnú vzorku.
- **Kontaminácia:** odhadované percento kontaminácie vzorky cudzou DNA pomocou nástroja verifyBamID2.

Uvedený je aj počet homozygotných a heterozygotných variantov, pomer týchto dvoch hodnôt (pomer Hom/Het), počet homozygotných a heterozygotných variantov na chromozóme X, ekvivalentný pomer na chromozóme X (pomer Het/Hom na chr X) a počet variantov na chromozóme Y.

Metriky Family link (Rodinné prepojenie) a Relationship coefficient (Koeficient príbuznosti), ktoré sa používajú na overenie vzťahu medzi jednotlivcami v rodine, sa zobrazujú iba v analýzach trojčlenných rodín (GermlineVar Family alebo GermVar Tertiary Family).

Metriky pohlavia sa používajú na kontrolu zhody medzi zadaným pohlavím a pohlavím zisteným v údajoch.

Metrika kontaminácie sa nepočíta pre pracovné súbory GermVar Tertiary a GermVar Tertiary Family.

Kontroly kvality týkajúce sa zhody pohlavia, kontaminácie a vzťahu sú k dispozícii na úrovni projektu alebo analýzy, aby bolo možné ľahko identifikovať akékoľvek zhody alebo nezhody.

Metriky zarovnania. V tejto časti sú uvedené hlavné metriky vypočítané na konci zarovnania, napríklad:

- Celkový počet sekvencií vo vzorke,
- Stredná veľkosť vložiek, vypočítaná na základe vzdialenosti medzi dvoma čítaniami rovnakého páru, ktorých príslušná dĺžka je zahrnutá,
- Priemerná dĺžka sekvencie,
- Podiel sekvencií úspešne zarovnaných s referenčným genómom,
- Miera duplikátov (len zachytenie),
- Podiel báz, ktorých kvalita je vyššia ako 30,
- Priemerné pokrytie cieľových oblastí, určené na základe analýzy manifestu.

Pokrytie. Stĺpcový graf znázorňujúci podiel cieľových oblastí pokrytých pri rôznych prahoch,

View as table



zobrazených na osi x. Kliknutím na určené tlačidlo môžu používatelia zobraziť hodnoty pokrytia v tabuľkovom formáte, ktoré je možné skopírovať a vložiť.

Grafy QC. Grafy z analýzy kvality FastQC, ktoré ilustrujú najmä obsah GC v sekvenciách, ako aj kvalitu báz podľa ich polohy v sekvenciách.

Štatistiky variantov. Zoznam štatistík týkajúcich sa variantov zistených v analýze, ktoré zahŕňajú najmä:

- Počet prechodov (Ts) a transverzií (Tv) medzi zistenými substitúciami,
- Pomer Ts / Tv,

- Celkový počet zistených mutácií,
- Podiel SNV, MNV a indelov v tomto celkovom počte.

12.1.3. Správa o monitorovaní identity „ID-Check“

Karta „ID check“ je venovaná výsledkom monitorovania identity v karte QC pracovných súborov volajúcich varianty.

Výsledky sa zobrazujú v tabuľkovej forme s nasledujúcimi informáciami:

- Popis každého markera na základe genomických súradníc, napríklad: chr-start-ref-alt,
- označenie (rsID, vlastné označenie), ak ho zadal používateľ,
- pozorovaný genotyp (GT) alebo frekvencia alely (VAF),
- hĺbka pokrytia v mieste markera (COV),
- kvalita databázy (QUAL).

Ak je do analýzy zahrnutých viacero vzoriek, stĺpce špecifické pre vzorku (GT/VAF, QUAL, COV) sa opakujú pre každú vzorku individuálne.

Výberom variantu v tabuľke sa otvorí prehliadač IGV v novej karte, zameraný na genómové súradnice variantu.

12.1.4. Technický report

Technická správa uvádza zoznam nástrojov použitých počas analýzy a ich verzie. Je dostupná na karte QC analýzy, potom v časti „Technická správa“.

Konkrétne opisuje:

- všeobecné informácie o konfigurácii a metadátach projektu a analýzy, ako napríklad typ sekvenovania, verzia genómu a verzia pracovnej sady, ktoré sa použili;

[QC report](#) [ID check](#) [Technical report](#)

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 použitom manifeste, ako je názov, veľkosť, dĺžka (bp) a vyplňanie pridané do analýzy;
- podrobnosti o použitých databázach a bioinformatických nástrojoch.

12.1.5. Správa UPD

Správa UPD je k dispozícii v sekcii QC kompatibilných zárodočných analýz (napr. GermVar a súvisiace pracovné postupy). Zhŕňa výsledky detekcie UPD a poskytuje vizuálne prvky na podporu interpretácie.

Správa UPD zvyčajne obsahuje:

- Prehľad výsledkov: indikácia, či analýza detegovala udalosť podobnú UPD a typ udalosti (napr. úplná alebo segmentálna; pri analýzach tria môže správa tiež podporiť rodičovský pôvod a podtypy udalostí, ako je UPD | Iso / UPD | Het).
- Zoznam udalostí: zoznam detegovaných udalostí s ich chromozomálnym umiestnenením a veľkosťou a pridruženým skóre, ktoré používa algoritmus.
- Vizualizačné grafy: grafy založené na genotypoch SNP (napr. frekvencia B-alely (BAF) a signál homozygotnosti/ROH), ktoré pomáhajú potvrdiť prítomnosť dlhých homozygotných oblastí a hraníc.

Informácie o kontrole kvality: metriky týkajúce sa spoľahlivosti detekcie (napr. hustota SNP / počet SNPs) a varovania, ak údaje nie sú dostatočné na spoľahlivé volanie.

Detegované udalosti UPD sa tiež zobrazujú v Large Variant Viewer (LVV) so špecifickými štítkami (napr. UPD | Iso a UPD | Het).

Skóre UPD

Skóre UPD je numerický indikátor používaný na kvantifikáciu toho, ako silne vzorka vykazuje signál podobný UPD pozdĺž genómu. Detekcia UPD v SeqOne sa spolieha na vzory genotypu SNP:

- V singletonovej (sólo) analýze algoritmus hľadá veľké oblasti homozygotnosti (ROH) a náhle zmeny v distribúcii SNPs pozdĺž chromozómov.
- V analýze tria sa vzor SNP navyše používa na určenie rodičovského pôvodu udalosti a na rozlíšenie izodizómie (UPD | Iso) od heterodizómie (UPD | Het).

V praxi skóre UPD odráža silu a konzistenciu signálu podporujúceho udalosť UPD (napr. jasný vzorec ROH/BAF a dobre ohraničené hranice). Vyššie skóre naznačuje, že udalosť je s väčšou pravdepodobnosťou skutočnou udalosťou UPD, zatiaľ čo nižšie hodnoty naznačujú slabšie alebo hlučnejšie signály.

Skóre UPD sa musí vždy interpretovať spolu s prvkami poskytnutými na karte Správa UPD (sekcia QC), vrátane:

- vizualizačných grafov UPD (napr. frekvencia B-alely / signál homozygotnosti),
- veľkosti a chromozomálneho umiestnenia udalosti,
- metrík kvality analýzy (napr. hustota SNP / počet SNPs).

Graf frekvencie B-alely (BAF) zo správy UPD

Graf frekvencie B-alely (BAF) je vizualizácia genotypov SNP pozdĺž genómu a používa sa na hodnotenie alelickej nerovnováhy a straty heterozygotnosti (LOH).

Pre každú pozíciu SNP zodpovedá BAF podielu sekvenčných čítaní podporujúcich alelu "B" a pohybuje sa od 0 do 1:

- Hodnoty blízke 0 zodpovedajú homozygotnému genotypu (AA).
- Hodnoty blízke 0,5 zodpovedajú heterozygotnému genotypu (AB).
- Hodnoty blízke 1 zodpovedajú homozygotnému genotypu (BB).

V normálnej diploidnej oblasti graf BAF zvyčajne ukazuje tri pásy okolo 0, 0,5 a 1. V dlhých oblastiach homozygotnosti (ROH) alebo izodizómie UPD je heterozygotný pás (okolo 0,5) znížený alebo chýba, čo podporuje prítomnosť signálu podobného UPD.

Dôležité obmedzenia:

- Signál podobný UPD možno tiež pozorovať v kontexte príbuznosti rodičov (viacero oblastí ROH naprieč genómom).
- Volanie UPD je založené predovšetkým na distribúcii SNPs a nemusí samo o sebe stačiť na odlíšenie niektorých vzorov UPD od udalostí súvisiacich s počtom kópií; v prípade potreby vykonajte krížovú kontrolu s inými dostupnými dôkazmi (napr. pokrytie / výsledky CNV).

13. Interpretácia výsledkov analýzy

13.1. Výsledky analýzy

Používatelia majú prístup k analýzám projektu kliknutím na ikonu „Analýzy“ v príslušnom poli alebo kliknutím na ikonu s rovnakým názvom v ľavom hornom rohu používateľského panela.

Keď je analýza dokončená a je k dispozícii na zobrazenie, zobrazí sa so stavom „*Interpretácia – Nová*“.

Výsledky analýzy sú potom k dispozícii vo forme niekoľkých správ, ktoré je možné otvoriť priamo kliknutím na ne.

Existuje niekoľko variácií stavu „Interpretácia“ v závislosti od pokroku dosiahnutého pri interpretácii výsledkov:

- **Nová** – tabuľka variantov ešte nebola zobrazená.
- **Prebieha** – tabuľka variantov bola zobrazená, ale žiadne varianty ešte neboli validované
- **Hodnotenie** – bola validovaná aspoň jedna varianta
- **Overené hlásenie** – bola vygenerovaná správa o analýze

13.1.1. Správa interpretácie analýzy Pracovný postup

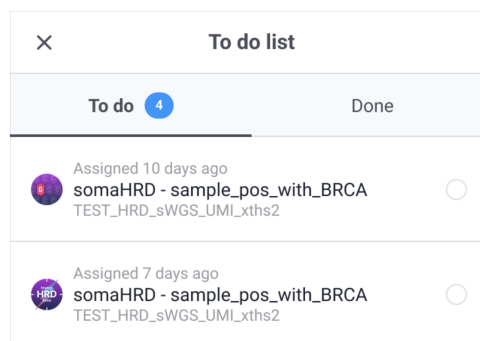
13.1.1.1. Priradovanie analýz na interpretáciu v rámci programu

Správcovia entít môžu na karte Analýzy projektu priradiť jednu alebo viacero analýz, ktoré majú interpretovať rôzni používatelia entít. Na to môže správca vybrať jednu alebo viacero analýz a

potom kliknúť na tlačidlo „ Edit assignment“ (Priradiť používateľom) v dolnej časti stránky.

Okno „Assignment“ (Pridelenie) vám umožňuje priradiť jedného alebo viacerých používateľov k vybraným analýzám vyhľadávaním identifikátora používateľa (užívateľského mena) prostredníctvom špeciálneho vyhľadávacieho poľa. Priradenie je možné tiež upraviť pridaním alebo odstránením používateľov. Používatelia, ktorým je analýza priradená, sú uvedení v stĺpci „Assigned to“ (Priradené k). Používatelia, ktorí sa nikdy neprihlásili do SeqOne, nemôžu byť priradení, kým si nevytvoria svoje prihlasovacie údaje.

Používatelia majú prístup k analýzam, ktoré im boli pridelené (karta „To do“ – Úlohy), a k overeným analýzam (karta „Done“ – Hotové) vo svojom zozname úloh, ktorý je dostupný prostredníctvom ikony  v pravom hornom rohu stránky. Pri každej novej úlohe sa zobrazí oznámenie.

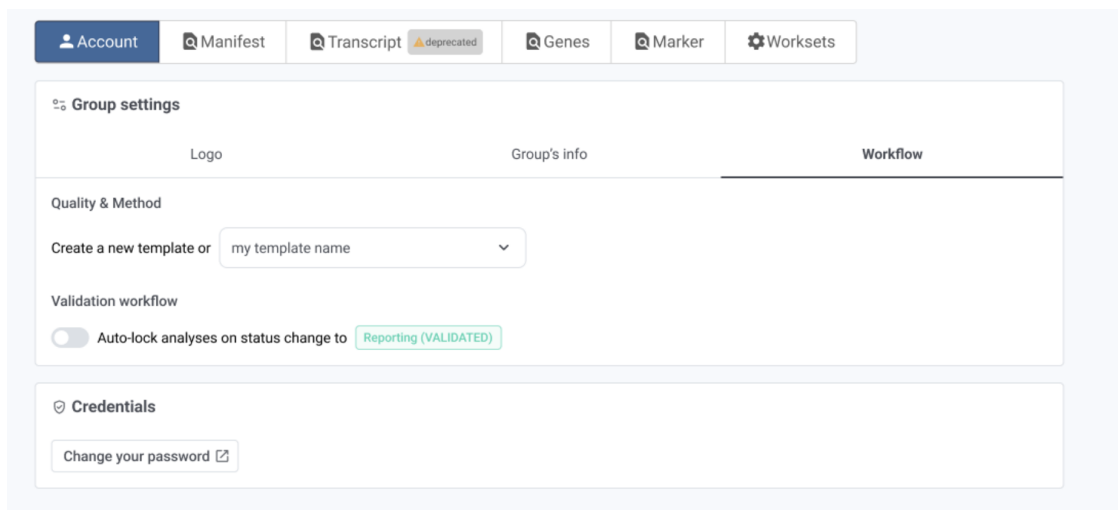


Používatelia majú prístup k prideleným analýzam priamo kliknutím na ne. Svoju úlohu môžu overiť kliknutím na príslušné začiarkavacie políčko v karte „To do“ (Úlohy), čím ju presunú do karty „Done“ (Hotové). Táto akcia je vratná.

Analýzy sa zobrazujú v nasledujúcom poradí podľa ich stavu: dokončené, prebiehajúce alebo neúspešné. Pre každú analýzu sú uvedené podrobnosti o dátume zadania, názve analýzy a názve projektu.

13.1.1.2. Zamknutie analýzy

Správca entity môže analýzu uzamknúť alebo odomknúť. V *nastaveniach* entity, na karte „Účet“ a potom „Pracovný postup“, môže aktivovať automatické uzamknutie validovaných analýz pomocou funkcie „Automaticky uzamknúť analýzy pri zmene stavu na Správa (VALIDOVANÉ)“:



Na konci interpretácie, keď používateľ zmení stav analýzy na „Reporting – validated“ (Správa – overené), analýza sa automaticky uzamkne. Uzamknuté analýzy sa zobrazujú so stavom „Reporting – Validated & Locked“ (Reporting – Validated & Locked), ktorý je viditeľný pomocou ikony  **Reporting** **VALIDATED & LOCKED**.

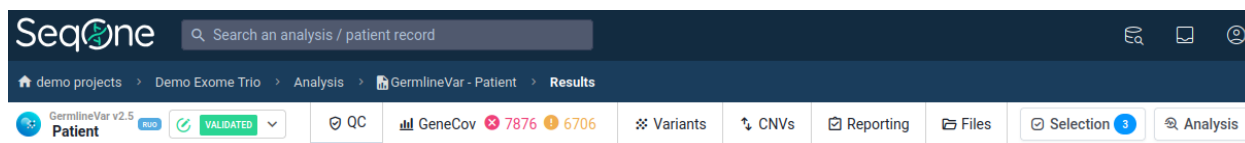
Keď je analýza uzamknutá, používatelia nemôžu z rozhrania vykonávať nasledujúce akcie:

- pridať alebo upraviť variantné vyhodnotenia vo VKB,
- pridávať alebo upravovať komentáre spojené s variantmi,
- vytvoriť novú správu z okna *Výber*,
- vytvoriť novú správu HRD,
- ručne upravovať stav analýzy,
- vykonanie jednej z nasledujúcich akcií v paneli analýzy:
 - upraviť priradenie testu (v časti „Priradenie“)
 - zmeniť pohlavie pacienta
 - pridať alebo zmeniť HPO
 - zmeniť panely analýzy
 - vymazať klinickú správ

Ak sa používateľ pokúsi vykonať niektorú z týchto akcií, bude vyzvaný, aby odomkol analýzu, ak má povolenie, alebo aby kontaktoval svojho správcu, ak povolenie nemá. Používateľ, ktorý chce odomknúť analýzu, musí mať povolenie „Odomknúť analýzu“.

13.1.1.3. Bočný panel pre správu analýz

Bočný panel analýzy, známy aj ako „zásuvka analýzy“, je prístupný z akejkoľvek karty analýzy.



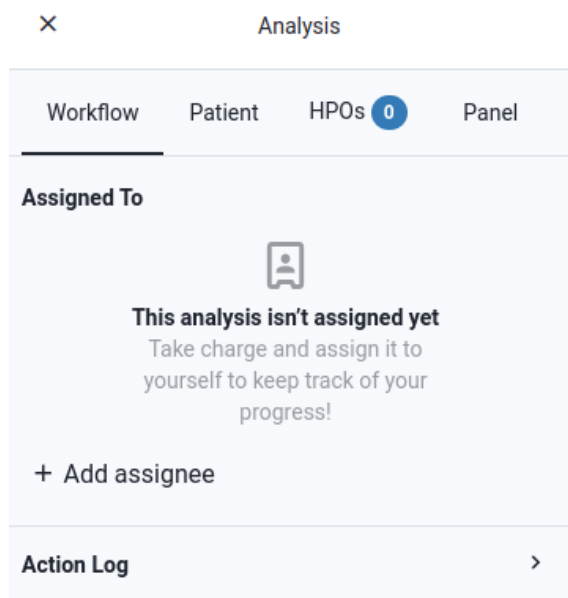
Kliknutím na tlačidlo „ Analysis“ v pravom hornom rohu banneru kariet analýzy môže používateľ priamo zobrazíť a spravovať údaje analýzy a HPO pacientov bez opustenia aktuálnej obrazovky.

Analýza (Analysis Drawer) obsahuje 4 zobrazenia:

- **Pracovný postup:** umožňuje organizovať a sledovať pracovný postup analýzy.
- **Pacient:** umožňuje zobrazíť a upravovať metadáta pacienta
- **HPO:** umožňuje pridávať alebo upravovať termíny HPO v analýze, dostupné len pre analýzy v konštitučnom kontexte.
- **Panel:** umožňuje aplikovať zoznam génov na analýzu a filtrovať výsledky analýzy na základe tohto zoznamu.

Modálny pohľad na pracovný tok

Toto zobrazenie vám umožňuje organizovať a sledovať pracovné postupy používateľov pre interpretovanú analýzu.



Pridelené. Toto zobrazenie zobrazuje úlohy pridelené pre aktuálnu analýzu. Úlohy môžu pridelať iba správcovia alebo používatelia s oprávnením „Spravovať pridelenia“ (pozri časť **Správa správcov – prezentácia rozhrania**).

Keď sú úlohy priradené k aktuálnej analýze, zobrazí sa zoznam všetkých priradených používateľov spolu s dátumom a časom dokončenia. Tento zoznam sa automaticky aktualizuje hneď, ako používateľ dokončí úlohu.

Používatelia môžu upravovať len svoje vlastné úlohy zaškrtnutím alebo odškrtnutím príslušného políčka. Pri prechádzaní kurzorom nad úlohou, ktorá je len na čítanie, sa zobrazí nápoveda

„Môžete dokončiť len svoje úlohy“.

Ak neboli pridelené žiadne úlohy, v zobrazení sa zobrazí „Táto analýza ešte nie je pridelená“.

Ak sa neoprávnený používateľ pokúsi priradiť úlohu kliknutím na tlačidlo „+ Pridať priradenú osobu“, zobrazí sa chybová správa „Na priradenie analýzy musíte mať oprávnenia správcu. V prípade potreby kontaktujte správcu platformy a požiadajte o príslušný prístup.“

Protokol akcií. Banner „Protokol akcií analýzy (Audit Trail)“ zobrazuje históriu všetkých akcií vykonaných v stave HRD v analýze SomaHRD, filtroch, profiloch filtrov a obmedzených génových paneloch v tabuľke variantov, s podrobnými informáciami o type akcie, používateľovi, časovej pečiatke a popise akcie.

Najnovšie akcie sa zobrazujú v hornej časti zoznamu. Pagination umožňuje používateľom efektívne navigovať v histórii akcií. Celú históriu je možné stiahnuť vo formáte .csv prostredníctvom ikony „Stiahnuť protokol“.

Modálne zobrazenie pacienta

Toto zobrazenie vám umožňuje zobraziť a upravovať metadáta pacienta, ako je pohlavie (žena, muž alebo neznáme), genetická skupina predkov GnomAD (len germline) alebo indikácia (len somatická).

Zobrazenie a úprava indikácie je k dispozícii len pre **somatické** analýzy:

- Ak bola indikácia zadaná pred analýzou (pozri **Metadáta vzorky**), bude automaticky definovaná pre vzorku. Následne bude označená hviezdíčkou a zobrazí sa ako prvá v zozname.
- Ak sa indikácia počas interpretácie zmení, varovná správa upozorní, že vybraná indikácia sa líši od pôvodnej indikácie, a farba pilulky v podstĺpci „A“ v stĺpci Klinické dôkazy sa aktualizuje (pozri **Výber indikácie**).
- Ak nie je k dispozícii žiadna indikácia na analýzu, zobrazí sa správa „Žiadna indikácia k dispozícii“.

Zobrazenie a úprava „GnomAD genetic ancestry group“ (Skupina genetických predkov GnomAD) je k dispozícii len pre analýzy **zárodočných buniek**. Predvolená hodnota je prázdna a je považovaná za kombinovanú populáciu GnomAD.

Zmeny vykonané v tomto zobrazení je potrebné uložiť alebo zrušiť kliknutím na príslušné ikony „Uložiť zmeny/Zrušiť“ v dolnej časti okna, aby bolo možné zmeniť karty.

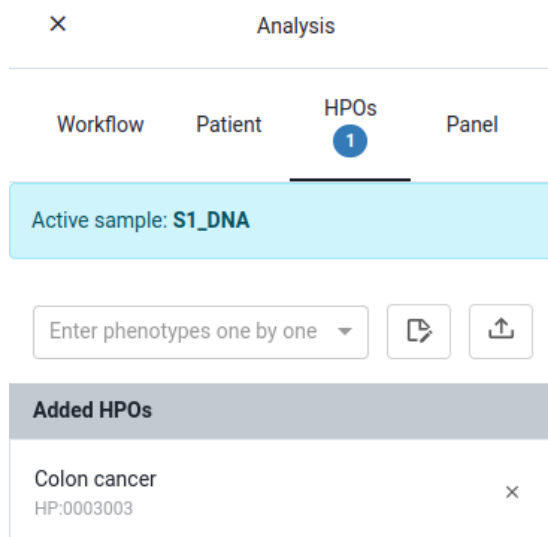
The screenshot shows a modal window titled 'Analysis' with a close button (X) in the top left. It has four tabs: 'Workflow', 'Patient' (selected), 'HPOs', and 'Panel'. Under the 'Patient' tab, there is a section 'Active sample: sample2'. Below this, there is a 'Sex' section with three buttons: 'Female', 'Male', and 'Unknown'. The 'Female' button is highlighted. Below the 'Sex' section is an 'Indication' section with a dropdown menu showing 'Ovarian endometrial cancer (Female reprod...'. Below the dropdown, there is a red warning message: 'Unsaved changes detected. Save now to avoid losing your updates.' At the bottom of the 'Indication' section, there are two buttons: 'Cancel' and 'Save modifications'. Below the 'Indication' section is a section titled 'GnomAD genetic ancestry group' with a magnifying glass icon. Below this title is a list of ancestry groups: 'African/African American', 'Amish', 'Ashkenazi Jewish', 'Admixed American', 'East Asian', and 'European (Finnish)'. The 'African/African American' group is highlighted.

Modálny pohľad HPO

Tento pohľad, ktorý je k dispozícii iba v projektoch Germline, vám umožňuje zobraziť, upraviť a priradiť termíny HPO súvisiace so vzorkou, buď ručne, alebo na základe klinického opisu pacienta.

Používateľ si môže pred analýzou zobraziť zoznam termínov HPO, ktoré sú už spojené so vzorkou. Ak nie sú k dispozícii kódy HPO, zobrazí sa správa „Zatiaľ žiadne HPO“.

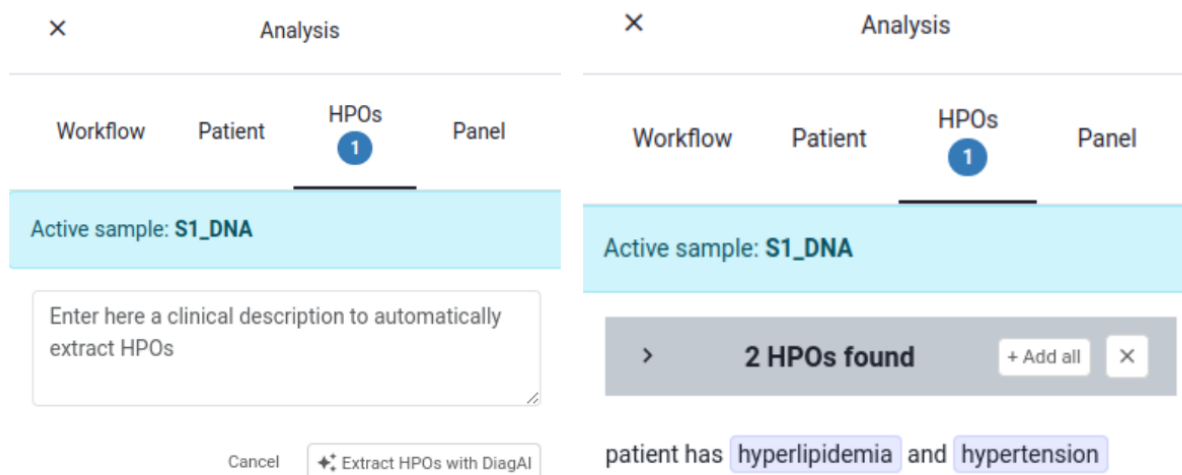
Ak chcete pridať termíny ručne, môžete ich zadať do vyhľadávacieho poľa „Zadajte fenotypy jeden po druhom“ vyhľadávaním fenotypu alebo kódu HPO. Môžete tiež importovať textový súbor obsahujúci kódy HPO kliknutím na ikonu „“ (pozri tiež časť **Metadáta vzorky**).



Obrázok 8: Zobrazenie HPO v paneli analýzy

Používateľ môže pomocou DiagAI generovať termíny HPO na základe klinického opisu pacienta.

Kliknutím na ikonu „“ (Vytvoriť HPO) sa zobrazí dialógové okno, v ktorom sa používateľ vyzve, aby opísal klinický profil pacienta (tento popis je obmedzený na 4000 znakov) a potom klikol na „Extract HPOs with DiagAI“ (Extrahovať HPO s DiagAI), aby sa vygeneroval zoznam nájdených HPO. Termíny HPO sú farebne zvýraznené a používateľ ich môže priradiť k pacientovi tak, že na ne umiestni kurzor myši a klikne na „Add“ (Pridať), alebo môže pridať všetky nájdené termíny HPO kliknutím na „Add all“ (Pridať všetky) (pozri tiež časť **Úprava podľa vzorových metadát s analytickým panelom**).

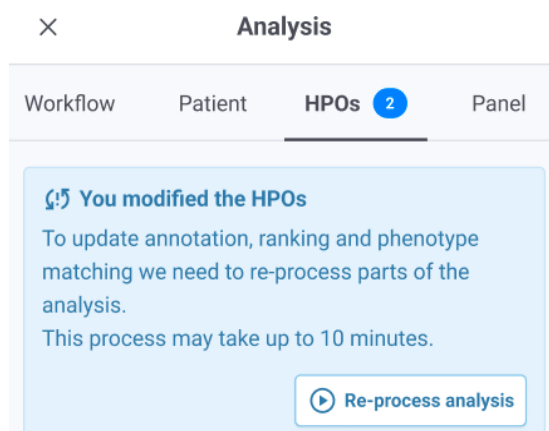


Obrázok 9: Extrahovanie termínov HPO pomocou DiagAI z pohľadu HPO v paneli analýzy

Keď sa HPO pridajú alebo upravia z tohto zobrazenia, zobrazí sa modrý banner, ktorý navrhuje, aby sa analýza spracovala znovu s novými HPO.

Výberom možnosti „Re-process analysis“ (Opätovné spracovanie analýzy) môže používateľ aktualizovať skóre DiagAI a klasifikáciu variantov (Short list alebo SmartPick) podľa nových zadaných HPO.

Tento krok opätovného spracovania na niekoľko minút zmrazí tabuľku variantov, aby sa aktualizovala klasifikácia variantov a stĺpec DiagAI.

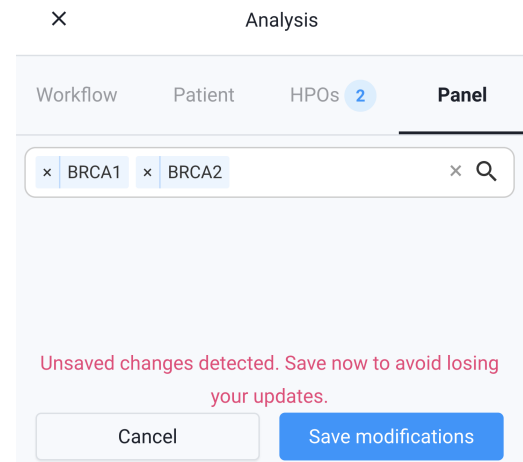


Modálny panel Zobrazenie

Toto zobrazenie je k dispozícii pre všetky pracovné sady okrem pracovných sád analýzy CNV (CNVExome, CNVCapture a SomaCNVCapture).

Poskytuje prístup k funkcii „Obmedzené panely“, ktorá umožňuje obmedziť a uzamknúť analýzu na vybrané gény, podobne ako v prípade *in silico* panelu.

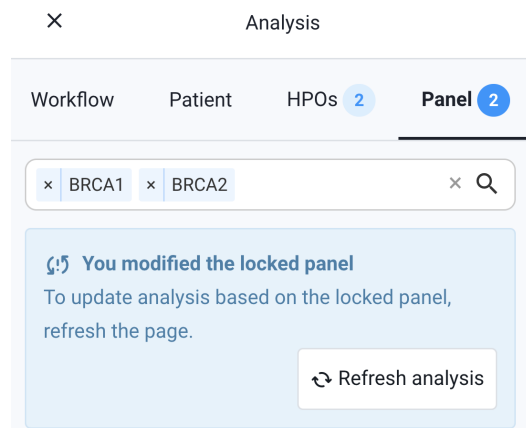
Na tento účel môže používateľ z roletového menu vybrať jeden alebo viacero zoznamov génov, ktoré sa majú použiť, z tých, ktoré boli predtým importované (pozri odsek **Gény** v časti **Správa používateľských parametrov**). Uložením vybraných panelov prostredníctvom funkcie „Uložiť zmeny“ sa na celú analýzu uplatní filter podľa vybraného génového panelu (panelov).



Obmedzené panely je možné vybrať hneď po spustení analýzy (pozri časť **Spustenie novej analýzy**) alebo z pohľadu Panel po analýze.

Keď sa pridajú alebo upravia genové panely, modrý banner upozorní používateľa na zmenu („Upravili ste uzamknutý panel“) a navrhne obnovenie analýzy s novými panelmi prostredníctvom funkcie „Obnoviť analýzu“.

Kliknutím na „Obnoviť analýzu“ sa používateľ rozhodne aktualizovať tabuľku variantov, kartu CNV (ak je to relevantné) a kartu GeneCov podľa nového panelu.



13.1.1.4. Bočný panel na správu hodnotení

Panel „ Selection“ (Analýza variantov) je prístupný z akejkoľvek karty v rámci analýzy. Umožňuje používateľovi zobrazíť malé varianty a CNV anotované vo VKB pre aktuálnu analýzu, vygenerovať súbor JSON („Generate file“) a upraviť klinickú správu („Generate report“).

CNV validované vo VKB sú k dispozícii aj v tejto sekcii, avšak len v prípade CNV z analýz exómu, genómu a CGH Microarray.

Malé varianty validované vo VKB sú odlíšené logom , zatiaľ čo CNV sú zobrazené logom  a slová „zisk“ alebo „strata“.

×

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. Prehľad výsledkov analýzy na karte Dashboard

Pre niektoré pracovné súbory je používateľom k dispozícii karta „Dashboard“, ktorá sumarizuje všetky výsledky a metriky QC a zobrazuje sa štandardne pri otvorení analýzy. Karta „Dashboard“ poskytuje prehľad výsledkov analýzy v viacerých sekciách. Dostupné sekcie sa líšia v závislosti od typu pracovného súboru. Táto karta je k dispozícii pre SomaCGP, SomaVar Tertiary a SomaLBx.

Pacient. Táto časť obsahuje súhrn metadát vzorky: názov vzorky (ID vzorky), pohlavie a indikácia patológie (indikácia).

Kontroly kvality. Táto sekcia poskytuje informácie o niekoľkých kľúčových metrikách týkajúcich sa kvality vzorky a sekvenovania:

- Čistota nádoru,
- Ploidia vzorky,
- percento báz, ktorých kvalita je väčšia alebo rovná 30 (% Q30 báz),
- priemerné pokrytie oblastí záujmu (priemerné pokrytie ROI) pred a po spracovaní UMI,
- percento oblastí záujmu s pokrytím väčším ako 200X (ROI > 200X) pred a po spracovaní UMI.

Signatúry. Časť o signatúrach uvádza kľúčové údaje pre analýzy TMB, MSI a HRD:

- Výpočet TMB (mutácia nádoru): počet mutácií/Mb.
- Mikrosatelitová nestabilita: percentuálny podiel nestabilných miest.
- Stav HRD (Homologous Recombination Deficiency – deficit homologickej rekombinácie).

Stavy skóre TMB a MSI môže používateľ vybrať prostredníctvom roletových menu umiestnených pod každým skóre.

Výber. V tejto časti sa zobrazujú varianty vyhodnotené v databáze VKB pre aktuálnu analýzu.

Navrhované spoločnosťou SeqOne. Genomické zmeny zistené v analýze sú prezentované ako zredukovaný zoznam klinicky relevantných variantov. Varianty sú vybrané, ak ich trieda ComPerMed je 5, alebo ak majú tvrdenie Tier I zodpovedajúce indikácii pacienta, alebo tvrdenie Tier II pre experimentálne terapie súvisiace s touto indikáciou, alebo tvrdenie Tier I pre inú indikáciu. Používateľ môže pridať varianty z tejto kategórie do sekcie Výber kliknutím na „Pridať“.

Súbor JSON je možné vygenerovať kliknutím na „Generovať súbor JSON“ a obsahuje všetky informácie zo stránky „Dashboard“.

13.1.3. Podrobné pokrytie podľa exónu

Na karte *GeneCov* môžu používatelia zobraziť mieru pokrytia génov zahrnutých v manifeste analýzy pre každý exón, či už ide o génový panel alebo exóm.

Metodika

Tento modul pretína manifest používateľa s anotáciami RefSeq pre všetky gény zahrnuté v paneli alebo exóme, pričom zachováva 10 bp na každej strane, aby pokrylo spojovacie miesta, a potom pokračuje vo výpočte hĺbky pokrytia pre každú bázu zahrnutú v tomto pretínaní pomocou nástroja *mosdepth*.

Sú nastavené rôzne prahové hodnoty, aby sa každému exónu priradil stav (pokrytý, varovanie, zlyhanie) podľa percenta báz pod týmito prahovými hodnotami pokrytia.

Stav je definovaný na základe základného pokrytia v exóne vrátane susedných 10 bp. Ak sa používajú UMI, pokrytie sa vypočíta po deduplikácii.

Prahové hodnoty, ktoré sa vzťahujú na rôzne typy analýz, sú uvedené v tabuľke nižšie*:

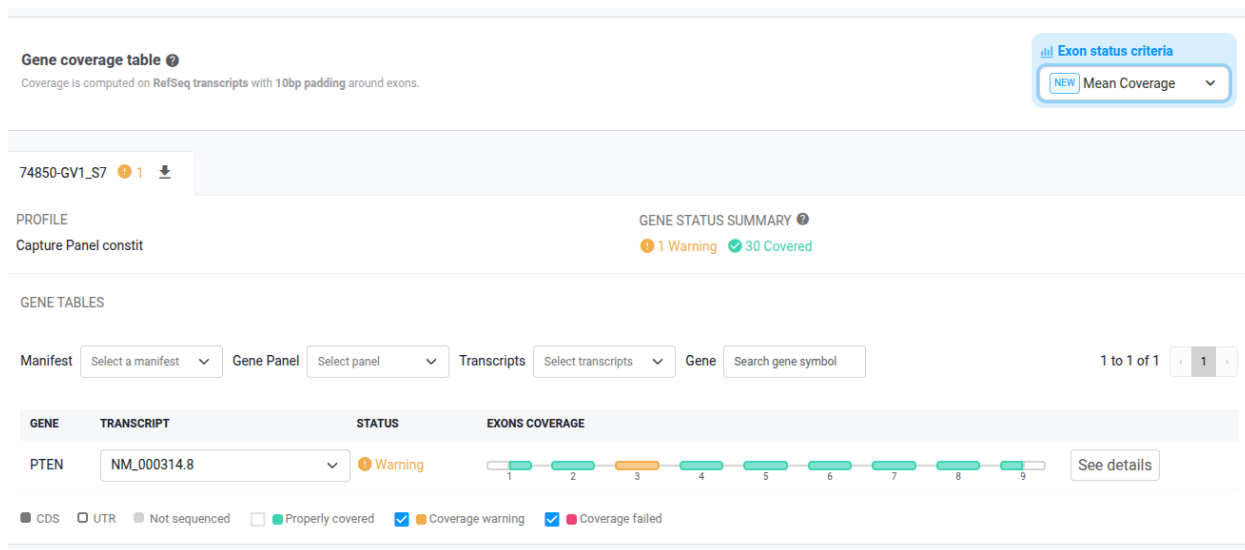
	GermlineVar /Germline Family/GermVar/GermVar Family				SomaVar		SomaRNA		SomaVar LF		SomaCGP	
	Panel		Exóm		Panel		Panel		Panel		Panel	
	Varovanie	Neúspešné	Varovanie	Neúspešné	Varovanie	Neúspešné	Varovanie	Neúspešné	Varovanie	Neúspešné	Varovanie	Neúspešné
Zachytávanie	Pokrytie 100x < 100	Pokrytie 30x < 100	Pokrytie 30x < 100 %	Kryt. 10x < 100 %	Kryt. 500x <100%	Kryt. 200x <100%	Pokrytie 200x <100 %	Kryt. 50x <100%	Pokrytie 5000x <100 %	Pokrytie 2500x <100 %	Pokrytie 100x < 100 %	Pokrytie 50x < 100 %
Amplicon	Pokrytie 100x < 100 %	Pokrytie 30x < 100 %			Pokrytie 1000x < 100 %	Pokrytie 500x < 100 %	Pokrytie 200x <100 %	Pokrytie 50x <100%				

Tabuľka 3: Prahové hodnoty pokrytia pre modul *Gene Coverage* podľa typu analýzy.

*Ak má používateľ vopred definované vlastné metriky kontroly kvality z parametrov projektu, tieto metriky kontroly kvality sa uplatnia na karte *GeneCov* (pozri časť **Konfigurácia predvolených metrík kontroly kvality projektu**).

Zobrazenie a interpretácia

Každý gén je možné zobrazíť na grafickom znázornení a výsledky pokrytia génov je možné filtrovať na základe génov, in silico panelov a zoznamov transkriptov. Cieľové exóny sú znázornené buď zelenou farbou (*pokryté*), oranžovou farbou (*varovanie*) alebo červenou farbou (*neúspešné*) podľa vyššie opísaných prahových hodnôt.



Obrázok 10: Podrobnosti o karte GeneCov

Možnosti filtrovania

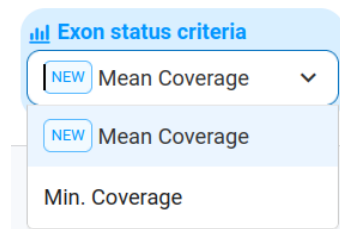
Ponuka filtrov vám umožňuje:

- obmedziť zobrazenie na podmnožinu regiónov pomocou súboru .bed, ktorý bol predtým importovaný do rozhrania, prostredníctvom filtra „Manifest“,
- použiť zoznam génov pomocou filtra „Gene Panel“ (Panel génov),
- použiť zoznam transkriptov pomocou filtra „Transcripts“ (Transkripty),
- vyhľadávať konkrétny gén pomocou filtra „Gene“.

Metódy výpočtu pokrytia

Používateľ môže vybrať metódu výpočtu pokrytia, ktorá sa má použiť, pomocou roletového menu „Exon status criteria“ v pravom hornom rohu stránky:

1. **Priemerné pokrytie** (predvolené): Na určenie stavu pokrytia sa používa priemerná hĺbka pokrytia v celom exóne.
2. **Minimálne pokrytie**: Na určenie stavu pokrytia sa používa minimálny podiel báz, ktoré spĺňajú špecifickú prahovú hodnotu pokrytia. Tým sa zaručuje, že určité percento báz má aspoň minimálnu požadovanú hĺbku pokrytia.



Rozbaľovacia ponuka **"Gene status criteria"** umožňuje používateľom použiť tieto metódy výpočtu na úrovni génov:

- **Mean coverage (default):** Na určenie stavu pokrytia používa priemernú hĺbku pokrytia u všetkých exónov génu.
- **Worst exon status:** Určuje stav pokrytia génu na základe najhoršie fungujúceho exónu. Ak ktorýkoľvek exón nesplní prah pokrytia, gén sa považuje za nepokrytý. V predchádzajúcich verziách to bola jediná podporovaná metóda.

Zobrazené údaje

Exóny, ktoré nie sú alebo sú čiastočne pokryté dizajnom (menej ako 80 % prekryvania medzi manifestným intervalom a hranicami exónu RefSeq), sa považujú za necieľové. Tieto exóny sú v grafe znázornené sivou farbou (nesekvenované).

Podrobný pohľad na každý gén možno použiť na zobrazenie metrík pokrytia exónov kliknutím na „Zobraziť podrobnosti“: priemerné pokrytie, medián, percento každej oblasti pokrytej 10x, 30x, 50x, 100x, 200x, 300x a 500x.

Pre každý exón umožňuje okno IGV používateľom zobrazíť pokrytie sekvencie a zarovnanie podrobnejšie.

Export údajov GeneCov

Kompletné údaje analýzy pokrytia sú k dispozícii ako súbor `sample_genecov.v2.tsv.gz`, prístupný na karte Súborné analýzy.

Na karte GeneCov môže používateľ exportovať analýzu pokrytia podľa metódy výpočtu pokrytia (priemerné pokrytie alebo minimálne pokrytie)

pomocou ikony „“ vedľa názvu vzorky v ľavom hornom rohu karty vzorky.

Download genes from S1_DNA as .tsv file



Genes status

☒ Failed ☒ Warnings ☐ Covered

Transcripts

☒ Canonical ☐ All

Close

Download

Používateľ môže zvoliť export údajov podľa stavu pokrytia génu (neúspešné, varovanie, pokryté) a exportovať iba kanonický transkript (kanonický) alebo všetky dostupné transkripty (všetky) pred kliknutím na Stiahnuť, aby získal požadovaný súbor.

Dáta sa exportujú ako súbor .tsv obsahujúci nasledujúce stĺpce:

- Gén (`gene_symbol`),

- Číslo génu podľa HGNC (gene_id),
- Kanonický transkript (is_canonical_transcript),
- Stav transkriptu (transcript_status)
- Exón (exon_number)
- Typ exónu CDS alebo UTR (exon_type)
- Stav exónu – pokrytý alebo nepokrytý (exon_status)
- Oblasť pokrytá manifestom alebo nie (target)
- Genómová pozícia (interval_id)
- Číslo chromozómu (chr)
- Počiatočná pozícia (start)
- Konečná pozícia (end)
- Predný alebo zadný reťazec (strand)
- % pokrytie pri 1, 10, 20, 30, 50, 100, 200, 300, 500 a 1000X (cov_)
- Priemerné pokrytie (cov_avg)
- Stredné pokrytie (cov_med)

13.1.4. Malé varianty

Metodika

Detekcia malých variantov sa vykonáva kombináciou viacerých nástrojov, ktoré sa môžu líšiť v závislosti od pracovnej sady.

Či už ide o analýzu exómu alebo génového panelu, identifikácia variantov je určená manifestom (vo formáte BED) spojeným s projektom, vrátane intervalov oblastí záujmu a vypíňania aplikovaného na tieto oblasti (pozri časť **Vytvorenie projektu**).

Okrem toho sa varianty volajú s ohľadom na kvalitu zarovnania čítaní na ľudskom genóme (verzia GRCh37 alebo GRCh38). Takzvané čítania „MAPQ0“ s kvalitou zarovnania (mapovania) rovnou 0, pretože sa zarovnávajú na viacerých miestach v genóme, sa nezohľadňujú pri volaní variantov pre pracovné súbory somatickej analýzy. V analýzach „Germline“ sa čítania MAPQ0 zohľadňujú pri volaní variantov, ak sa pri vytváraní projektu zadá manifest špecifikujúci oblasti zarovnania s nízkou kvalitou v metadátach oblasti mapovania s nízkou kvalitou. Varianty detegované z čítaní MAPQ0 sú v tabuľke variantov identifikované pomocou popisu „_mapq0“ viditeľného v stĺpci CAL (Caller).

SNVs / indels. Na detekciu SNV a indelov:

- *Freebayes*¹ sa používa ako jediný volajúci variantov v pracovných súboroch GermlineVar (exóm), GermlineFamily (exóm), SomaRNA a SomaVar LF. Nižšie prahové hodnoty detekcie sú nasledovné:
 - VAF $\geq 10\%$ v GermlineVar pre exóm,
 - VAF $\geq 5\%$ v GermlineVar pre génové panely,
 - VAF $\geq 1\%$ v SomaRNA,
 - VAF $\geq 0,1\%$ v SomaVar LF,
- Je doplnený *GATK HaplotypeCaller*² v pracovnej sade GermlineVar (panel).
- Jeho somatický ekvivalent, *Mutect2*³, sa používa spolu s Freebayes v pracovných súboroch SomaVar, SomaHemato a SomaCGP. Minimálna detekčná prahová hodnota je VAF $\geq 1\%$ pre tieto dva volajúce programy v SomaVar.
- *Sentieon*⁴ sa používa ako hlavný variant caller v pracovných súboroch GermVar, GermVar Family a SomaLBx.
- *Clair3*⁵ sa používa ako jediný variant caller v pracovnej sade GermlineVarLR.

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 sa spoliehajú na bioinformatické komponenty Sentieon pri kľúčových krokoch sekundárnej analýzy. Najmä:

- Zarovnanie sa vykonáva pomocou optimalizovanej implementácie BWA-MEM2 od spoločnosti Sentieon a duplicitné čítania sú označené.
- Zárodočné volanie (SNVs/Indels) sa vykonáva pomocou Sentieon DNAScope⁶.
- Sentieon TNScope⁷ možno použiť na detekciu mozaikových variantov s nízkym VAF (v závislosti od konfigurácie GermVar).

Poznámka k regiónom s nízkou mapovateľnosťou (MAPQ0): vo vysoko homológnych regiónoch (napr. SMN1/SMN2) môžu byť varianty odfiltrované filtrovaním strojového učenia vyvolávača. V prípade potreby poskytuje SeqOne voliteľný modul MAPQ0 (modul regiónov s nízkou kvalitou) založený na Freebayes s prispôbenými prahmi na zvýšenie citlivosti v používateľom definovaných regiónoch.

Analýzy SomaLBx sa spoliehajú na bioinformatické komponenty Sentieon pri kľúčových krokoch sekundárnej analýzy tekutej biopsie. Najmä:

- Zarovnanie čítania sa vykonáva pomocou Sentieon (BWA-MEM) a pipeline zahŕňa spracovanie/konsenzus UMI založené na Sentieon na podporu vysoko citlivých pracovných postupov ctDNA.
- Somatické volanie krátkych variantov (SNVs/Indels) sa vykonáva pomocou Sentieon TNScope⁷ v režime iba pre nádor (tumor-only), s podporou veľmi nízkej alelovej frakcie (VAF $\geq 0,1$ % v závislosti od konfigurácie).
- Poznámka: genotypy variantov sú vo výsledkoch SNV/indel SomaLBx štandardne hlásené ako heterozygotné, pretože Sentieon TNScope nevykonáva predikciu genotypu.

SVs - Structural Variants. Vloženia a vynechania, ktorých dĺžka je blízka alebo väčšia ako dĺžka sekvenčného čítania, sa detekujú pomocou identifikátora SV GRIDSS⁸ v pracovných súboroch GermlineVar (panel/exóm), SomaVar, SomaVar LF SomaHemato a SomaCGP. Tieto udalosti sa detekujú pomocou identifikátora SV Delly⁹ v pracovných súboroch GermVar a GermVar Family. V somatických pracovných súboroch sa detekujú iba vloženia a vynechania menšie ako 300 bp, zatiaľ čo v pracovnom súbore GermlineVar (pre verzie ≥ 2.0), GermVar a GermVar Family (verzie ≥ 3.5) sa uplatňuje limit 10 kb. Varianty detekované nástrojom GRIDSS sa detekujú bez ohľadu na VAF.

FLT3-ITD. Špecifický detekčný modul pre interné tandemové duplikácie génu FLT3 (FLT3-ITD) je k dispozícii iba vo worksets SomaVar a SomaHemato, od verzie v2.2. Tento modul je založený 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. Nakoniec, spoločnosť SeqOne vyvinula aj modul na detekciu inzercíí transpozitelných elementov (*AluMEI*), ako sú Alu elementy. Používa sa v pracovných súboroch GermlineVar (panel/exóm), GermlineFamily (panel/exóm), GermVar (panel/exóm/genóm), GermVar Family (panel/exóm/genóm), SomaHemato a SomaVar. Na varianty detegované nástrojom AluMEI sa neuplatňuje žiadna nižšia prahová hodnota VAF.

Kombinované varianty *Freebayes*, *HaplotypeCaller* (alebo *Mutect2*) a *FiLT3r* caller sú k dispozícii v konečnom súbore analýzy *final.vcf.gz*. Varianty špecifické pre *GRIDSS*, *FiLT3r* alebo *AluMEI* sú k dispozícii v príslušných súboroch VCF (**gridss.vcf.gz**, **filt3r.vcf.gz** a **variants.alumei.vcf.gz**), ktoré sú tiež k dispozícii na stiahnutie na karte *Súbory analýzy*.

V konkrétnom prípade *FiLT3r* :

- Varianty *FiLT3r*, ktoré sú spoločné pre *Freebayes* a *Mutect2*, sú k dispozícii v konečnom VCF analýzy (**final.vcf.gz**),
- varianty jedinečné pre *FiLT3r*, ako aj varianty spoločné pre *FiLT3r* a *GRIDSS* sú uvedené vo *FiLT3r* VCF (**filt3r.vcf.gz**).

Dva ďalšie moduly vyvinuté spoločnosťou SeqOne potom umožňujú:

- Zachovanie multinukleotidových variantov (MNV), ktoré kódujú a sú obmedzené na jediný kodón počas procesu rozkladu variantov podobných delínom.
- Rekalibrácia alelickej frekvencie variantov indelov a delinov detegovaných nástrojom *Freebayes* a *GATK* (nástroj *Kraf*).

Tieto dva moduly sa týkajú iba pracovných súborov GermlineVar (panel/exóm), SomaVar, SomaRNA a SomaVar LF.

Kraf používa stratégiu jediného k-meru bez zarovnania. Rekalibrovaná alelová frekvencia indelov a delinov sa zachováva len vtedy, ak je vyššia ako počiatočná alelová frekvencia variantu.

Varianty sú potom hlavne anotované pomocou nástroja *Variant Effect Predictor*¹¹ (VEP) a doplňujúce informácie sú pridávané z nasledujúcich databáz:

- ClinVar
- OMIM (komerčná verzia) + PanelApp Australia + PanelApp England
- CKB
- dbSNP
- COSMIC (komerčná verzia)
- GnomAD (exómové + genómové údaje)
- Dfam
- DGV
- LOVD

Okrem toho sú k dispozícii nasledujúce *in silico* skóre:

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

- Skóre spliceovania (*MaxEntScan*, *GeneSplicer*, *dbscSNV*)
- Skóre konzervácie (*GERP++*, *PhastCons100way*, *PhastCons20way*, *PhyloP100way*, *Vertebrate*, *PhyloP20way Mammalian*, *SiPhy*)
- Predikcie funkčnosti (*REVEL*, *DANN*, *FATHMM*, *FATHMM-MKL*, *LRT*, *MetaLR*, *MetaSVM*, *MutationAssessor*, *MutationTaster 2*, *PROVEAN*, *SIFT*, *PLI*)

Vizualizácia a interpretácia

Karta Varianty  vám umožňuje zobraziť, filtrovať a anotovať varianty získané počas analýzy. Tieto sú zobrazené v tabuľkovej forme, pričom každý riadok zodpovedá inému variantu. Počet variantov vybraných po filtrovaní (vybrané varianty) a celkový počet detegovaných variantov (celkový počet variantov) sa zobrazujú v pravom hornom rohu záhlavia tabuľky variantov pomocou nasledujúcej ikony:  9/22 . Paginácia v dolnej časti stránky vám umožňuje prechádzať z jednej stránky výsledkov na ďalšiu. Ak analýza používa obmedzený panel, hlavička tabuľky bude obsahovať aj celkový počet variantov v obmedzenom paneli (celkový počet variantov v obmedzenom paneli) .

Táto navigácia je uľahčená klávesovými skratkami pomocou klávesov ← a → na klávesnici.

V tabuľke sa štandardne zobrazuje niekoľko popisných stĺpcov. Používatelia môžu tabuľku exportovať ako súbor TSV kliknutím na  .

Pre každú variantu môže byť k dispozícii niekoľko transkriptov a ich počet sa zobrazuje vľavo od každého riadka. Varianty sú anotované na transkriptoch RefSeq, s výnimkou mitochondriálnych variantov, ktoré sú anotované na *Ensembl*.

Transkripty MANE Select a MANE Clinical Plus sú označené ikonou,

resp.  alebo  , spolu s popisom.

TRANSCRIPT	HEVSC
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

Pre rovnakú variantu, s výnimkou mitochondriálnych variantov, sú transkripty zoradené podľa nasledujúceho poradia:

- vplyv ['vysoký' > 'stredný' > 'nízky' > 'pozmeňujúci']
- typ transkriptov [„NM“, „XM“, „NR“, „XR“] (pre transkripty RefSeq)
- typ transkriptu MANE [Mane Select > MANE Clinical Plus > non MANE]
- alfanumerické poradie

V prípade mitochondriálnych variantov sú transkripty zoradené v nasledujúcom poradí:

- vplyv ['vysoký' > 'stredný' > 'nízky' > 'pozmeňujúci']
- ak je vplyv „pozmeňujúci“:
 - efekt „non_coding_transcript_exon_variant“
 - biotyp [kódovanie proteínu > Mt_rRNA > Mt_tRNA]
- alfanumerické poradie

DiagAI. Tento stĺpec zobrazuje skóre od 0 do 100 na základe súboru DiagAI s využitím strojového učenia. Skóre DiagAI je založené na štyroch komponentoch:

- Algoritmus UP², ktorý predpovedá patogenitu variantu
- Matica PhenoGenius, ktorá hodnotí zhodu medzi génom a HPO pacienta (ak boli zadane počas analýzy)
- Súlad medzi pozorovaným genotypom a spôsobom prenosu ochorenia opísaným v databázach GenCC, OMIM a PanelApp (Austrália a England).
- Kvalita variantu

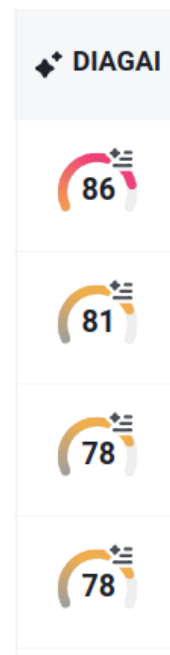
Čím je skóre bližšie k 100, tým viac sa predpokladá, že varianta je patogénna a má diagnostický význam. Čím je bližšie k 0, tým viac sa predpokladá, že varianta je benígna.

Dve ikony tiež upozorňujú na určité zaujímavé varianty:

- Ikona „“ znamená, že varianta má vyššiu pravdepodobnosť, že je príčinou ochorenia (AI Smart Pick).
- Ikona „“ označuje, že variant je prítomný v krátkom zozname variantov s vysokou pravdepodobnosťou, že je príčinou ochorenia pacienta.

Termíny HPO je možné zadať pred spustením analýzy úpravou metadát vzorky (pozri časť **Metadáta vzorky**). Je možné ich upraviť aj počas interpretácie pomocou modálneho zobrazenia HPO v bočnom paneli analýzy (pozri časť **Modálne zobrazenie HPO**). V tomto prípade môže používateľ opätovne spracovať skóre DiagAI a klasifikáciu variantov (krátky zoznam alebo SmartPick) podľa nových zadaných HPO pomocou funkcie „Opätovné spracovanie s novými HPO“.

Ak používateľ klikne na ikonu skóre DiagAI, stránka variantov sa otvorí priamo v karte DiagAI s podrobným popisom skóre (pozri časť DiagAI na stránke Variantov nižšie).

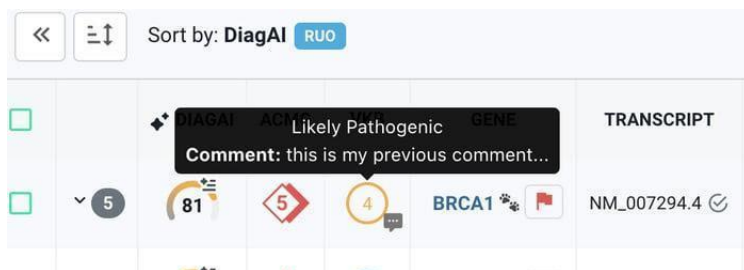


ACMG. Klasifikácia patogenity každej varianty sa vypočítava ako súčasť poloautomatického procesu podľa odporúčaní vydaných American college of Medical Genetics and Genomics. Používatelia môžu klasifikáciu revidovať v špeciálnej záložke na stránke varianty, ktorá sa potom použije na vygenerovanie konečného hodnotenia varianty.

ComPerMed. Klasifikácia patogenity každej varianty sa vypočíta ako súčasť poloautomatického procesu podľa odporúčaní vydaných belgickým odborným výborom ComPerMed (Komisia pre personalizovanú medicínu). Používatelia majú prístup k podrobným informáciám o klasifikácii v špeciálnej karte na stránke varianty.

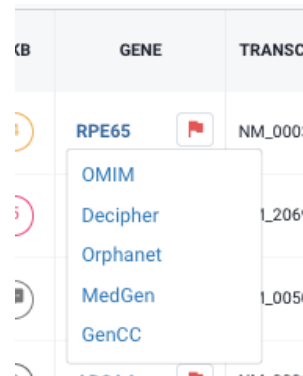
AMP. (Analýza SomaVar verzia v1.8 alebo novšia, iba SomaVar LF alebo SomaCGP). Tento stĺpec sa zobrazuje štandardne v analýzach, ktoré obsahujú anotácie CKB, a zobrazuje úroveň AMP každej varianty, ak je k dispozícii (úroveň I alebo II). Zobrazená úroveň AMP je založená na najlepších klinických dôkazoch vo všetkých typoch, bez ohľadu na vybranú indikáciu.

VKB. Hodnotenia vykonané používateľom počas predchádzajúcich analýz sa ukladajú do znalostnej bázy používateľa. Toto hodnotenie sa zobrazuje v stĺpci VKB tabuľky pre každý nový výskyt variantu. Nástrojová tipka rýchlo zobrazí triedu a komentár súvisiaci s hodnoteným alebo predtým hodnoteným variantom.



Gene (Gén). Názov génu podľa [HGNC](#) a súvisiaca vlajka OMIM:

- Červená vlajka: „Gén OMIM s fenotypmi a klinickými charakteristikami“,
- Červená vlajka s obrysom: „OMIM gén s fenotypmi, ale bez klinických charakteristík“,
- Červená bodka: „Gén OMIM bez fenotypov alebo klinických znakov“.



Gény, o ktorých je známe, že sú imprintingovými génmi podľa PanelApp (Austrália), sú zvýraznené ikonou .

Súvisiaca nápoveda uvádza, či je gén exprimovaný materinsky, otcovsky, či je neznámy, alebo či sa uvádza viacero imprintingov.

Kliknutím na vlajočku vedľa názvu génu sa dostanete priamo na stránku OMIM. Kliknutím na

názov génu sa zobrazí roletové menu s odkazmi na databázy OMIM¹², Decipher¹³, Orphanet¹⁴, MedGen¹⁵ a GenCC¹⁶. Každý odkaz vedie priamo na stránku zodpovedajúcu vybranému génu.

Informácie týkajúce sa OMIM a databáz sú k dispozícii iba pre analýzy zárodočných buniek.

HGVSC. Tento stĺpec poskytuje HGVS anotáciu variantu na komplementárnej DNA (cDNA) a na proteíne.

HGVSC

c.10580G>A
p.R3527Q

Effect (Vplyv). Tento stĺpec opisuje vplyv variantu na proteín a vplyv na transkript.

Vplyv sa počíta podľa 4 úrovní:

- **High** (vysoký vplyv, červená farba): varianta má vysoký rušivý vplyv na proteín, pravdepodobne spôsobuje skrátenie proteínu, stratu funkcie alebo spúšťa jeho rozpad spôsobený nezmyselnou sekvenciou.
- **Moderate** (stredný vplyv, žltá farba): nerušivá varianta, ktorá pravdepodobne modifikuje účinnosť proteínu.
- **Low** (nízky vplyv, sivá farba): varianta, ktorá sa považuje za neškodnú alebo je nepravdepodobné, že by ovplyvnila správanie proteínu.
- **Modifier** (modifikujúci vplyv, modrá farba): všeobecne nekódujúce varianty alebo varianty ovplyvňujúce nekódujúce oblasti alebo gény, pre ktoré je ťažké predpovedať vplyv alebo neexistujú dôkazy o vplyve.

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

Pre každú variantu môže byť v stĺpci Vplyv k dispozícii niekoľko vplyvov na transkript. V tomto prípade sa zobrazí len vplyv s najvyšším dôsledkom podľa VEP¹⁷, za ktorým nasleduje číslo (napr. +2) pre ostatné dostupné vplyvy. Podrobnosti o nich sú uvedené v príslušnej nápovede.

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 (Prenos). Tento stĺpec označuje spôsob prenosu chorôb spojených s génmi uvedenými v databázach GenCC a OMIM. Farebné značky slúžia na rýchlu identifikáciu vzorov prenosu na základe predpovedí DiagAI:

- „De Novo“
- „Compound Het – Verified“ (overený)
- „Recessive Hom“
- „Compound Het – Suspected“ (podozrivý)

Značky De Novo, Zložená Het – overená a Recessívna Hom sa zobrazujú iba v rodinných analýzach.

Pomocou filtra je možné vybrať varianty s týmito značkami (pozri **Filtrovanie variantov** – sekcia Prenos a Fenotypy).

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

Transmission
AR, AD

Leber congenital amaurosis 9
OMIM:7638289

AD, AR | OMIM
AD | 3 2 GenCC
AR | 3 GenCC

Leber congenital amaurosis 2
OMIM:0019200

AD, AR | OMIM
AR | 3 GenCC

Leber congenital amaurosis 5
MONDO:739002

AR | 3 GenCC

Retinis pigmentosa
OMIM:10939229

AD, AR | OMIM

Definitive to Moderate
Supportive
Limited to Unknown

Kliknutím na odkazy režimov prenosu v tomto stĺpci sa otvorí okno, v ktorom používatelia nájdu choroby uvedené s ich názvami a odkazmi na databázy (OMIM, MONDO – ak OMIM nie je k dispozícii – alebo GenCC).

Choroby sú usporiadané v špecifickej hierarchii: najskôr sa zobrazia choroby, ktoré sa nachádzajú v OMIM aj GenCC, potom nasledujú choroby, ktoré sa nachádzajú len v OMIM, a nakoniec choroby, ktoré sa nachádzajú len v GenCC.

Každý záznam obsahuje farebne označené značky: značky OMIM majú zelené okraje, zatiaľ čo značky GenCC používajú systém semaforu (zelená/žltá/červená) na označenie sily dôkazov:

- Zelená: Definitívna + Silná + Stredná
- Žltá: Podporná
- Červená: Obmedzená, Zvieracie, Neznáma

Clinical Evidences (Klinické dôkazy). (Analýza SomaVar v1.8 alebo novšia, iba SomaVar LF alebo SomaCGP). Tento stĺpec sa zobrazuje štandardne v analýzach obsahujúcich anotácie CKB. Je rozdelený do troch podstĺpcov, z ktorých každý obsahuje celkový počet daného typu dôkazov: A (akčnosť), D (diagnóza) a P (prognóza). Stĺpec A zobrazuje aj počet citlivých (vľavo) a rezistentných (vpravo) terapeutických tvrdení ako odznak.

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

Hodnoty v týchto podstĺpcoch sú farebne označené:

- Rezistencia, ak je zistená (> 0 tvrdení), sa vždy zobrazuje červenou farbou bez ohľadu na indikáciu, v ostatných prípadoch šedou farbou.
- Citlivosť je zvýraznená zelenou farbou len v prípade, ak je vybraná príslušná indikácia, inak je zobrazená šedou farbou.

Ak aspoň jeden dôkaz tohto typu zodpovedá indikácii vybranej v tabuľke variantov (panel Indikácia), v popise pred označením indikácie sa zobrazí začiar knutie. Ak sa nezistí žiadna zhoda, označenie sa nezobrazí.

Phenotype (Fenotypy). Tento stĺpec, dostupný pre projekty Germline, umožňuje asociáciu génu s termínmi HPO pomocou matice PhenoGenius 17. Ak sú termíny HPO asociované so vzorkou, zobrazí sa najvyššie percento zhody medzi génom a fenotypom spolu s pomerom termínov, s ktorými je gén asociovaný. Ak žiadne gény nezodpovedajú zadaným HPO, stĺpec vráti správu „no match“ (žiadna zhoda); ak vzorka nemá žiadne termíny HPO, stĺpec zobrazí správu „see HPO“ (pozri HPO). V oboch prípadoch sa však v príslušnom zobrazovacom okne navrhnú známe termíny pre gén.

The image shows two screenshots of the Phenogenius interface. The top screenshot displays 'Patient phenotypes matching P3H1' with a list of HPO terms and their matching percentages: Short ribs (48%), Short long bone (47%), Lambdoidal craniosynostosis (0%), Vitreous floaters (0%), and Hypoplastic vertebral bodies (0%). The bottom screenshot shows 'Known phenotypes described for SLC25A22' with a list of HPO terms: Epileptic encephalopathy, Encephalopathy, EEG with burst suppression, Focal myoclonic seizure, and EEG with spike-wave complexes. The interface includes a search bar, a list of results, and a 'Phenotypes' column on the right.

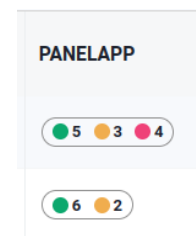
Phenotype	Match Percentage	Ratio
HP:0000773 Short ribs	48%	2/5
HP:0003026 Short long bone	47%	2/5
HP:0004443 Lambdoidal craniosynostosis	0%	no match
HP:0100832 Vitreous floaters	0%	no match
HP:0008479 Hypoplastic vertebral bodies	0%	no match

Phenotype	Match Percentage	Ratio
Abnormal bone collagen fibril morphology	24%	1/5
Abnormal cartilage collagen	no match	no match
Type 1 collagen overmodification	no match	no match
Decreased calvarial ossification	no match	no match
Femoral bowing	no match	no match

Obrázok 11: Zobrazenie informácií v Phenogenius

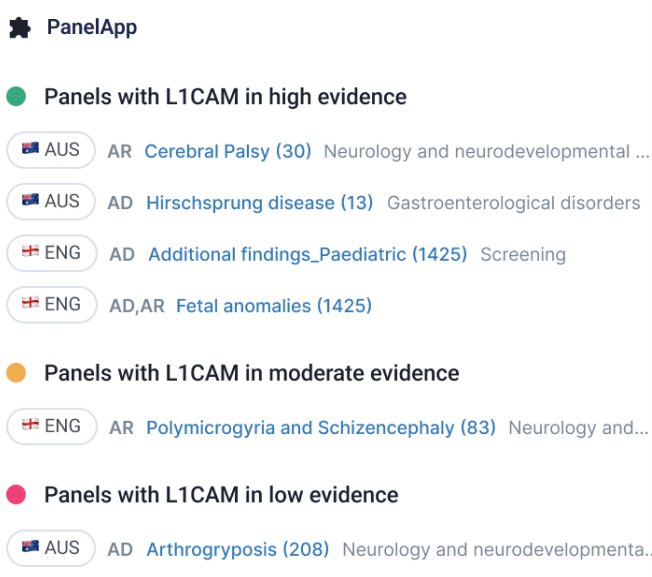
PanelApp. Tento stĺpec integruje anotácie z panelovej databázy *PanelApp Australia*¹⁸ a *PanelApp England*¹⁹. Uvádza počet panelov spojených s génom, pre ktorý bola zistená varianta, podľa klasifikačnej úrovne opísanej nižšie.

Ak sa gén v paneli nenachádza, v stĺpci sa zobrazí „N/A“.



V súvisiacom vyskakovacom okne sa zobrazia podrobnosti o asociácii génu/fenotypu, ako aj spôsob prenosu spojený s fenotypom, spôsob prenosu spojený s fenotypom a databáza, z ktorej je panel odvodený, PanelApp Australia (**AUS**) alebo PanelApp England (**ENG**) (**obrázok 12**).

Každému génu bude pridelený jeden alebo viac panelov podľa trojstupňového klasifikačného systému založeného na dôkazoch o asociácii génu a fenotypu: vysoká dôkaznosť je označená **zelenou farbou**, stredná dôkaznosť **žltou farbou** a nízka dôkaznosť **červenou farbou**.



Obrázok 12: Vizualizácia informácií PanelApp.

¹⁸ <https://panelapp-aus.org>

¹⁹ <https://panelapp.genomicsengland.co.uk/>

ClinVar Germline. Tento stĺpec zobrazuje informácie získané z ClinVar počas analýzy. Poskytuje celý rad informácií, vrátane triedy patogenity, podrobností a počtu podaní pre každú triedu ACMG, stavu revízie a preferovanej podmienky/choroby ClinVar. Kliknutím na bunku v tabuľke variantov bude používateľ presmerovaný na príslušnú stránku ClinVar.



Ďalšie informácie nájdete na stránke https://www.ncbi.nlm.nih.gov/clinvar/docs/review_status/

GnomAD. Tento stĺpec zobrazuje 6 prvky z databázy GnomAD pre konkrétnu variantu v populácii: alelovú frekvenciu (AF), počet aliel (AC), počet homozygotov (HOM), počet hemizygótov (HEM), frekvencia hlavnej alely (MAF) a pre projekty GRCh37 verziu GnomAD, pre ktorú sa táto varianta nachádza (VER).

Kliknutím na hodnotu zobrazenú v stĺpci AF sa používateľ presmeruje do databázy GnomAD vo verzii v4 alebo, ak to nie je možné, vo verzii v2. Ak varianta nie je uvedená v databáze, hodnota zobrazená v tomto stĺpci bude N/A.

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

V stĺpci VER sa zobrazí upozornenie, ak je táto varianta prítomná v GnomAD, ale nespĺňa všetky filtre kvality aspoň pre jeden z dvoch súborov údajov (exóm alebo genóm). V tomto prípade sa zobrazené hodnoty prispôbia hodnotám dostupným pre túto variantu a zobrazia sa filtre, ktoré spustili upozornenie, spolu s krátkou správou uvádzajúcou problematický súbor údajov.

Skóre UP². Funkčné prediktívne skóre UP² (Universal Pathogenicity Prediction) v rozmedzí od -1 do 1 klasifikuje patogenitu variantov pomocou prístupu strojového učenia podľa troch kategórií: P (patogénny: skóre > 0,1), VUS (Variant of Unknown Significance: skóre ≥ -0,1 a ≤ 0,1) alebo B (Benign: skóre < -0,1). Skóre UP² sa štandardne zobrazuje v tabuľke variantov, ale používateľ

môže stĺpec UP² deaktivovať v paneli správy stĺpcov kliknutím na ikonu „



“

CI-SpliceAI. Tento stĺpec obsahuje predikčné skóre vypočítané pomocou prístupu strojového učenia na identifikáciu variantov s vplyvom na alternatívne splajzovanie spôsobom *CI-SpliceAI*²⁰.

SpliceAI				
AG	DG	AL	DL	Link
0.01	0.02	0.62	0.84	↗
Position: -2bp				
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	↗
Open variant in SpliceAI Lookup				
No splicing impact				
SpliceAI score was not significant (Delta score <0.2)				

Skóre sa vypočíta pre podmnožinu variantov spĺňajúcich nasledujúce kritériá: varianty splice podľa VEP (splice_*) s frekvenciou GnomAD menšou ako 10⁻³ a spĺňajúce nasledujúce kritériá kvality (QA > 20, AO > 10, DP > 20).

Ak sa vypočíta, stĺpec poskytuje 4 delta skóre pre danú variantu, pre prírastok akceptora (AG), stratu akceptora (AL), prírastok donora (DG) a stratu donora (DL).

Delta skóre variantu, ktoré sa pohybuje od 0 do 1, zodpovedá pravdepodobnosti, že variant modifikuje zostrih. Farebný kód sa uplatňuje na prahové hodnoty delta skóre: **zelená** (>0,2 ; vysoká presnosť), **žltá** (>0,5; odporúčaná) a **červená** (> 0,8; vysoká presnosť). Ak je delta skóre nižšie ako 0,2, považujeme skóre za nevýznamné. Stĺpec potom špecifikuje „Žiaden vplyv na zostrih“.

Pre každú variantu je k dispozícii odkaz na príslušné vyhľadávacie rozhranie *SpliceAI*²¹.

²⁰ https://github.com/YStrauch/CI-SpliceAI__Annotation

²¹ <https://spliceailookup.broadinstitute.org/>

REVEL. Predikčné skóre REVEL je založené na komplexnej metóde predikcie patogenity missense variantov pomocou kombinácie skóre z 13 individuálnych nástrojov: MutPred, FATHMM v2.3, VEST 3.0, PolyPhen-2, SIFT, PROVEAN, MutationAssessor, MutationTaster, LRT, GERP++, SiPhy, phyloP a phastCons.

REVEL ⓘ

B

P

Skóre sa môže pohybovať od 0 do 1 pre missense variant, pričom vyššie skóre odzrkadľuje väčšiu pravdepodobnosť, že variant spôsobuje ochorenie. Použité prahové hodnoty sú nasledovné: skóre REVEL > 0,5 pre **P** (patogénny) a skóre REVEL < 0,5 pre: **B** (benígny). Skóre môže používateľ zobrazit' v tabuľke variantov zaškrtnutím políčka REVEL v časti „Stĺpce“ v sekcii „Funkčná predikcia“.

ALM. Skóre predikcie AlphaMissense²² je k dispozícii v stĺpci ALM v časti „Stĺpce“ pod „Funkčná predikcia“. AlphaMissense je predikčný model, ktorý hodnotí patogenitu missense variantov pomocou hlbokého učenia. Model je postavený na štrukturálnych predikciách AlphaFold s cieľom zlepšiť jeho presnosť.

REVEL ⓘ
ALM ⓘ

Likely Pathogenic score : 0.78

P

B

AlphaMissense integruje evolučnú konzerváciu a štrukturálny kontext, aby posúdil potenciálny vplyv missense variantov na funkciu proteínu.

Skóre sa pohybuje od 0 do 1, pričom vyššie hodnoty naznačujú väčšiu pravdepodobnosť, že varianta je škodlivá. Použité prahové hodnoty sú: skóre ALM > 0,564 pre P (pravdepodobne patogénne) a skóre ALM < 0,34 pre B (pravdepodobne benígne). V ostatných prípadoch (0,34 < skóre ALM < 0,564) bude označené ako **VUS** (nejednoznačné).

Frekvencia vzorky. Tento stĺpec zobrazuje pre každú variantu jej frekvenciu v projekte (podstĺpec PROJECT) a vo všetkých projektoch entity s podobnými metadátami (podstĺpec „ALL“).

Frekvencia vzoriek projektu (PROJECT) zodpovedá počtu vzoriek, v ktorých bola zistená varianta, v pomere k celkovému počtu vzoriek analyzovaných v projekte.

Frekvencia vzoriek ALL zodpovedá percentuálnemu podielu analyzovaných vzoriek, v ktorých sa variant nachádza na entite, v rámci rovnakého metadátového projektu. Kliknutím na hodnotu zobrazenú v stĺpci „VŠETKO“ môže

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

²² <https://alphamissense.hegelab.org/>

používateľ získať prístup k názvom projektov a počtu vzoriek, v ktorých sa varianta nachádza, ako aj podrobnosti o všetkých podobných projektoch metadát (možnosť „Zobraziť všetky projekty“).

Podstĺpec VŠETKY je v tabuľke variantov štandardne skrytý, ale je možné ho znovu aktivovať v paneli správy stĺpcov kliknutím na ikonu x a aktivovaním tlačidla „Entita frekvencie vzorky (Entita)“.

Exón. Tento stĺpec uvádza celkový počet exónov v géne podľa príslušného transkriptu, ako aj počet ovplyvnených mutáciou vo forme zlomku. Nástrojová tipka uvádza vzdialenosť medzi variantom a najbližším miestom spliceovania. Ak je variant nad miestom spliceovania, hodnota bude kladná, a ak je pod ním, bude záporná.

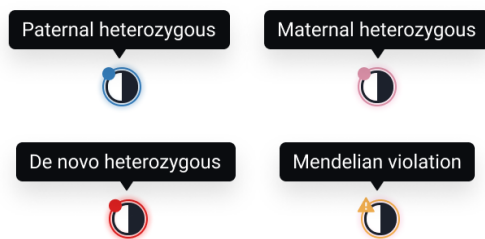
GT (Genotype). Tento stĺpec zobrazuje genotyp predpovedaný programami na detekciu variantov. Rôzne možné piktogramy sú popísané v časti **Filtrovanie variantov**.

V trojitej analýze (GermlineFamily), ak je možné určiť pôvod heterozygotnej varianty, bude to označené nasledujúcim farebným kódom:

- **ružová** → materský heterozygotný pôvod
- **modrá** → heterozygotný pôvod po otcovej strane
- **červená** → *De novo* heterozygotná varianta
- **oranžová** → Mendeliho porušenie

Ak nie je prenos určený (napr. ak sú obaja rodičia heterozygotní) alebo v prípade homozygotných variantov, farba sa nepoužíva.

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



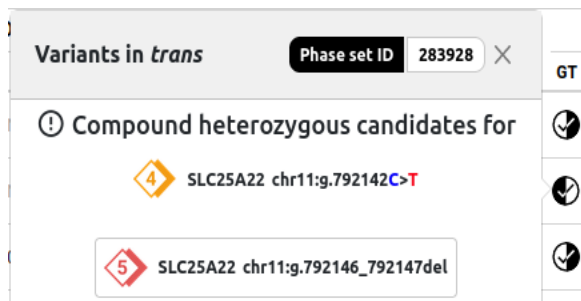
Obrázok 13: Farebný kód použitý na heterozygotné genotypy na základe prenosu GermlineFamily

Fázovanie variantov je k dispozícii pre analýzy GermVar Tertiary a GermVar Tertiary Family pomocou súborov VCF zo sekvenátorov Oxford Nanopore Technologies. Keď sú varianty

fázované, ikony genotypu sa zobrazujú takto: . Začiarknutie označuje, že variant je fázovaný. Fázované varianty s čiernou časťou na tej istej strane a patriace do rovnakého „Phase ID“ sú vo fáze, t. j. na rovnakom reťazci.

Kliknutím na genotyp získate zoznam cis alebo trans variantov, ak sú prítomné. Zobrazené sú len 5 variantov s najvyšším hodnotením DiagAI, ako je uvedené na obrázku.

Kompletný zoznam variantov je k dispozícii po kliknutí na „Prehliadať všetky varianty“.



VAF (Variant Allele Frequency). Tento stĺpec označuje frekvenciu variantnej alely, t. j. pomer čítaní podporujúcich alternatívnu alelu (AO) k celkovému počtu čítaní v danej pozícii (COV), vyjadrený v percentách.

Analýzy používajúce pracovné sady SomaVar verzie ≥ 2.5 , SomaCGP verzie ≥ 0.7 a SomaVar LF verzie ≥ 2.1 zobrazujú nasledujúce dodatočné informácie:

- *nápoveda zobrazuje hodnoty VAF, AO a COV vypočítané dvoma volajúcimi variantami Freebayes a Mutect2, ktoré detekujú variantu.*
- *varovná ikona  upozorňuje používateľa na akýkoľvek nesúlad medzi 2 variantnými volajúcimi pre varianty detegované s nízkym VAF. Pre varianty zobrazujúce VAF menej ako 5 % v stĺpci VAF sa teda zobrazí varovanie v prípade nesúladu VAF > 10 % medzi VAF vypočítanými programami Freebayes a Mutect2.*

AO (Alternate Observation). Tento stĺpec udáva počet čítaní nesúcich variantu.

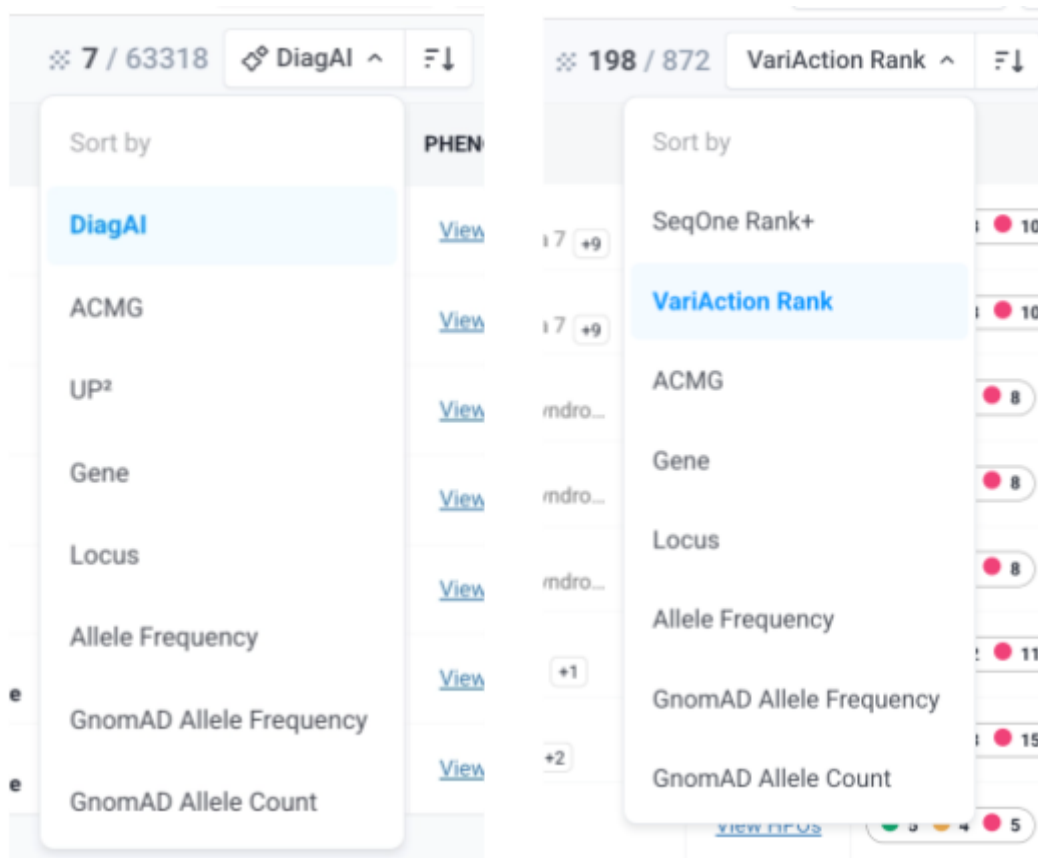
CAL (Caller). Tento stĺpec uvádza názov *volajúceho variantu* (volajúcich variantov), ktorý (ktoré) detegoval variant.

Cov (Coverage). Tento stĺpec uvádza celkový počet čítaní v pozícii variantu.

SOR. Strand Odds Ratio (SOR) je metrika vláknovej predpojatosti (strand-bias) odvodená z testu symetrického pomeru šancí (Symmetric Odds Ratio), ktorá meria nerovnováhu čítaní podporujúcich variant medzi priamymi (forward) a spätnými (reverse) vláknami. Vyššia hodnota SOR naznačuje vyššiu pravdepodobnosť artefaktu, zatiaľ čo nižšia hodnota SOR naznačuje vyváženejšiu podporu vlákien a je konzistentnejšia so skutočným variantom.

Triedenie variantov

Na karte Varianty je k dispozícii niekoľko možností triedenia variantov. Tieto možnosti sú dostupné v roletovom menu v pravom hornom rohu tabuľky. Pre každú možnosť triedenia je možné triediť vzostupne alebo zostupne kliknutím na ikonu triedenia . V predvolenom nastavení sú všetky možnosti triedené zostupne, s výnimkou možnosti „Locus“, ktorá je triedená vzostupne.



Obrázok 14: Okno výberu možnosti triedenia variantov

DiagAI Score. Táto možnosť triedenia sa štandardne používa pri všetkých analýzach germline a umožňuje triediť varianty podľa ich skóre DiagAI (nahradzuje SeqOne Rank+ v analýzach vykonaných po 2. októbri 2024).

SeqOne Rank+. Algoritmus prioritizácie SeqOne Rank+, založený na staršom modeli umelej inteligencie ako DiagAI Score, zostáva k dispozícii a štandardne sa uplatňuje na všetky analýzy vykonané pred 2. októbrom 2024.

Je založený na 4 hlavných komponentoch:

- Nová implementácia klasifikácie ACMG, založená na strojovom učení tréňovanom na ClinVar (skóre UP²).
- Fenotyp pacienta (ak sú s vzorkou spojené termíny HPO)
- Súlad medzi pozorovaným genotypom a spôsobom prenosu patológie opísanej v databázach GenCC, OMIM a PanelApp (Austrália).
- Kvalita variantov

VariAction Rank. Algoritmus VariAction Rank na prioritizáciu somatických variantov je navrhnutý v analýzach SomaVar od verzie v1.8. Tento systém používa anotáciu JaxCKB na výpočet skóre prioritizácie variantov na základe ich klinickej použiteľnosti. Tieto skóre prioritizácie zohľadňujú vplyv variantu, jeho patogenitu pomocou skóre ComPerMed, úroveň AMP, špecifickosť variantu s molekulárnym profilom, ako aj súlad s indikáciou, ak bola choroba zadaná pred spustením analýzy.

ACMG. Varianty je možné zoradiť podľa ich klasifikácie ACMG výberom tejto možnosti. Predvolene sú varianty zoradené v zostupnom poradí od patogénnej triedy 5 po benígnu triedu 1.

Lokus. Varianty je možné zoradiť podľa ich genómovej pozície výberom tejto možnosti.

Frekvencia alely. Varianty je možné triediť podľa frekvencie alely výberom tejto možnosti. Predvolene sú varianty triedené v zostupnom poradí. Je možné triediť v ascendentnom poradí

kliknutím na ikonu triedenia  .

Frekvencia alely GnomAD, počet aliel GnomAD. Varianty je možné zoradiť podľa frekvencie alely GnomAD (AF) alebo počtu aliel (AC) v ascendentnom alebo descendentnom poradí (ak sú k dispozícii údaje GnomAD).

UP². Varianty je možné zoradiť podľa skóre patogenity UP², čo je zložka DiagAI Score.

Gene. Varianty možno zoradiť vzostupne alebo zostupne podľa abecedy podľa symbolu génu. Pri variantoch ovplyvňujúcich prekrývajúce sa gény je symbol génu použitý na zoradenie ten, ktorý je spojený s variantom s najvyšším dopadom, na základe pravidiel prioritizácie transkriptov opísaných vyššie. Pri rozšírení zoznamu transkriptov môže používateľ vidieť druhý prekrývajúci sa gén, ktorý sa môže javiť ako nezoradený. Zoradenie sa tiež nemusí prejaviť, ak je na tabuľku variantov aplikovaných viacero filtrov súčasne.

Konfigurácia stĺpcov variantov v tabuľk

Stĺpce zobrazené v tabuľke variantov je možné konfigurovať v ponuke „Organise your columns“

(Usporiadajte stĺpce), ktorá je dostupná prostredníctvom tlačidla „“ (Usporiadajte stĺpce) v pravom hornom rohu tabuľky variantov.

Zobrazenie stĺpcov môže používateľ zmeniť kliknutím na ikony „“ (Aktívny/Neaktívny) v ľavej časti, čím aktivuje požadované stĺpce. V pravej časti sa zobrazia aktívne stĺpce v analýze a umožňuje používateľom vytvárať a ukladať profily stĺpcov z roletového menu kliknutím na ikonu „Select profile“ (Vybrať profil).

Aktívne stĺpce je možné preskupiť kliknutím a pretiahnutím na požadované miesto. Predvolené sú stĺpce usporiadané podľa profilov „Germline“ a „Somatic“.

Keď používatelia prejdú kurzorom nad názvom profilu stĺpca v prednastavenom roletovom menu, zobrazí sa nápoveda, ktorá workset vytvorila tento profil. Týmto spôsobom sú používatelia upozornení, ak ich profil obsahuje stĺpce, ktoré nie sú kompatibilné s aktuálnou pracovnou sadou.

Ak chcete vytvoriť nový profil stĺpca, stačí zadať názov do poľa „Vytvoriť novú predvolenú sadu údajov“ a kliknúť na „Uložiť“.

Úprava vzorových metadát z panela analýzy worksetu

Používateľ môže upravovať metadáta vzorky priamo z analýzy a zadávať pohlavie pacienta alebo HPO (len pre analýzy zárodočných buniek) z priečinka Analýza, ktorý je dostupný kliknutím

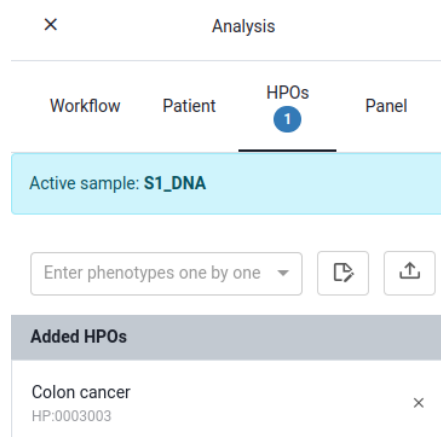
na tlačidlo  Analysis v pravom hornom rohu stránky (pozri **Bočný panel pre správu analýz**).

Z pohľadu pacienta má používateľ prístup k úprave pohlavia (pozri časť **Modálny pohľad pacienta**).

V analýzach projektu Germline má používateľ z pohľadu HPOs prístup k viacerým funkciám úpravy termínov HPOs:

- Zobrazenie termínov, ktoré boli zadane pred spustením analýzy
- Ručné pridávanie kódov HPO alebo fenotypov prostredníctvom vyhľadávacieho poľa „Zadajte fenotypy jeden po druhom“
- Import textového súboru obsahujúceho kódy HPO
- Generovanie termínov HPO na základe klinického opisu pacienta vyvolaním DiagAI prostredníctvom

 . Táto funkcia vyzve používateľa, na popis klinického profilu pacienta v angličtine (opis je obmedzený na 4000 znakov). Potom po kliknutí na „Extract HPOs with DiagAI“ (Extrahovať HPO s DiagAI) sa vygeneruje zoznam HPO zvýraznených farbou. Nakoniec môže používateľ priradiť všetky navrhované termíny HPO k pacientovi kliknutím na „Add all“ (Pridať všetko) alebo ich vybrať jednotlivo tak, že umiestni kurzor myši na vybrané termíny a klikne na „Add“ (Pridať).



Pridávanie a úprava HPO z tohto zobrazenia dáva používateľovi možnosť opätovného spracovania skóre DiagAI a klasifikácie variantov (krátky zoznam alebo SmartPick) podľa nových HPO zadanych prostredníctvom funkcie „Opätovné spracovanie analýzy“.

Filtrovanie variantov

Vyhľadávanie môžete spresniť na karte Varianty pomocou filtrovacích kategórií uvedených nižšie. Panel filtrov je prístupný cez ikonu „ Filters“ umiestnenú v ľavom hornom rohu tabuľky variantov alebo stlačením klávesy F na klávesnici. Filtre sú k dispozícii na karte „Basic“ (Základné). Sekcia „Active Filters“ (Aktívne filtre) umiestnená v hornej časti panelu filtrov zobrazuje počet filtrov, ktoré sú momentálne v použití.

Filtry génov a variantov

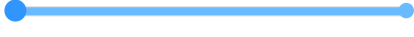
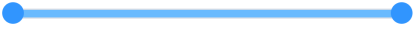
Názov filtra a prehľad	Popis
Genes ⓘ Search gene symbols	Obmedzuje zoznam variantov na jeden alebo viac génov.
Variants ⓘ Search by HGVS	Obmedzuje zoznam variantov na jeden alebo viacero variantov podľa nomenklatúry HGVS alebo HGVS.

Sekcia Vzorky a kvalita

Názov filtra a prehľad	Popis
Coverage ⓘ 1 1 836	Definuje minimálnu hranicu pre hĺbku pokrytia (COV). Maximálna hodnota zobrazená týmto filtrom zodpovedá najvyššej hodnote pokrytia identifikovanej medzi variantmi.
Allele Frequency ⓘ 0% 0 100 100%	Filtruje varianty podľa frekvencie mutovanej alely v analyzovanej vzorke.
Quality ⓘ Low Medium High	Priemerná kvalita (phred skóre) každej bázy nesúcej mutovanú alelu, ktorá udáva spoľahlivosť detekcie variantov: - Stredná ($\geq \text{QUAL}=20$): 1 šanca zo 100 na nesprávne určenie, - Vysoká ($\geq \text{QUAL}=30$): 1 šanca z 1000 na nesprávne určenie. Tento filter sa vzťahuje na všetky vzorky.
Quality ⓘ All Low $\geq Q8$ Medium $\geq Q20$ High $\geq Q30$	Pri analýzach GermVar Tertiary a GermVar Tertiary Family na údajoch od Oxford Nanopore Technologies kvalita založená na skóre Phred indikuje spoľahlivosť detekcie variantov. $\geq \text{QUAL}=8$ (nízka): 1 šanca z 8 na nesprávne určenie, $\geq \text{QUAL}=20$ (stredná): 1 šanca zo 100 na nesprávne určenie, $\geq \text{QUAL}=30$ (vysoká): 1 šanca z 1000 na nesprávne určenie.

	Tento filter sa vzťahuje na všetky vzorky.
Genotype ⓘ ● Homozygous 387 ● Heterozygous 1163 ○ None 1942 ⦿ Biallelic Variant 222 ⦿ Biallelic Variant 84 ⦿ Unknown 76 ? Unknown 0	Filtruje varianty podľa pozorovaného genotypu.
Alternate Observations ⓘ 0 0 458	Definuje počet sekvencií nesúcich mutáciu (alternatívne pozorovanie alebo AO). Maximálna hodnota zobrazená týmto filtrom zodpovedá najvyššej hodnote AO identifikovanej medzi variantmi.
Strand Odds Ratio ⓘ 0 5 0 1 2 3 4 ≥5 min 0 — max 5	Filtruje varianty podľa skóre strand odds ratio (SOR) vypočítaného v mieste varianty.

Sekcia Frekvencia

Názov a prehľad filtra	Popis
▼ Sample frequency ⓘ all  all rare	Eliminuje varianty na základe ich výskytu v skúmanom projekte, od najčastejších (všetky) po najmenej časté (vzácne).
▼ Relative pop frequency ⓘ all 10-2 10-3 10-4 10-5 10-6 absent 	Vyberá varianty podľa frekvencie ich mutovanej alely v kombinovanej populácii projektov ExAc a gnomAD.
▼ Absolute pop frequency ⓘ 0 1609917 	Vyberá varianty na základe počtu aliel v kombinovanej populácii ExAc a GnomAD.
▼ Absolute homozygous pop frequency ⓘ 0 803576 	Vyberá varianty na základe ich homozygotnej frekvencie v kombinovanej populácii ExAc a GnomAD.

Sekcia genetika

Názov filtra a prehľad	Popis
Impact ⓘ Low Moderate High ☒ Keep <i>Modifier</i>	Vplyv mutácie na proteín. <i>Modifikátor: ovplyvňuje nekódujúce gény</i>
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	Termíny SO (Sequence Ontology), ktoré opisujú najrušivejšie dôsledky na úrovni génov. Tento filter vám umožňuje vybrať vplyv mutácie na gén a možnosti sú usporiadané do logických kategórií, aby sa zlepšila navigácia: : <i>Strata funkcie, Kódovanie, Oblasť šplhania, Tichá varianta, Intronová, UTR, ncRNA, Upstream/Downstream, Regulačná, Ostatné.</i>
Gene panel ⓘ Select panel	Vyberte a použite jeden alebo viacero génových panelov. Tento filter zobrazuje všetky genové panely dostupné pre danú entitu. Tento filter nie je k dispozícii, ak bol na analýzu použitý obmedzený panel z karty <i>Panel</i> v paneli správy analýzy (pozri časť Modálne zobrazenie panelu). Správa informuje používateľa, že filter „Génový panel“ nemožno použiť s obmedzeným panelom.
Panel in-silico ⓘ Dyslipidemia_panel_1_Twist★	Vyberie a použije panel in-silico. Tento filter zobrazuje všetky dostupné súbory BED pre skupinu. Názov súboru použitého počas analýzy je zvýraznený hviezdíčkou. Tento filter zachováva iba varianty, ktorých počiatočná pozícia (začiatok) sa nachádza v jednom z intervalov vybraného súboru bed.

PanelApp ⓘ Epi ENG Blepharophimosis ptosis and epicanthus ENG Early onset or syndromic epilepsy ENG Epidermodysplasia verruciformis	Umožňuje filtrovať varianty podľa zoznamu in-silico panelov v katalógoch PanelApp Australia a PanelApp England. Odznak identifikuje databázu, z ktorej in-silico panel pochádza: AUS (Austrália) alebo ENG (Anglicko). V oboch katalógoch nie sú k dispozícii na výber superpanely, ani panely, ktoré neobsahujú aspoň jeden gén. Napríklad panel obsahujúci iba STR nebude možné vybrať.
Gene occurrences ⓘ Usable only if they are fewer than 2000 selected variants.	Definuje prah pre počet výskytov daného génu v tabuľke variantov.
Chromosomes & MT ⓘ X 25 Y 0 M 0	Zachováva varianty umiestnené na pohlavných chromozómoch a/alebo mitochondriálnom chromozóme (M).
RefSeq transcripts ⓘ BARD1_2transcripts.txt	Vyberá zoznam transkriptov RefSeq na filtrovanie variantov.

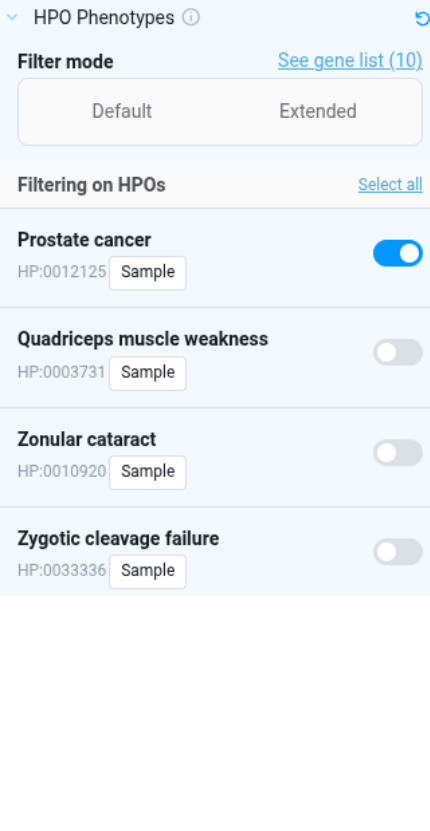
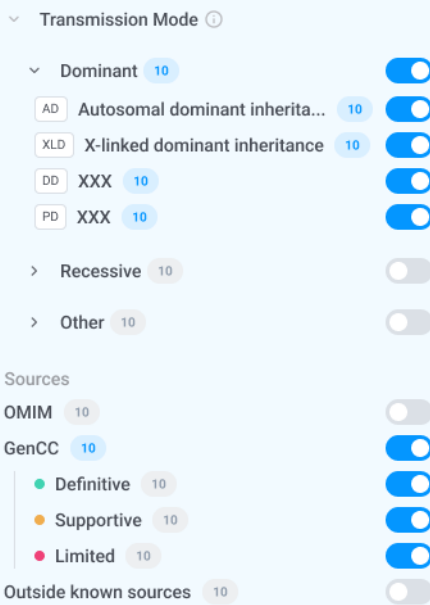
Sekcia Databáza

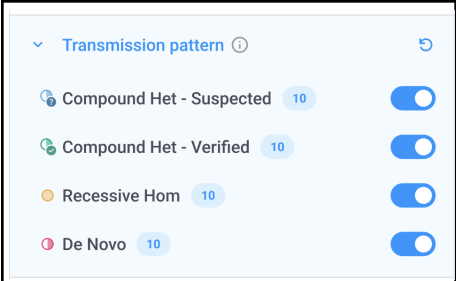
Názov filtra a prehľad	Popis
Variant Databases ⓘ dbsnp 31 cosmic 31 civic 30 lovd 10 clinvar 3	Vyberá varianty anotované nasledujúcimi databázami: dbsnp, cosmic, lovd, clinvar a civic.

▼ 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

Sekcia prenosu a fenotypov

Názov filtra a prehľad	Popis
▼ Description keywords ⓘ	Používa kľúčové slová na charakterizovanie fenotypu (fenotypov) spojených s variantmi prostredníctvom vyhľadávania v databázach OMIM, Medgen a Litvar. Táto funkcia nezohľadňuje HPO, pretože na ne je určený špecifický filter (pozri nižšie). K dispozícii sú dva režimy vyhľadávania: zjednotenie a prienik.

	Umožňuje filtrovanie variantov na základe predtým zadanych termínov HPO, buď pred analýzou vzorky (pozri Metadáta vzorky), alebo po analýze z karty HPO v paneli analýzy (pozri Modálne zobrazenie HPO). Zhody sa identifikujú pomocou nášho prístupu PhenoGenius, ktorý sa pri výbere génov opiera o maticu HPO-gén. V režime Filtrovanie sú k dispozícii 2 možnosti: • Predvolené filtrovanie, ktoré vyvažuje presnosť a počet vrátených génov. • Rozšírené filtrovanie maximalizujúce presnosť s väčším počtom génov Možnosť „Zobraziť zoznam génov“ poskytuje zoznam génov spojených s HPO vybranými v tomto filtri. Táto funkcia je k dispozícii len v prípade, ak sú zaškrtnuté termíny HPO. Otvorí sa vyskakovacie okno, v ktorom sa zobrazí ID každého génu, symbol, poradie Phenogenius a skóre Phenogenius. Zoznam génov si môžete stiahnuť pomocou ikony Stiahnuť alebo skopírovať pomocou tlačidla Kopírovať/Vložiť.
	Umožňuje používateľom filtrovať varianty pomocou dedičných prepojení z chorôb súvisiacich s ľudskou lekárskou genetikou. Varianty je možné filtrovať podľa spôsobov prenosu (dominantný, recesívny, iný) a zdrojov prenosu (OMIM, GenCC alebo iné databázy).

	Umožňuje výber variantov podľa heterozygotných kompozitných (podozrivých alebo overených), homozygotných recesívnych alebo de novo spôsobov prenosu na základe DiagAI. Upozorňujeme, že filtre „Compound Het – Verified“, „Recessive Hom“ a „De novo“ sú k dispozícii iba v rodinných analýzach.
---	--

Sekcia klasifikácie

Názov filtra a prehľad	Popis
New Unknown variant 5 Pathogenic 4 Likely Pathogenic 3 Uncertain Significance 2 Likely Benign 1 Benign Possible artifact ☐ Analysis only ⓘ	Umožňuje vybrať neznáme varianty, ktoré neboli vyhodnotené (New - Unknown variant) alebo validované vo VKB vo všetkých analýzach entity. Ak je zaškrtnuté políčko „Analysis only“, varianty sa budú vyhľadávať iba v analýze, ktorá sa momentálne anotuje.
5 Pathogenic 5 4 Likely Pathogenic 1 3 Unknown significance (VUS) 56 2 Likely Benign 49 1 Benign 6	Klasifikácia patogenity pre varianty zárodočných buniek, vypočítaná na stupnici od 1 (benígna) do 5 (patogénna).
5 Pathogenic 7 4 Likely Pathogenic 6 3 VUS 102 2 Likely Benign 1 1 Benign 7	Vyberá somatické varianty podľa ich triedy patogenity ComPerMed na stupnici od 1 (benígne) do 5 (patogénne).
Actionable 1346 Diagnostic 0 Prognostic 0	Vyberá varianty podľa 3 typov dôkazov (CKB): akčné, diagnostické, prognostické.

AMP/ASCO/CAP Tier ⓘ I II 1346 0	Vyberá varianty podľa ich priradenej úrovne AMP (úrovne I, II a N/A) iba v analýze somatického kontextu.
---	--

Sekcia Skóre

Názov filtra a prehľad	Popis
CI-SpliceAI ⓘ all > 0.2 > 0.5 > 0.8 Keep unknown	Vyberte delta skóre. Políčko „Zachovať neznáme“ zachová varianty s hodnotou „N/A“, ale vylúči varianty bez vplyvu na spliceovanie (varianta „Bez vplyvu na spliceovanie“).
Conservation ⓘ Low Conserved Highly Conserved	Stupeň konzervácie miesta, vypočítaný z viacerých skóre konzervácie.
Metasvm ⓘ Tolerated Damaging Highly Damaging	Zohľadňuje skóre 10 aplikácií na predikciu variantov a maximálnu frekvenciu pozorovanú v populácii 1000genomes.
Universal Pathogenicity Predictor (UP²) ⓘ Benign VUS Pathogenic -1 0 1 Min Max -1 1	Vyberá varianty podľa univerzálneho prediktora patogenity (UP²).
LOEUF ⓘ Intolerant Tolerant 0 1 2 Min Max 0 2 Keep unknown	Vyberá varianty podľa skóre LOEUF.

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

Vyberte indikáciu

V SomaVar, SomaVar LF a SomaCGP je možné vybrať preddefinovanú indikáciu patológie z databázy JaxCKB z roletového menu, ktoré je k dispozícii na karte *Pacient* v bočnom paneli správy analýzy (pozri časť **Modálne zobrazenie pacienta**).

Ontologický systém je založený na *ontológii chorôb*²³ a obsahuje termíny špecifické pre JaxCKB. Informácie o citlivosti v podstĺpci „A“ stĺpca Klinické dôkazy (pozri časť **Stránka variantov**) sú zvýraznené zelenou farbou len vtedy, ak indikácia pre daný variant zodpovedá vybranej indikácii. Ak sa pre variant nájdu tvrdenia o rezistencii, príslušná bunka v stĺpci „A“ bude zvýraznená červenou farbou bez ohľadu na indikáciu. V opačnom prípade bude sivá.

Ak bolo ochorenie vybrané počas načítania vzorky a pred analýzou (pozri **Metadáta vzorky**), stane sa predvolenou indikáciou definovanou pre vzorku a v karte *Pacient* na bočnom paneli správy analýzy. Je označené ikonou hviezdičky a zobrazuje sa ako prvé v zozname. Zmena indikácie na bočnom paneli správy analýzy neaktualizuje predvolené ochorenie pre vzorku (Metadáta ochorenia).

Analysis

Workflow Patient HPOs Panel

Active sample: CTDNA_toy

Sex

Female Male Unknown

Indication

★ Colorectal cancer (Colorectal cancer) × ▾

²³ <https://disease-ontology.org/>

Použitie preddefinovaného profilu filtra

Používatelia môžu vybrať preddefinovaný profil filtra z roletového menu kliknutím na ikonu „Vybrať profil“. Platforma SeqOne ponúka niekoľko preddefinovaných profilov filtrov ako predvolené: *GermVar*, *Moja voľba*, *Výskum*, *DiagAI*.

Úprava a uloženie profilov filtrov

Používatelia môžu uložiť aktuálnu polohu kurzorov pre každý filter ako profil filtra tak, že profilu zadajú názov a kliknú na „Uložiť ako nový profil“. Profil filtra bude potom k dispozícii v rovnakom roletovom menu v kategórii Súkromné.

Pri ukladaní profilu filtra sa otvorí formulár na vytvorenie a používateľ môže vybrať tlačidlo s názvom entity, aby bol prístupný ostatným používateľom entity. Ak používateľ vyberie tlačidlo „Súkromné“, profil bude prístupný len používateľovi, ktorý ho vytvoril.

Iba používateľ, ktorý pôvodne vytvoril profil filtra, ho môže upraviť kliknutím na „Aktualizovať“; ak profil upraví iný používateľ v organizácii, bude požiadaný, aby ho uložil pod iným názvom. Ak používateľ nemá oprávnenie na úpravu verejného profilu filtra, bude požiadaný, aby kontaktoval správcu svojej organizácie. Profily filtrov SeqOne nemožno upravovať a ukladať pod iným profilom filtra.

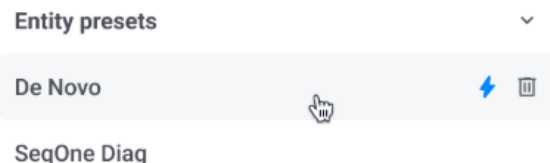
Jeden alebo viacero filtrov je možné vynulovať. Stačí kliknúť na ikonu „“ (Vynulovať všetky filtre) vpravo od filtra alebo v pravom hornom rohu okna „Filters“ (Filtry).

Oblíbené profily filtrov

Používatelia môžu nastaviť profily filtrov ako oblíbené kliknutím na ikonu  vpravo od profilu, aby sa zobrazili ako skratka priamo v lište rýchleho prístupu nad tabuľkou variantov.

Ak chcete odstrániť profil filtra z oblíbených, stačí

kliknúť na rovnakú ikonu . Profil Moje výbery je predvolene definovaný ako oblíbený.

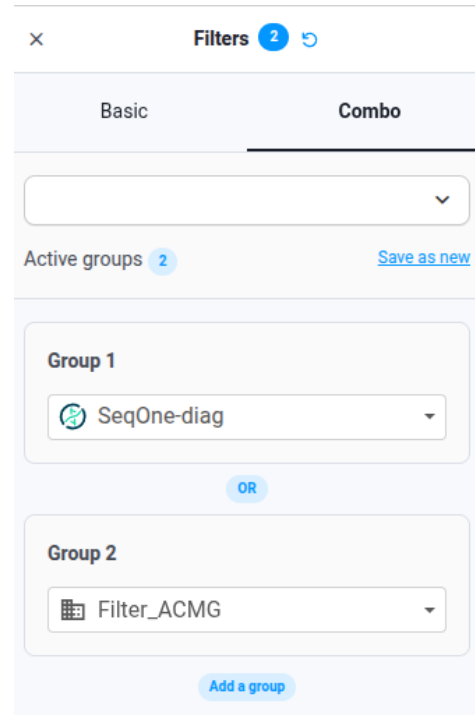


Kombinovanie profilov filtrov

Profile filtrov je možné kombinovať pod podmienkou **OR** kliknutím na kartu Combo v paneli filtrov a potom na „Vytvoriť kombináciu“. V predvolenom nastavení sa vo formulári na vytvorenie zobrazujú 2 skupiny profilov filtrov, ale používatelia môžu pridať ďalšie kliknutím na „Pridať skupinu“.

Po výbere profilov sú používatelia vyzvaní, aby uložili svoj kombinovaný profil kliknutím na „Uložiť ako nový“. Rovnako ako v prípade profilov filtrov môže používateľ sprístupniť svoj kombinovaný profil všetkým používateľom entity vo formulári na vytvorenie. Iba používateľ, ktorý kombinovaný profil pôvodne vytvoril, ho môže upraviť kliknutím na „Aktualizovať“.

Ak profil upraví iný používateľ v entite, bude tento používateľ požiadaný, aby ho uložil pod iným názvom. Ak používateľ nemá oprávnenie na úpravu verejného kombinovaného profilu, bude požiadaný, aby kontaktoval správcu svojej entity.



The screenshot shows a 'Filters' window with a 'Combo' tab selected. At the top, there's a dropdown menu. Below it, 'Active groups' is shown as 2, with a 'Save as new' link. The main area contains two groups: 'Group 1' with a dropdown showing 'SeqOne-diag' and 'Group 2' with a dropdown showing 'Filter_ACMG'. An 'OR' operator is placed between the two groups. At the bottom right, there is a button labeled 'Add a group'.

Filtrovanie profilov SeqOne

Môj výber. Tento profil filtra umožňuje zobraziť varianty hodnotené ako 3 (VUS), 4 (pravdepodobne patogénne) a 5 (patogénne) priamo v VKB v interpretovanej analýze.

DiagAI. Tento profil filtra, dostupný v analýzach germline (GermlineVar, GermlineFamily, GermVar Tertiary, GermVar Tertiary Family), zobrazuje zredukovaný zoznam variantov s vysokou pravdepodobnosťou, že sú príčinou ochorenia pacienta. Tento profil využíva algoritmus prioritizácie DiagAI a výskyt variantu v projekte (filter Sample Frequency).

Môže sa použiť, ak sú alebo nie sú termíny HPO spojené so vzorkou, ale jeho citlivosť sa zvyšuje v prítomnosti termínov HPO. V prípade analýzy rodiny sa profil filtra automaticky prispôbí veľkosti analyzovanej rodiny, pretože zohľadňuje, že výskyt variantov v projekte môže byť väčší. Používatelia môžu tiež kedykoľvek upraviť filter Sample Frequency v projekte.

Zadajte HPO, aby sa zobrazili varianty SmartPick s najvyššou pravdepodobnosťou spôsobenia ochorenia.

Rodinné profily. V analýzach GermlineFamily a GermVar Tertiary Family budú filtre *De Novo* (nové), *Recessive Hom* (nové) a *Compound Het* (nové) k dispozícii len pre analýzy trojíc, ktoré

zahŕňajú postihnutý indexový prípad, nepostihnutého otca a nepostihnutú matku. V opačnom prípade budú tieto filtre deaktivované.

Pri vytváraní a ukladaní osobného profilu sa neodporúča upravovať profily filtrov *De Novo* (nové), *Recessive Hom* (nové) a *Compound Het* (nové). Pri vytváraní variantu týchto profilov sa odporúča začať od začiatku.

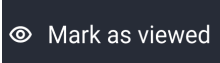
Skríning nositeľov. Tento profil filtra je k dispozícii iba v rodine Germline pre analýzy dvojíc (rodičia bez indexovej vzorky) a identifikuje recesívne gény s potenciálne patogénnou heterozygotnou variantou u oboch nepostihnutých rodičov pre testovanie pred počatím.

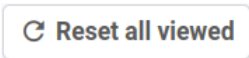
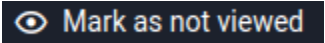
Somatický FFPE. Tento profil filtra zobrazuje zredukovaný zoznam somatických variantov v somatických analýzach (SomaVar, SomaCGP) vykonaných na FFPE DNA.

Somatický LF. Tento profil filtra zobrazuje zredukovaný zoznam somatických variantov v analýzach SomaVar LF.

Označiť ako prehliadnuté Funkcia

„Označiť ako prečítané“ je funkcia, ktorá používateľovi umožňuje selektívne skryť určité varianty.

Stačí vybrať požadované varianty a kliknúť na ikonu „ Mark as viewed“ (Označiť ako prečítané) zobrazenú v spodnej tabuľke variantov. Výber variantov sa potom zobrazí sivou farbou. Ak chce

používateľ výber zrušiť, môže buď kliknúť na ikonu „ Reset all viewed“ v hornej časti tabuľky variantov, čím sa zobrazenie obnoví, alebo vybrať varianty a kliknúť na „ Mark as not viewed“.

Používateľ môže skryť vyskakovacie okno kliknutím na krížik. Tým sa zruší výber všetkých variantov, zostanú však označené sivou farbou.

Výber variantov

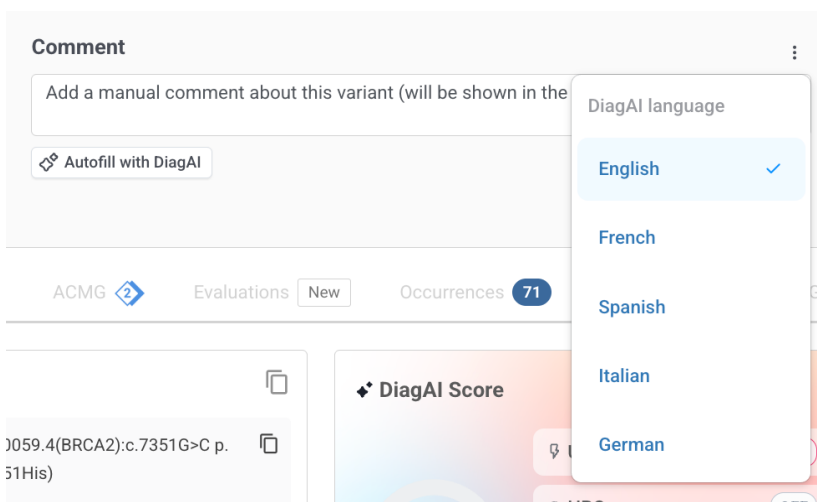
Varianty je možné pridať do panela „Výber“ vedľa akýchkoľvek vyhodnotených variantov (pozri časť nižšie). Na to je potrebné vybrať varianty zaškrtnutím políčok v ľavej časti tabuľky variantov a potom kliknúť na ikonu „Pridať do výberu“ v spodnej časti tabuľky.

Hodnotenie variantov

Varianty je možné rýchlo vyhodnotiť z tabuľky variantov kliknutím na stĺpec VKB a špecifikovaním ich patogénnej triedy a voliteľného komentára. Tieto hodnotenia a komentáre sú prístupné všetkým používateľom platformy v rámci tej istej entity prostredníctvom ikony „Výber“

v analýze (pozri časť **Bočný panel na správu hodnotení**). Variantu je možné vyhodnotiť aj z jej vyhradenej stránky.

Komentáre je možné zadávať ručne alebo generovať automaticky. Rozhranie obsahuje prázdne pole „Klasifikácia“ pre komentáre a tlačidlo „Automatické vyplnenie s DiagAI“ pre automatické generovanie. Počas generovania komentárov sa zobrazuje indikátor. Existujúce komentáre je možné vylepšiť pomocou možnosti „Vylepšiť pomocou DiagAI“. Každý vygenerovaný komentár je možné schváliť, upraviť alebo vynulovať. Po overení sa komentáre uložia spolu s triedou patogenity, dátumom a obsahom. Po overení je možné ich upraviť alebo odstrániť, pričom sa zachová ich história.



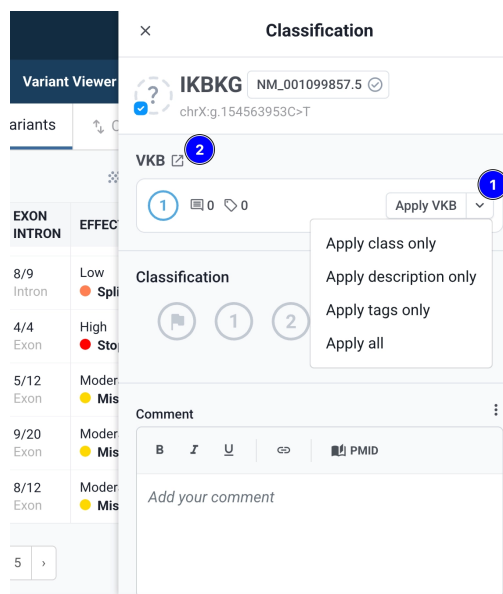
Ak chcete vybrať jazyk pre DiagAI, kliknite na tri bodky v pravom hornom rohu poľa pre komentáre a z roletového menu vyberte preferovaný jazyk.

Varianty možno klasifikovať aj kliknutím do stĺpca klasifikácie, ktorého záhlavie má vlajku. Ak chcete klasifikovať variant, kliknite na ikonu +. Otvorí sa bočný panel klasifikácie, ktorý používateľovi umožňuje pridať triedu patogenity, komentár a tag.

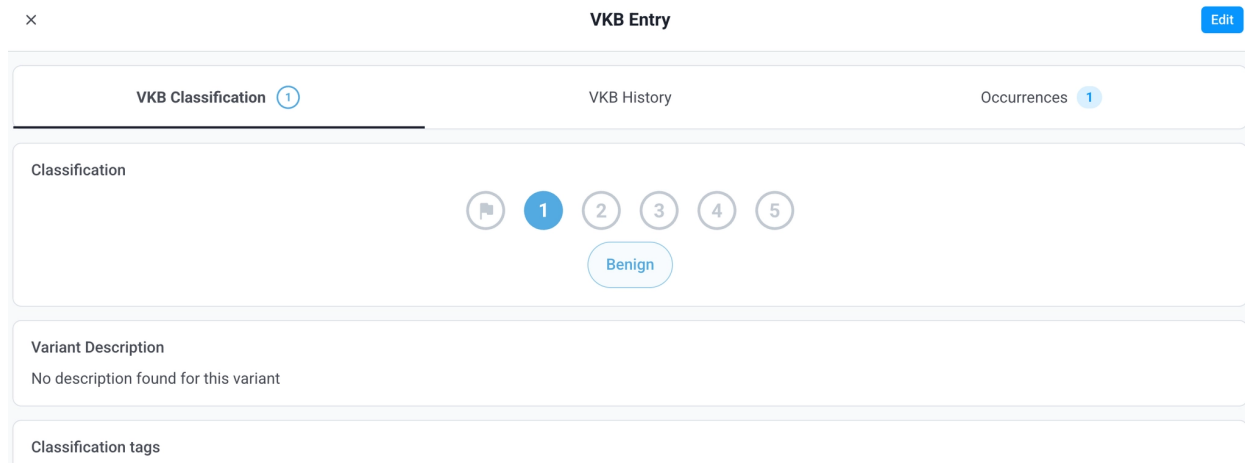
Ak chcete túto klasifikáciu pridať do databázy VKB, aktivujte prepínač "Add to VKB". Ak klasifikácia v databáze VKB už existuje, kliknutím na "Apply VKB" ju aplikujete na aktuálny záznam. Ďalšie možnosti sú k dispozícii kliknutím na šípku vedľa "Apply VKB" (pozri krok 1 na obrázku nižšie).

Klasifikáciu potvrdíte kliknutím na "Save" v pravom hornom rohu bočného panela.

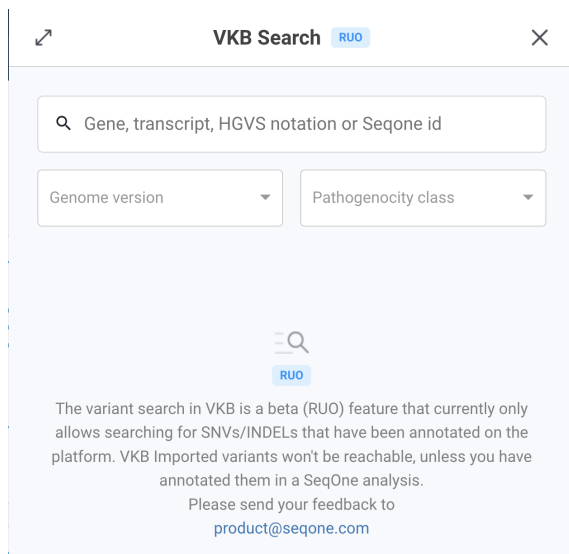
Ak chcete upraviť klasifikáciu priamo v databáze VKB, kliknite na presmerovací odkaz VKB (pozri krok 2 na obrázku nižšie) a otvorte okno "VKB Entry".



V okne "VKB Entry" kliknite na "Edit" vpravo hore, aby ste upravili hodnotenie, a potom kliknite na "Save" na použitie zmien. História VKB a výskyty variantu sú prístupné prechádzaním cez vyhradené karty.



VKB Search (Vyhľadávanie VKB). Bočný panel je k dispozícii po kliknutí na tlačidlo „VKB“ v lište vyhľadávania v hlavičke SeqOne. To vám umožňuje vyhľadávať anotované varianty v databáze VKB podľa názvu génu, transkriptu, anotácie HGVS alebo identifikátora SeqOne (Genome_variant_chromosome_location_REF_ALT). Vyhľadávanie je možné filtrovať podľa referenčného genómu (GRCh37 alebo GRCh38) a/alebo triedy patogenity.



Obrázok 15: Okno vyhľadávania VKB

Po kliknutí na vyhodnotenú variantu sa otvorí podrobnejšia stránka s informáciami o variante, ako je jej genomická pozícia, HGVS anotácie na komplementárnej DNA (cDNA) a na proteíne,

triedy patogenity, analýzy a súvisiace komentáre, kde bola vyhodnotená, a potom odkaz na analýzy.

BRCA1 NM_007294.3

c.2834_2836delinsC p.Ser945ThrfsTer6

gDNA chr17:g.41244712_41244714delinsG

cDNA NM_007294.3:c.2834_2836delinsC

Protein NP_009225.1:p.Ser945ThrfsTer6

VCF 17-41244712-TAC-G

Latest classification: 5

5 (7)
4 (0)
3 (0)
2 (0)
1 (0)
0 (0)

Apr 19, 2024
demo · GermlineVar

PP5: Found a Clinvar Pathogenic entry (<https://www.ncbi.nlm.nih.gov/clinvar/variation/54694>) having at least 1 star. PM1: Found pathogenic hotspot in ClinVar with 28 pathogenic variants out of 68 (73.68%) on 36bp window, above 66.67% threshold. PM2: GnomAD max frequency (0.0 for population) is inferior to 1/10000. PVS1: Predicted LoF variant (frameshift variant) in BRCA1 gene where there is 3418 out of 3641 pathogenic LoF variants found in Clinvar (93.88%) above 10.00% threshold.

Open on analysis

Obrázok 16: Okno s podrobnosťami o variante vo VKB

Stránka variantov

Kliknutím na variantu z tabuľky variantov sa otvorí bočný panel obsahujúci ďalšie informácie s niekoľkými kartami. Používatelia môžu otvoriť stránku varianty v inej karte kliknutím na „Otvoriť v novom okne“ v pravom hornom rohu. Bočný panel je možné zatvoriť pomocou klávesy „Esc“.

V tomto okne môžu používatelia hodnotiť a komentovať varianty. Podobne ako hodnotenia a komentáre v hlavnej tabuľke variantov, aj tieto hodnotenia a komentáre môžu vidieť všetci členovia skupiny, ktorí používajú platformu.

Z tejto stránky majú používatelia tiež prístup k informáciám o variante, vrátane genotypu (GT), alelickej frekvencie (VAF), hĺbky pokrytia (COV) a phred skóre (QA).

Karty na stránke variantov sú popísané nižšie.

Variant (Variantov). Karta je rozdelená na niekoľko blokov, ktoré sú popísané nižšie.

- **Nomenklatúra:** sú uvedené rôzne nomenklatúry (g., c., p., VCF entry) s možnosťou kopírovania informácií z celého bloku alebo z každého riadku do schránky pomocou ikony  .

- **DiagAI Score:** používa blok súhrnu výpočtu DiagAI Score, ktorý je k dispozícii na karte DiagAI Score na stránke variantov (pozri odsek **DiagAI Score**).
- **ClinVar:** informácie prevzaté zo stĺpca ClinVar Germline tabuľky variantov, ktoré uvádzajú triedu patogenity priradenú variantu, jeho klinický stav a počet záznamov v ClinVar (pozri

Vizualizácia a interpretácia v časti **Malé varianty**). Ikona  otvorí stránku ClinVar v novej karte.

- **Anotácie:** vplyv transkriptu, príslušný exón, vzdialenosť od najbližšieho miesta splicingu a typ mutácie.
- **Transmission (Prenos):** odkazy na databázy OMIM a GenCC, ak je s génom spojená jedna alebo viacero chorôb.
- **Kandidáti na zložené heterozygotné mutácie,** tento blok je rozdelený na 2 časti:
 - Sekcia *Krátké varianty* zoskupuje všetky kandidátske SNV nachádzajúce sa na tom istom géne (na základe ID génu) a zoradí ich podľa ich skóre DiagAI. Variant vybraný na stránke Varianty sa zobrazí ako prvý a v zozname je sivý.
 - Sekcia *CNV* uvádza, či boli nájdené CNV, zoradené podľa triedy ACMG. Táto sekcia je k dispozícii len v prípade, ak bola vykonaná analýza CNV, inak sa zobrazí varovná správa „CNV nie sú k dispozícii“.

Kliknutím na jeden z výsledkov sa otvorí nová karta so stránkou variantov (ak ide o malý variant) alebo stránkou CNV (ak ide o kopírovaný variant). Ak sa nenašli žiadne CNV, zobrazí sa varovná správa s textom „no CNVs found“ (nenašli sa žiadne CNV).

- **Nástroje:** odkazy na iné databázy (LitVar, LOVD, VarSome, dbSNP, Cosmic, GnomAD, CKB, UCSC, Decipher, Ensembl).
- **GnomAD:** ak je varianta známa v tejto databáze, zobrazí sa tabuľka s prehľadom frekvencie výskytu varianty v rôznych populáciách a priamy prístup k GnomAD je umožnený cez ikonu  .

Nomenclature

gDNA chr4(GRCh38)g.154611883G>C

cDNA NM_021870.3(FGG):c.323C>G

Protein NP_068656.2 p.Ala108Gly

VCF 4-154611883-G-C

Clinvar

Likely Pathogenic

Criteria Provided, Multiple Submitters, No Conflicts ★★☆☆

Review status **Reclassified by SeqOne**

Submissions **2 P | 9 LP | 2 VUS**

Disease Congenital Fibrinogen Deficiency, Thrombus

Annotations

Impact/Effect **Moderate** ● missense

Exon/Intron 4/9

Biotype protein coding

Transmission

AR

DISEASE	OMIM	MEDGEN
Hypofibrinogenemia, congenital	AR	N/A

Compound Heterozygous candidates

Short variants

79 4 New c.323C>G ● missen...

7 1 New c.*16C>T ● 3 prime ...

1 to 2 of 2 1

CNVs

No CNVs found

Tools

Variant database

Gene database

Genome database

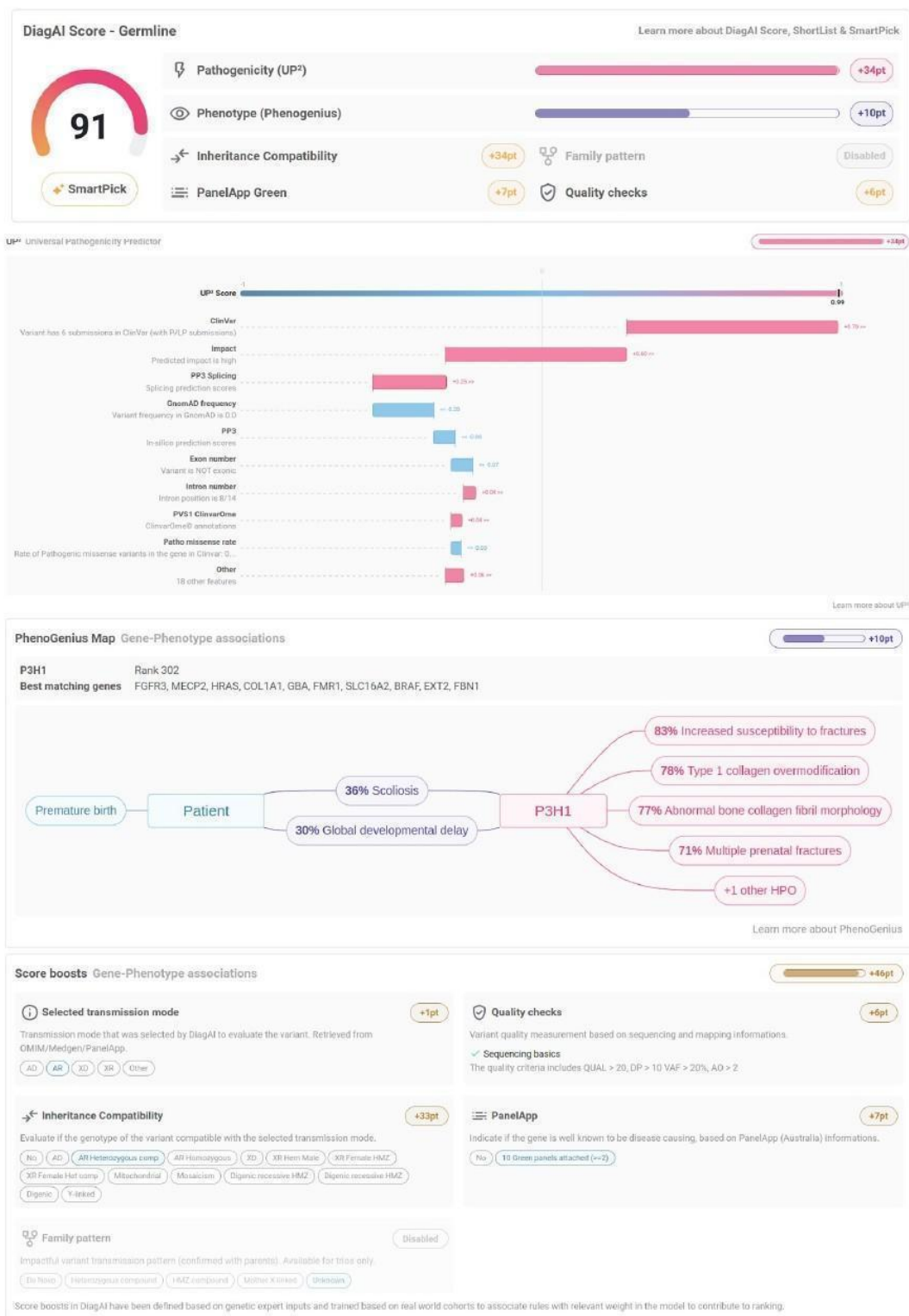
GnomAD

MAF 3.84e-3 ALL 3.03e-3

QC Exome | PASS QC Genome | PASS

Population	AF	AC	HOM
African/American population	6.13e-4	46	0
American population	6.67e-4	40	0
Amish	0.00e+0	0	0
Ashkenazi Jewish	1.52e-3	45	0
East Asian population	0.00e+0	0	0
Finnish population	1.06e-3	67	0
Middle Eastern population	6.61e-4	4	0
Non-Finnish European population	3.84e-3	4526	11
Other populations	1.63e-3	102	1
South Asian population	6.05e-4	55	0
Non-UKB combined population	2.38e-3	1854	5
Combined population	3.03e-3	4885	12

Obrázok 17: Karta Variant na stránke variantov



Obrázok 18: Vizualizácia karty DiagAI na stránke variantov

DiagAI. Karta obsahuje vysvetlenia skóre DiagAI priradeného variante. Táto karta je k dispozícii len pre analýzy GermlineVar, GermlineFamily, GermVar Tertiary a GermVar Tertiary Family (od verzie 1.8).

Karta DiagAI je rozdelená do 4 hlavných častí:

- **Skóre DiagAI** – blok s prehľadom **Germline** zobrazujúci skóre, stav SmartPick a zoznam hlavných charakteristík prispievajúcich k celkovému skóre a príslušný počet bodov.
- Vodopádový graf vysvetľujúci skóre UP² (Universal Pathogenicity Predictor).

Varovná správa  sa zobrazí keď používateľ navštívi stránku variantu, ktorého prepis sa líši od prepisu použitého na výpočet skóre UP².

- Blok vysvetľujúci **mapu PhenoGenius**, farebne označený, ak sú termíny HPO spojené so vzorkou. Blok obsahuje podrobnosti o poradí génov, zoznam 10 génov s najvyššou asociáciou a mapu PhenoGenius, farebne označenú podľa zodpovednosti medzi HPO pacienta a génom: modrá farba označuje iba HPO pacienta, fialová farba označuje HPO pacienta a zodpovedajúci gén, červená farba označuje iba HPO génu. Sekcia je sivá, ak pre pacienta nie sú zadane žiadne termíny HPO.
- Sekcia **Zvýšenie skóre** integruje 5 čiastkových funkcií, ktoré môžu alebo nemusia prispievať (vypnuté) k skóre: Režim prenosu, Kompatibilita dedičnosti, Rodinný vzor, PanelApp a Kontrola kvality.

ACMG. Karta umožňuje používateľom zobrazit' a skontrolovať kritériá použité počas automatizovanej klasifikácie ACMG. Môžu zmeniť výber zaškrtnutím rôznych požadovaných kritérií a potom použiť klasifikáciu na vyhodnotenie variantu kliknutím na „*Použiť na vyhodnotenie*“. Je tiež možné zmeniť úroveň dôkazov (*podporná, stredná, silná a veľmi silná*), aby sa každému kritériu priradil iný počet bodov, ktorý sa zohľadní pri výpočte konečného skóre ACMG. Každá úroveň dôkazov prideli: 1 bod za *podpornú*, 2 body za *strednú*, 4 body za *silnú* a 8 bodov za *veľmi silnú*. Body za každé kritérium sa potom spočítajú, aby sa variante priradil status ACMG:

- *Patogénna*: 10 bodov
- *Pravdepodobne patogénna*: 9 až 6 bodov
- *Neistý význam*: 5 až 0 bodov
- *Pravdepodobne neškodný*: -6 až -1 bod
- *Neškodný*: od -7 bodov

V prípade kritéria BA1 nemôže používateľ vybrať úroveň dôkazu, pretože ide o samostatnú triedu.

TP53 protein coding

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

sample

GT / VAF 1.25%

COV / QA 319 34

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

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).

Obrázok 19: Vizualizácia karty ACMG na stránke variantov

TP53. Táto záložka je k dispozícii len v prípade, ak je variant anotovaný na géne TP53. Táto záložka poskytuje informácie o mutáciách TP53 a ich klinickom význame podľa databázy IARC²⁴. Záložka TP53 je rozdelená do niekoľkých blokov, ktoré sú popísané nižšie.

Blok anotácií (*Anotácie*) zobrazuje nasledujúce informácie súvisiace s variantom:

- Miera mutácie: miera substitúcie spojená s polohou variantu
- Hotspot (áno/nie): prítomnosť variantu v mieste opakujúcej sa mutácie (hotspot)
- Štrukturálny motív: štrukturálny motív ovplyvnený mutáciou
- Doména: funkčná doména proteínu ovplyvnená mutáciou

Blok Funkčné údaje zobrazuje klasifikácie a funkčné informácie súvisiace s variantom:

- Trieda DNE: dominantne negatívny účinok (DNE) na transaktiváciu proteínu p53 divokého typu
- Trieda DNElof: funkčná klasifikácia pre stratu potlačenia rastu a negatívnych dominantných aktivít
- Trieda transaktivácie: funkčná klasifikácia založená na globálnej transkripčnej aktivite (TA)
- Trieda štruktúrnej funkcie: funkčné predikcie.

²⁴ <https://tp53.cancer.gov/>

Bloky „Grafika“ a „Podrobnosti o vzorkách“ poskytujú podrobnosti o vzorkách uvedených v databáze IARC pre danú variantu. Používateľ si môže zvoliť zobrazenie údajov spojených so somatickými alebo zárodočnými vzorkami výberom možnosti Somatické alebo Zárodočné.

Blok „Grafika“ zobrazuje dva grafy s roletovými menu, ktoré ukazujú rozloženie vzoriek podľa ich charakteristík.

Blok „Podrobnosti o vzorkách“ zobrazuje podrobné informácie o vzorkách v tabuľkovej forme.

ComPerMed. (len somatická analýza). Trieda patogenity každej varianty sa vypočíta podľa odporúčaní belgického odborného výboru ComPerMed (Komisia pre personalizovanú medicínu). Používateľ má prístup k podrobným informáciám o tejto klasifikácii na špeciálnej karte na stránke varianty.

Dôkazy. Karta obsahuje zoznam terapií spojených s variantom vybraným z databázy JaxCKB. Používateľ môže prechádzať jednotlivé riadky a nájsť podrobnosti o terapiách alebo indikáciách.

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
Compermed
Evidences
Evaluations
New
Occurrences
Scores
IGV
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
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

Obrázok 20: Vizualizácia karty Dôkazy na stránke variantov

Evaluations (Hodnotenia). Táto karta poskytuje prístup k histórii hodnotení a interpretácií pre variantu v priebehu času:

- V aktuálnej analýze (ľavá časť karty)
- V ostatných analýzách vykonaných v rámci skupiny (pravá časť karty)

Occurrences (Výskyty). Táto karta obsahuje zoznam všetkých ostatných vzoriek, v ktorých sa varianta vyskytuje, bez zohľadnenia výsledkov hodnotenia. Pre každý prípad sú uvedené nasledujúce informácie: názov vzorky, alelová frekvencia, odkaz na príslušnú analýzu a názov projektu.

Publications (Publikácie). Táto karta poskytuje prístup k vedeckej literatúre spojenej s variantom a génom. Kliknutím na podkartu "Variant" alebo "Gene" (Gén) zobrazíte publikácie PubMed odkazujúce na variant alebo gén. Odkaz PMID presmeruje na záznam PubMed a odkaz na plný text otvorí celý článok, ak je k dispozícii.

Scores (Skóre). Karta je k dispozícii pre varianty s efektmi missense, start lost alebo stop gained. S výnimkou týchto prípadov bude skóre sivé. Táto karta umožňuje zobrazíť výsledky predikčných skóre viacerých nástrojov, ako sú *REVEL*, *DANN*, *FATHMM*, *FATHMM-MLK*, *LRT*, *MetaLR*, *MetaSVM*, *MutationTaster 2*, *PROVEAN*, *SIFT* a *PLi*. Poskytuje tiež prístup k skóre konzervácie, ako sú *GERP++*, *PhastCons100way Vertebrate*, *PhastCons20way Mammalian*, *PhyloP100way Vertebrate*, *PhyloP20way Mammalian* a *SiPhy*.

IGV. Táto záložka umožňuje používateľovi zobrazíť čítania zarovnané s polohou variantu prostredníctvom webovej verzie prehliadača genómu IGV. V oblastiach s vysokým pokrytím čítania sú čítania štandardne „podvzorkované“, t. j. nezobrazujú sa všetky čítania. Používajú sa predvolené parametre zobrazenia, t. j. maximálne 100 čítaní na okno s 50 nukleotidmi. Tieto oblasti možno identifikovať podľa čiernej čiary pod profilom pokrytia (downsampling). Kliknutím

na ňu sa zobrazí počet čítaní odstránených z displeja (# downsampled). Z ikony  je k dispozícii ponuka s tromi možnosťami:

- *Otvoriť v novej karte IGV:* otvorí variant zobrazenia IGV v novej karte.
- *Nastavenie stopy:* umožňuje pridať do zobrazenia ďalšie vzorky z rovnakého alebo iného projektu.
- *Open local IGV:* presmeruje aplikáciu IGV Desktop na súradnice varianty, ktorá sa práve interpretuje. Na zobrazenie SNV alebo iných variantov je potrebné najskôr načítať vzorku BAM do aplikácie IGV Desktop. Táto možnosť je kompatibilná iba s verziami IGV Desktop ≥ 17 .

IGV podporuje rozdelené zobrazenia, napríklad kopírovaním dvoch súborov súradníc do vyhľadávacieho poľa lokusu, oddelených medzerou. Kliknutím na ozubené koliesko vpravo od prehliadača genómu sú k dispozícii rôzne nastavenia. Ak chcete napríklad priamo zobrazíť neusporiadané čítania (soft clips), vyberte možnosť „Show soft clips“ (Zobrazíť soft clips).

Gene (Gén). Karta obsahuje ďalšie informácie o géne a súvisiacu bibliografiu.

Vytvorenie klinickej správy

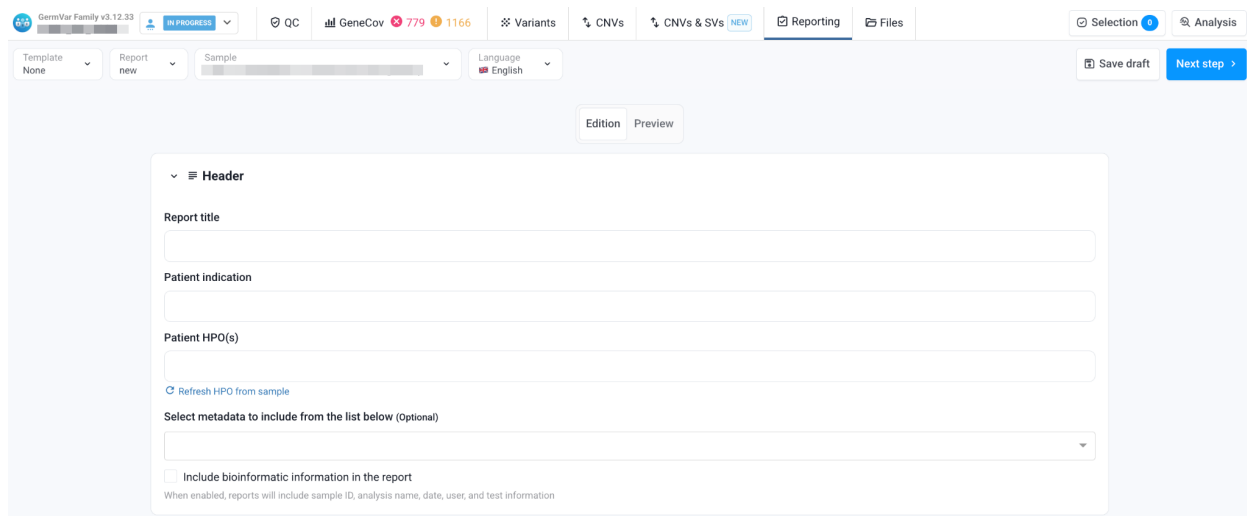
Používateľ môže upravovať exportovateľnú klinickú správu vo formáte PDF a .docx kliknutím na tlačidlo "Generate report" (Generovať správu) na karte "Selection" (Výber) počas navigácie v analýze.

V hornej časti stránky úprav môže používateľ vybrať šablónu, nastaviť stav správy na novú alebo koncept, vybrať vzorku (pre rodinné analýzy je možné vybrať viacero vzoriek) a vybrať jazyk úprav. Používateľ môže tiež prepínať medzi zobrazením "Edition" (Úprava) a zobrazením "Preview" (Náhľad) pred uložením konceptu alebo prechodom na ďalší krok (pomocou tlačidiel "Save Draft" a "Next Step").

Šablóny správ je možné vytvárať v sekcii "Reporting", pod "Template list" v nastaveniach subjektu (Entity settings).

V zobrazení úprav môže používateľ upravovať rôzne časti správy a vyberať viacero možností. V zobrazení náhľadu môže používateľ skontrolovať správu pred pokračovaním v úpravách, uložením konceptu alebo overením správy.

Používateľ môže uložiť koncept a znova k nemu prísť z karty "Selection" a potom "Generate report", rovnako ako pri vytváraní novej správy. Keď sa vygeneruje správa vo formáte PDF, automaticky sa zobrazí na karte "Reporting" (Správy) danej analýzy. Až po overení je možné správu exportovať ako .PDF alebo .DOCX. Používateľ môže správu vymazať alebo vytvoriť ďalšie správy pre tú istú analýzu.



Obrázok 21: Zobrazenie karty klinickej správy.

13.1.5. CNV

13.1.5.1. CNV v paneloch génov

Metodika

Detekcia CNV v údajoch génových panelov v GermlineVar, GermVar/GermVar Family, SomaVar a SomaHemato je založená na výbere kontrolných vzoriek z koanalyzovanej kohorty podľa ich korelácie.

Korelácia sa vypočíta z tabuľky pokrytia, ktorá sa zase vypočíta zo všetkých regiónov v manifeste, ale s vylúčením pohlavných chromozómov.

Kritériá, ktoré musí vzorka spĺňať, aby bola vhodná na analýzu CNV, sú nasledovné:

- Aby bola oblasť zachovaná v analýze CNV, minimálne pokrytie je 50 čítaní.
- Hraničná hodnota korelácie pre výber kontrolných vzoriek je 0,90 pre kontext zárodočných buniek a 0,75 pre kontext somatických buniek.
- Minimálny počet kontrolných vzoriek požadovaných pre každú vzorku je šesť.

Pred spustením CNVCapture alebo SomaCNVCapture je potrebné definovať pohlavie vzoriek (metadáta vzoriek), aby bolo možné vykonať volanie génov lokalizovaných na pohlavných chromozómoch. Nakoniec, na vykonanie analýzy CNVCapture alebo SomaCNVCapture na SeqOne je potrebných najmenej osem vzoriek.

Kritériá uplatňované na manifest používateľa, aby bol panel génov spôsobilý na analýzu CNV, sú nasledovné:

- regióny manifestu sú anotované z kanonického transkriptu určeného podľa zdroja NCBI MANE Select. Ak nie je k dispozícii žiadny transkript MANE Select, anotácia je založená na zdroji NCBI RefSeq.
- Analyzujú sa len exóny, pri ktorých je ≥ 80 % prekryvanie medzi manifestným intervalom a exónom. Ak sa žiadny génový exón neprekryva ≥ 80 % s manifestným intervalom, gén nebude zahrnutý do analýzy CNV.
- Analyzujú sa iba intróny, pri ktorých je ≥ 60 % prekryvanie medzi manifestným intervalom a intrónom.

Analýza CNVCapture (verzia ≥ 2.4) zahŕňa oblasti nachádzajúce sa 20 kb nad a pod génmi pokrytými manifestom, ak sú takéto oblasti zahrnuté v návrhu manifestu. Konkrétne sa analyzuje každý z 20 segmentov s veľkosťou 1 kb, ktoré tvoria oblasť s veľkosťou 20 kb, ak sa aspoň 20 % segmentu prekryva s počiatočným manifestom.

Zobrazenie a interpretácia

Na karte CNV sa zobrazí graf predstavujúci zoznam génov, v ktorých boli identifikované variácie kópií. Táto karta je k dispozícii v pracovných súboroch opísaných v časti Metodika vyššie.

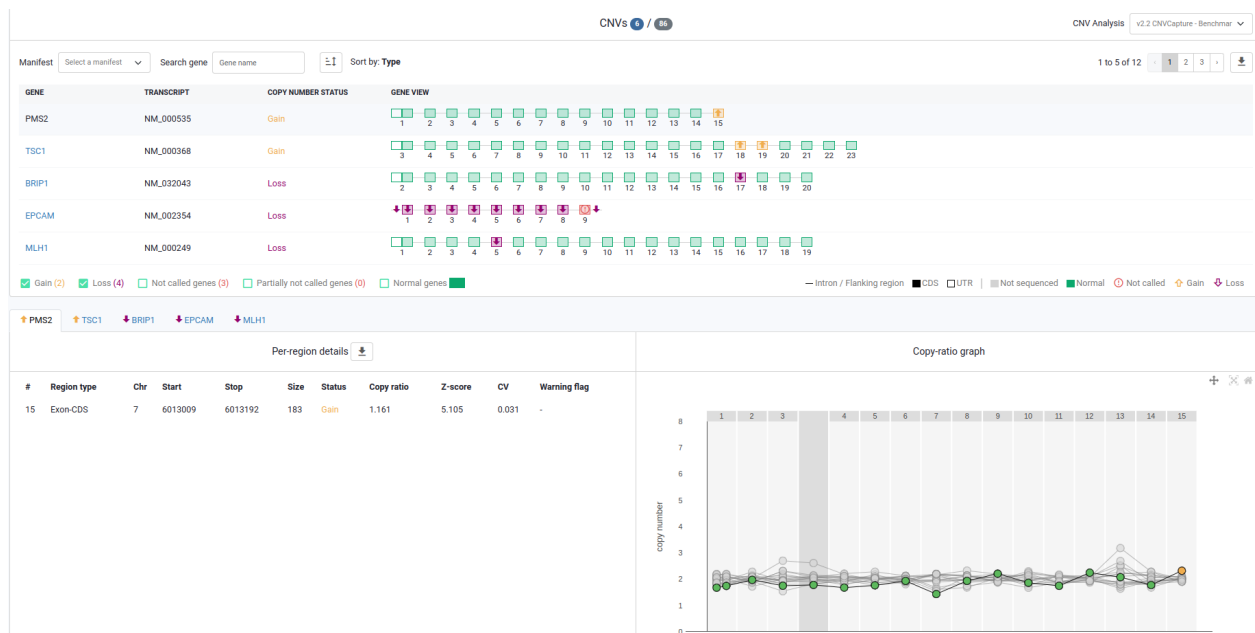
Terminológia používaná na opis stavu identifikovaných génov je nasledovná:

- Zisk: gén je ovplyvnený čiastočnou alebo úplnou amplifikáciou
- Strata: gén nesie čiastočnú alebo úplnú deléciu
- Normálny: gén nenesie žiadnu CNV
- Čiastočne nevyvolaný: gén je čiastočne vyvolaný
- Nenazvaný: gén je vylúčený z analýzy CNV

Stav čiastočne neoznačený sa priradzuje génu, ak je stav jeho kópie normálny, ale aspoň jeden exón nebol označený (so stavom „neúspešný“).

Napriek tomu, že je gén vhodný na analýzu CNV, je z analýzy vylúčený (neoznačený), ak:

- je gén na pohlavnom chromozóme a metadáta o pohlaví neboli vyplnené,
 - všetky exóny zlyhajú podľa kritérií uvedených nižšie: počet kontrol, percento variácie.
- Stav „neúspešný“ génu je definovaný iba na CDS. Intronové a upstream/downstream regióny sú vylúčené, ak sú volané exóny CDS.



Obrázok 22: Vizualizácia karty CNV v analýze GermlineVar.

Filtry a možnosti triedenia

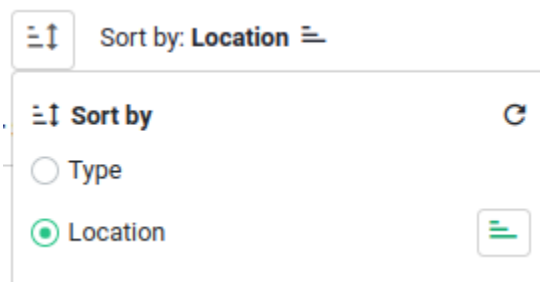
Ponuka filtrov môže byť použitá na:

- Obmedzenie zobrazenia na podmnožinu regiónov/génov pomocou súboru BED, ktorý bol predtým nahraný do rozhrania pomocou funkcie „Manifest“
- Vyhľadávanie konkrétneho génu pomocou funkcie „Search gene“ (Vyhľadávanie génu)

K dispozícii sú dve možnosti triedenia CNV:

- **Typ** (predvolené): aktuálna logika triedenia (zisk, potom strata)
- **Poloha**: slúži na triedenie CNV podľa chromozómov a ich počiatkovej polohy, v ascendentnom alebo descendentnom poradí,

kliknutím na ikonu triedenia



Používateľ môže tiež vybrať gény podľa ich stavu: Zisk, Strata, Normálne gény, Čiastočne neoznačené gény alebo Neoznačené gény.

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

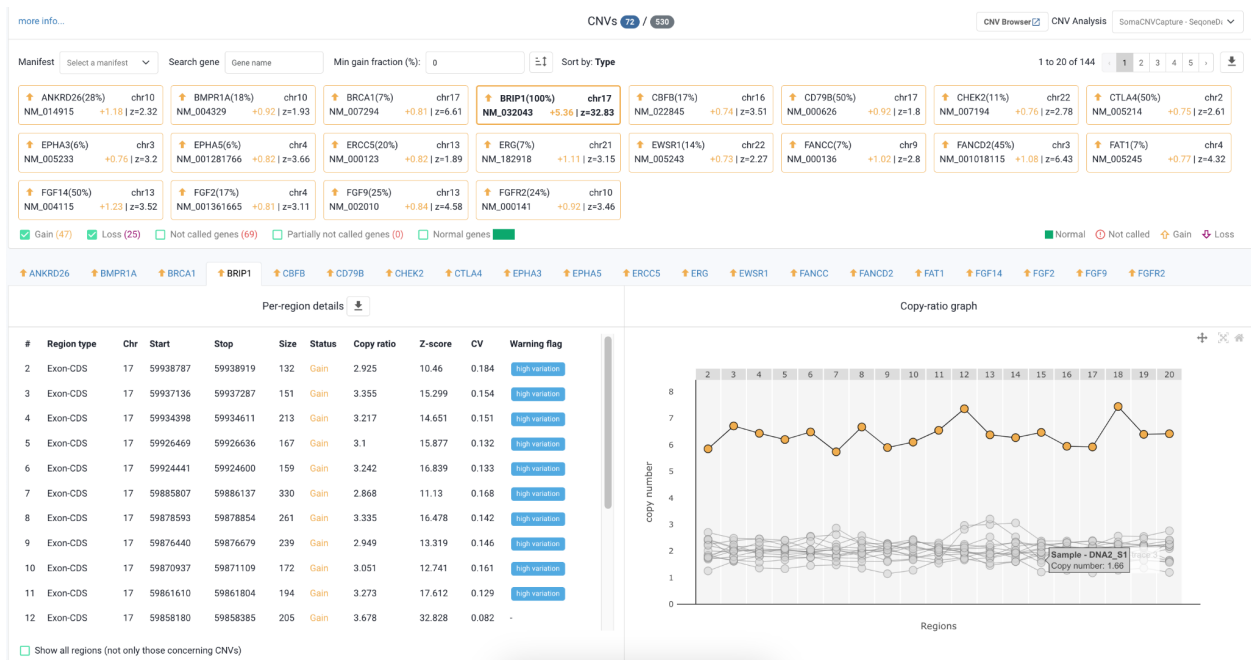
Podrobnosti podľa regiónu

Tabuľka Podrobnosti podľa regiónu zobrazuje podrobnosti pre každý región pre každý gén:

- **#**: číslo regiónu (exóny a intróny)
- **Typ regiónu**: typ regiónu (upstream, utr-exonic, coding-exonic, intronic, downstream)
- **Chr**: chromozóm
- **Začiatok/koniec**: počiatková a koncová pozícia CNV
- **Veľkosť**: veľkosť CNV (bp)
- **Stav**: stav regiónu (zisk = duplikácia, strata = delécia, normálny = bez CNV, neoznačený = región nebol zahrnutý do analýzy CNV)
- **Pomer kópií**: pomer počtu kópií k normálu (n=2).
- **Z-skóre**: počet štandardných odchýlok oddeľujúcich výsledok od priemeru. Vysoké kladné alebo záporné hodnoty naznačujú vysokú spoľahlivosť identifikácie CNV.
- **CV**: koeficient variácie merajúci relatívnu disperziu v rámci vzoriek kohorty.
- **Varovná vlajka**: Podľa nasledujúcich kritérií sa môže občas zobraziť chybová správa:
 - *Low zscore (Nízky z-skóre)*: absolútna hodnota Z-skóre <4
 - *High variation (Vysoká variabilita)*: koeficient variácie (CV) > 0,1
 - *Low coverage (Nízke pokrytie)*: pokrytie regiónu pre túto vzorku <100
 - *Number of controls (Počet kontrol)*: Počet kontrol nie je dostatočný na to, aby bolo možné vykonať volanie v tejto oblasti

- **Percentuálna variabilita:** Variabilita medzi vzorkami je príliš veľká na to, aby bolo možné vykonať volanie pre túto oblasť
- **Hĺbka čítania:** nedostatočné pokrytie regiónu (<50) a pomer kópií >0,1
- **Chýbajúce pohlavie:** oblasť sa nachádza na pohlavnom chromozóme a metadáta o pohlaví neboli zadávané.

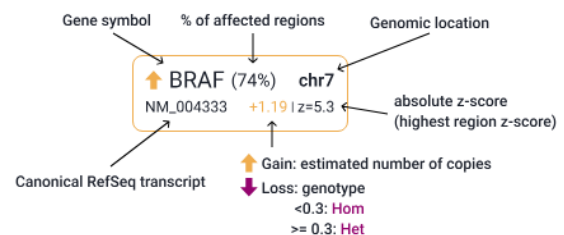
Upozornenie: oblasť bohatá na mäkké klipy, ako napríklad veľký indel alebo Alu insert, môže byť interpretovaná ako pokles pokrytia pracovnými sadami CNVCapture a SomaCNVCapture a detegovaná ako strata kópie, ak je veľkosť vloženia rovnaká ako veľkosť volanej oblasti.



Obrázok 23: Vizualizácia karty CNV v analýze SomaVar

V prípade analýz **SomaCNVCapture** je metóda detekcie podobná metóde CNVCapture, avšak mierka použitá na grafické znázornenie variantov je gén, nie exón. Znázornenie je zobrazené ako séria políčok. Tá obsahuje symbol génu, genomickú pozíciu, kanonický Refseq transkript, absolútny z-skóre a v prípade zisku počet kópií alebo v prípade straty genotyp.

Nad filtrami v sekcii „viac informácií...“ je k dispozícii vyskakovacie okno s vysvetlením týchto políčok.

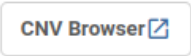


Percentuálny podiel zobrazený v každom poli odráža podiel génu (minimálny podiel zisku) ovplyvnený detegovanou CNV. Pomocou pomeru častí je možné filtrovať výsledky. Predvolene je filter nastavený na 0 % a vzťahuje sa iba na zisky kópií.

Delécie zistené týmto postupom nepodliehajú tomuto filteru a mali by sa interpretovať s opatrnosťou. Schopnosť tohto nástroja zistiť variácie kópie génu závisí od povahy a kvality vzorky, najmä od frakcie nádoru.

Upozornenie: Ak sa na rovnakom géne zistí amplifikácia aj delécia, gén sa štandardne považuje za amplifikovaný, a preto podlieha filteru minimálneho zisku opísanému vyššie.

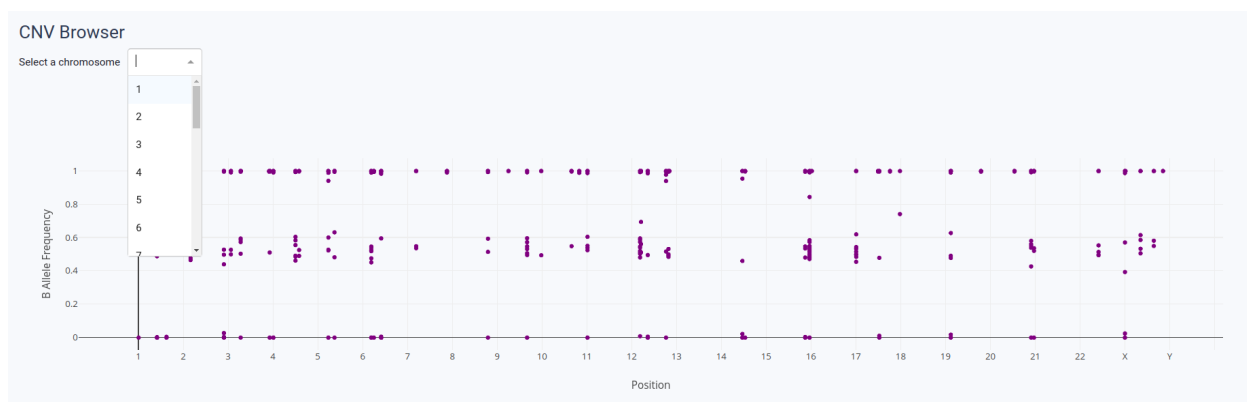
Prehliadač CNV

Ikona  je k dispozícii na karte zobrazenia CNV v analýzach GermlineVar, GermlineFamily, SomaVar alebo SomaHemato, kde otvorí prehliadač CNV v novej karte. Graf prehliadača CNV zobrazuje frekvenciu alely B (BAF) SNP (jednonukleotidový polymorfizmus) prítomných v analyzovanej vzorke a prítomných v súbore údajov SNP polí od Illumina a SNP z iných chrbtových kostí, ako je Agilent's OneSeq, aby sa uľahčila vizualizácia určitých typov udalostí, ako je LoH (strata heterozygotnosti).

BAF je frekvencia nereferenčnej alely (alela B) pozorovaná pre heterozygotný SNP v zárodočnej línii.

Graf štandardne zobrazuje BAF variantov (os y) v celom genóme (os x), rozdelených podľa chromozómov (1–22, X a Y). Filter umožňuje používateľovi zobrazíť varianty na danom chromozóme v ľavom hornom rohu.

V kontexte analýz CNV sa tieto informácie BAF môžu použiť na zobrazenie zmeny v počte kópií špecifických pre alelu, straty heterozygotnosti alebo alelickej nerovnováhy.



Obrázok 24: Vizualizácia karty prehliadača CNV v analýze GermlineVar.

Export údajov CNV

Ikonu  (Exportovať údaje) v pravom hornom rohu môžete použiť na export údajov analýzy CNV vo forme súboru .tsv, ktorý obsahuje nasledujúce stĺpce:

- gén (gene_name),
- transkript RefSeq (refseq_id),
- stav génu (status),
- chromosóm (chr),
- genómové súradnice CNV: počiatočná a koncová pozícia,
- časť génu ovplyvnená CNV (gene_fraction_cnv),
- pomer počtu kópií (copy_ratio) k normálu (n=2),
- počet štandardných odchýlok oddeľujúcich výsledok od priemeru (zscore)
- počet exónov prekrývajúcich CNV (exons) v prípade analýzy CNVCapture.

Pre každý gén je možné pomocou ikony „“ vedľa položky „Per-region details“ (Podrobnosti o regióne) exportovať údaje analýzy CNV pre vybraný gén vo forme súboru .tsv, ktorý obsahuje nasledujúce polia:

- gén (gene_name),
- chromosóm (chr),
- genómové súradnice: počiatočná a koncová pozícia,
- číslo exónu alebo intrónu (exon_intron_number),
- stav,
- typ (exón alebo intrón, type),
- súvisiace chybové správy (warning_flags, fail_flags),
- upravená kritická hodnota pravdepodobnosti (padj),
- kritická hodnota pravdepodobnosti (pvalue),
- pomer počtu kópií k normálu (pomer),
- mieru relatívnej disperzie v rámci vzoriek (variabilita),
- štandardná odchýlka výsledku od priemeru (zscore)

13.1.5.2. CNV v SomaCGP a SomaLBx Analýza

Metodika

SomaCGP a SomaLBx používajú PureCN²⁵ na detekciu CNV a LOH (strata heterozygotnosti). PureCN identifikuje somatické varianty z údajov o detekcii variantov s cieľom posúdiť ploidiu a čistotu nádoru vo vzorke. Tieto informácie sa používajú na určenie počtu kópií každej udalosti.

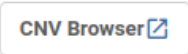
Vizualizácia a interpretácia

Karta vizualizácie CNV v terciárnej analýze SomaVar funguje rovnako ako v SomaCGP a SomaLBx, ako je opísané nižšie.

CNV sú zobrazené v tabuľke, ktorú je možné filtrovať. Každá CNV zobrazuje nasledujúce informácie:

- **Gén:** názov génu
- **Umiestnenie:** cytoband
- **Súradnice udalosti / Veľkosť udalosti:** súradnice udalosti a veľkosť
- **Transkript:** referenčný transkript
- **Exóny:** ovplyvnené exóny
- **Typ:** typ udalosti (LoH, strata alebo zisk)
- **Počet kópií:** absolútny počet kópií
- **Funkcia:** funkcia génu, ak je to relevantné, onkogén alebo tumor supresorový gén (TSG)

Používateľ môže vyhľadávať CNV ovplyvňujúce daný gén prostredníctvom vyhľadávacieho poľa a zvoliť si, či sa majú zobrazovať gény prítomné v LOH regiónoch, zaškrtnutím políčka „Zobraziť gény v LOH regiónoch”. Tabuľku výsledkov je možné exportovať pomocou ikony na stiahnutie v pravom hornom rohu.

Ikona  (Prejsť na LOH) otvorí v novej karte CNV Browser. Táto karta zobrazuje graf znázorňujúci frekvenciu alely B (BAF) SNP v analyzovanej vzorke a LOH SNP z vybraných publikácií a OneSeq Agilent backbone. Druhý graf znázorňuje údaje logRatio generované PureCN. Viac podrobností nájdete v časti **CNV: Gene Panels (CNV: Génové panely) pod CNV browser**.

²⁵ <https://github.com/lima1/PureCN>

13.1.5.3. CNV v exónoch a genónoch

Karta vizualizácie CNV pre exómové údaje je k dispozícii v analýzach GermlineVar, GermlineFamily, GermVar a GermVar Family pri použití súborov FASTQ ako primárneho vstupu a spustení analýzy CNVExome. Táto karta tiež zobrazuje výsledky CNV pre genómové údaje v analýzach GermVar a GermVar Family s použitím súborov FASTQ, ako aj v analýzach GermVar Tertiary a GermVar Tertiary Family s použitím súborov VCF alebo údajov z mikroatray CGH.

Metodika pre exóm

Táto metodika sa vzťahuje konkrétne na nástroj CNVExome. Rovnako ako v prípade analýzy CNV na údajoch génového panelu, proces identifikácie variantov v tomto prípade závisí od analýzy hĺbky pokrytia. Analýza CNVExome pomocou nástroja [GATK GermlineCNVCaller](#) je však individuálna, pretože každá vzorka sa porovnáva s modelom HHMM (*Hierarchical Hidden Markov Model*) vypočítaným nástrojom SeqOne.

Aby sa zohľadnili technické odchýlky pri príprave vzoriek, tento model je založený na veľkej kohorte od jednej do niekoľkých stoviek vzoriek sekvenovaných laboratóriom pomocou:

- Rovnakého protokolu prípravy knižnice
- Rovnakú sadu na zachytávanie
- Rovnakú sekvenčnú platformu a dĺžku čítania

Tieto požiadavky a variabilný počet vzoriek potrebných na vytvorenie modelu možno vysvetliť potrebou kohorty, ktorá ponúka homogénnu úroveň pokrytia. Je potrebné zohľadniť viacero faktorov, ako napríklad hĺbku sekvenovania, stupeň variability v rámci kohorty a reprodukovateľnosť protokolov prípravy vzoriek.

Pred začatím analýzy CNVExome je potrebné vyplniť metadáta o pohlaví (pozri **Metadáta vzorky**), aby bolo možné overiť ploidiu pohlavných chromozómov. Ak nie sú vyplnené, na rozhraní CNV sa zobrazí varovná správa „Ploidy“.

Obmedzenie: Dokumentácia GATK GermlineCNVCaller odporúča vylúčiť oblasti PAR (pseudoautozomálne oblasti), t. j. oblasti PAR1 a PAR2, ak je ich príspevok k celkovému pokrytiu chromozómu X neprimeraný. V skutočnosti je ploidia ťažko vypočítateľná pre konce chromozómu X zodpovedajúce oblastiam PAR pre mnoho modelov vypočítaných programom SeqOne. V dôsledku toho sa ploidia pre tieto oblasti, ktoré sú umelo detegované ako 1 kópia, nepočíta správne, zatiaľ čo celý chromozóm X je detegovaný ako 2 kópie.

Používatelia môžu kontaktovať tím podpory používateľov na adrese support@seqone.com, aby sa dozvedeli viac o postupe nastavenia CNVExome vo svojej skupine.

Metodika pre genóm

Pre verzie až do v3.6 analýzy GermVar a GermVar Family na genómových údajoch integrujú nástroj Sentieon CNVScope²⁶ na detekciu CNV. Od verzie v3.7 pracovné sady GermVar a GermVar Family používajú nástroj Cue²⁷ na volanie CNV na genómových údajoch.

Údaje o segmentálnej duplikácii sa detegujú pomocou Sentieon's SegDup Caller, ktorý je špecificky navrhnutý na identifikáciu udalostí počtu kópií vo vysoko homológnych genomických oblastiach na údajoch WGS iba pomocou GermVar. Tento vyvolávač sa zameriava na vybraný súbor klinicky relevantných génov, o ktorých je známe, že sú náročné pre štandardné prístupy detekcie variantov. Výsledné udalosti sú automaticky integrované do analytického potrubia a vizualizované v prehliadači CNVs & SVs ako CNV (zisk aj strata), čo umožňuje identifikáciu variantov v genomických oblastiach, ktoré boli predtým pre štandardnú analýzu neprístupné.

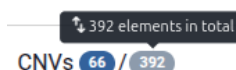
Zobrazenie a interpretácia

Interpretácia je k dispozícii na karte **CNV** pre analýzy CNVExome, GermVar a GermVar Family, GermVar Tertiary alebo GermVar Tertiary Family. Rozhranie je rozdelené do dvoch častí: tabuľka výsledkov CNV vľavo a podrobná stránka CNV s viacerými kartami vpravo.

Tabuľka výsledkov CNV

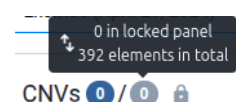
Tabuľka výsledkov zobrazuje identifikované CNV spolu s ponukou filtrov na základe ich prítomnosti v morbidných génoch OMIM [<https://omim.org/>], veľkosti, frekvencie výskytu CNV v kohorte alebo LogRatio v prípade analýzy MicroArray CGH.

Počet filtrovaných CNV (vybrané CNV) a celkový počet detegovaných CNV (celkový počet prvkov) sa zobrazujú v hornej časti tabuľky CNV.



392 elements in total
CNVs 66 / 392

Ikona visiaceho zámku označuje, že analýza používa obmedzený panel, a podrobnosti o CNV sa zobrazujú v príslušných nápovedách.



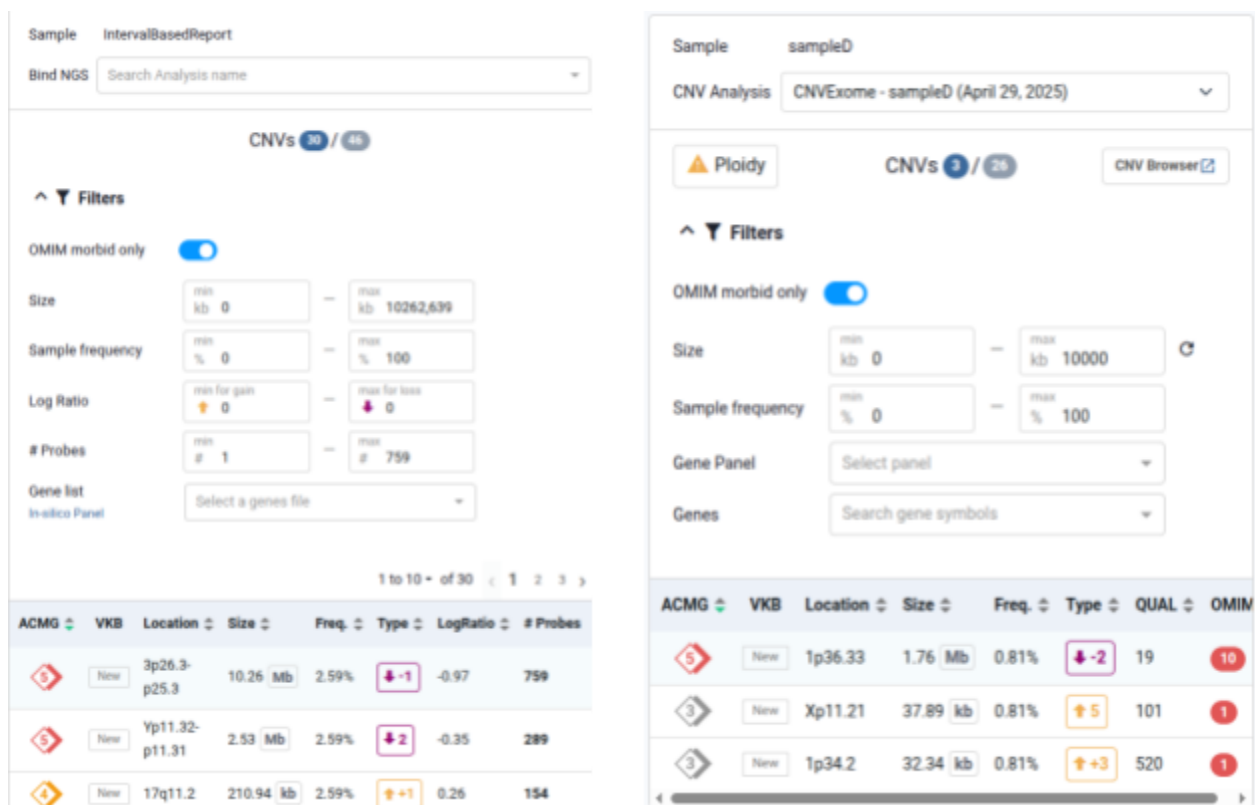
0 in locked panel
392 elements in total
CNVs 0 / 0

Varovná správa „ Ploidy “ (Ploidy je neplatná) sa zobrazí, ak sa zistí abnormálna hodnota ploidie a/alebo ak pred spustením analýzy neboli vyplnené metadáta o pohlaví (pozri **Metadáta vzorky**).

²⁶ <https://support.sentieon.com/appnotes/cnvscope/>

²⁷ <https://github.com/PopicLab/cue>

Kliknutím na ikonu „Ploidy“ môže používateľ zobrazíť ploidiu každého chromozómu a pohlavie pacienta. Abnormálne hodnoty sú zvýraznené žltou farbou.



Obrázok 25: Prehľad detegovaných CNV a filter menu analýzy CGH (vľavo) a štandardnej analýzy (vpravo).

V prípade analýzy CGH MicroArray v kombinácii s SNP array (CGH+SNP) sa zistené udalosti straty heterozygotnosti (LOH) uvádzajú v tabuľke výsledkov CNV. Tieto udalosti zodpovedajú strate heterozygotnosti bez zmeny počtu kópií (CN-LOH). Inými slovami, oblasť si zachováva dve kópie, ale tieto sú identické, čo sa vyznačuje prítomnosťou homozygotných SNP.

V prípade analýzy CGH MicroArray je možné z roletového menu „Bind NGS“ vybrať analýzu non-CGH SNV (GermlineVar, GermlineFamily, GermVar Tertiary a GermVar Tertiary Family), aby sa zobrazili varianty, ktoré sa prekrývajú s detegovaným CNV.

Filtry

OMIM morbid only. Tento filter je štandardne aktivovaný, aby bolo možné ľahko identifikovať CNV ovplyvňujúce klinické gény, anotované OMIM Morbid na základe ich identifikátora NCBI (NCBI gene_id) alebo, ak to nie je možné, na základe názvu génu a synonym. Aby sa zobrazili všetky CNV detegované vo vzorke, stačí tento filter zrušiť.

Veľkosť. Používateľ môže filtrovať CNV podľa ich očakávanej veľkosti (v kb).

Frekvencia vzorky. Používateľ môže filtrovať CNV podľa ich veľkosti a frekvencie. Táto frekvencia zodpovedá výskytu CNV medzi všetkými vzorkami analyzovanými do dátumu analýzy a porovnanými s rovnakým modelom, vrátane vzoriek použitých na vytvorenie modelu. Frekvencia vzoriek je založená na 95 % prekryvaní a zohľadňuje rovnaký typ CNV (zisk/strata/vypustenie). Po prejdení kurzorom myši nad hodnotou zobrazenou v percentách sa používateľovi zobrazí príslušný podiel. Aktualizácia frekvencie po otvorení analýzy môže trvať niekoľko minút.

Panel génov. Podobne ako panel in silico, tento filter umožňuje používateľovi vybrať a použiť zoznam génov na filtrovanie CNV detegovaných v analýze pomocou roletového menu (pozri **Správa parametrov používateľa**). Tento filter nie je k dispozícii, ak bol na analýzu použitý obmedzený panel z karty Panel v bočnom paneli správy analýzy (pozri časť **Modálne zobrazenie panelu**).

Gény. Umožňuje filtrovať detegované CNV obsahujúce gény vybrané prostredníctvom tohto filtra.

Logaritmický pomer. Pri analýzach založených na údajoch MicroArray CGH sa logaritmický pomer (medzi surovým signálom a kontrolným signálom) prítomný vo vstupných súboroch Agilent zobrazuje v dodatočnom stĺpci. Súvisiaci filter umožňuje používateľovi vybrať minimálne hodnoty pre zisky kópií (minimálna hodnota zisku) a maximálne hodnoty pre straty kópií (maximálna hodnota straty). Udalosti LoH (strata heterozygotnosti) zostávajú zachované aj po použití tohto filtra.

#Probes (Sondy). Pre analýzy založené na údajoch CGH MicroArray stĺpec #Probes (Sondy) uvádza počet sond použitých v každej oblasti (exónovej, intergénovej alebo intrónovej). Tieto hodnoty sú prevzaté zo vstupného súboru IntervalBasedReport.xls a príslušný filter umožňuje používateľovi vybrať minimálny a maximálny počet použitých sond. V kombinovanej analýze CGH+SNP počet sond označuje počet homozygotných SNP v oblasti LOH.

Triedenie CNV

Zoznam CNV je možné triediť aj podľa triedy patogenity ACMG (ACMG), genómovej pozície (Location), veľkosti (Size), frekvencie vzorky (Freq.), počtu kópií (Type), skóre kvality (QUAL) v GermVar Tertiary a podľa LogRatio v CGH MicroArray (LogRatio).

ACMG. Trieda patogenity každej varianty sa vypočíta poloautomaticky podľa odporúčaní *Americkéj akadémie lekárskej genetiky a genomiky*. Trieda patogenity je predvolený režim triedenia v tabuľke CNV.

Udalosti LoH (strata heterozygotnosti) v CGH MicroArray sú vždy priradené do triedy 3 (VUS).

Typ (počet kópií). Zobrazená je variácia počtu kópií (napr.: +2, -1, +3) pre CNV detegované na autozomálnych chromozómoch (1 až 22), keďže počet kópií je predvolene nastavený na 2.

Pre CNV detegované na gonosómoch (pohlavné chromozómy X a Y) sa však zobrazuje absolútny počet kópií (napr. 0, 1, 2) a nie násobná zmena.

Na rozdiel od analýzy génového panelu nie je potrebné, aby metadáta vzoriek súvisiace s pohlavím obsahovali pohlavné chromozómy pri určovaní CNV exómu.

Pri kombinovaných analýzach CGH+SNP sú udalosti CN-LOH, ktoré zachovávajú dve kópie, v stĺpci Typ označené odznakom  .

QUAL. Skóre kvality Phred sa v CNVExome vypočíta pomocou GATK caller. Ak je CNV zložená z viacerých genomických intervalov, CNV sa priradí najvyššie skóre QUAL spomedzi všetkých intervalov. V prípade terciárnych analýz sa zobrazuje skóre QUAL získané zo vstupného súboru VCF (DRAGEN alebo ONT).

Prehliadač CNV

Ikona otvorí prehliadač CNV v novej karte, kde sa zobrazuje frekvencia alely B (BAF) SNP (jednonukleotidový polymorfizmus) prítomných v analyzovanej vzorke. Graf je popísaný v predchádzajúcej časti (CNV: Génové panely, pod prehliadačom CNV).

Podrobná stránka CNV

Súhrn CNV. Táto počiatočná karta obsahuje množstvo informácií a poznámok o vybranom CNV, vrátane:

- Podrobnosti o automatickej klasifikácii ACMG
- Klinické oblasti a gény zahrnuté v intervale ovplyvnenom variantom. Anotácia klinických oblastí je založená na databázach [ClinGen](#) a [OMIM](#). Predvolene sa zobrazujú len klinické gény s anotáciou OMIM Morbid. Ak chce používateľ zobrazíť všetky klinické gény, môže aktivovať funkciu „Zobraziť všetky gény“ v pravej časti stránky.
- Prekrývanie medzi detegovaným CNV a databázou [DGV](#).
- Malé varianty (SNV a indely) detegované v génoch ovplyvnených CNV. V predvolenom nastavení je aktivovaný filter SeqOne-diag. Ak chcete tento filter deaktivovať, kliknite na

SeqOne-diag a zrušte výber SeqOne-diag z profilov filtrov. Ak chcete zobrazit' všetky varianty, môžete aktivovať funkciu „Zobraziť všetky varianty“ v pravej časti stránky.

- Špeciálna sekcia pre znalostnú bázu variantov (VKB) na identifikáciu variantov s identitou nad 95 %, ktoré už boli validované v skupine, a ich príslušné hodnotenia.

ACMG Classification (Automatic) ⓘ
4 Likely Pathogenic (+0.9)

Genomic content
1A 0

Genes or Genomic regions
2C_1 +0.9

Gene number
3A 0

Literature and databases
No selected tags, only 40 is assessed.

Inheritance patterns
Not assessed

Clinical regions

Region name	Overlap	Link	Description
CHROMOSOME 16p13.3 DELETION SYNDROME, PROXIMAL	0.29%	OMIM	
CHROMOSOME 16p13.3 DUPLICATION SYNDROME	0.29%	OMIM	
ALPHA-THALASSEMIA/IMPAIRED INTELLECTUAL DEVELOPMENT SYNDROME, DELETION TYPE	0.29%	OMIM	

Clinical genes
Show all genes

Gene symbol	Databases	Phenotype(s)	ClinGen haploinsufficiency score	pLI (GnomAD)	Exons	Fraction overlap	Variants
PKD1	omim	[AD] Polycystic kidney disease, ad...	3 (Sufficient evidence for dosage pathogenicity)	0.12051	12/46	48.39%	

DGV overlaps

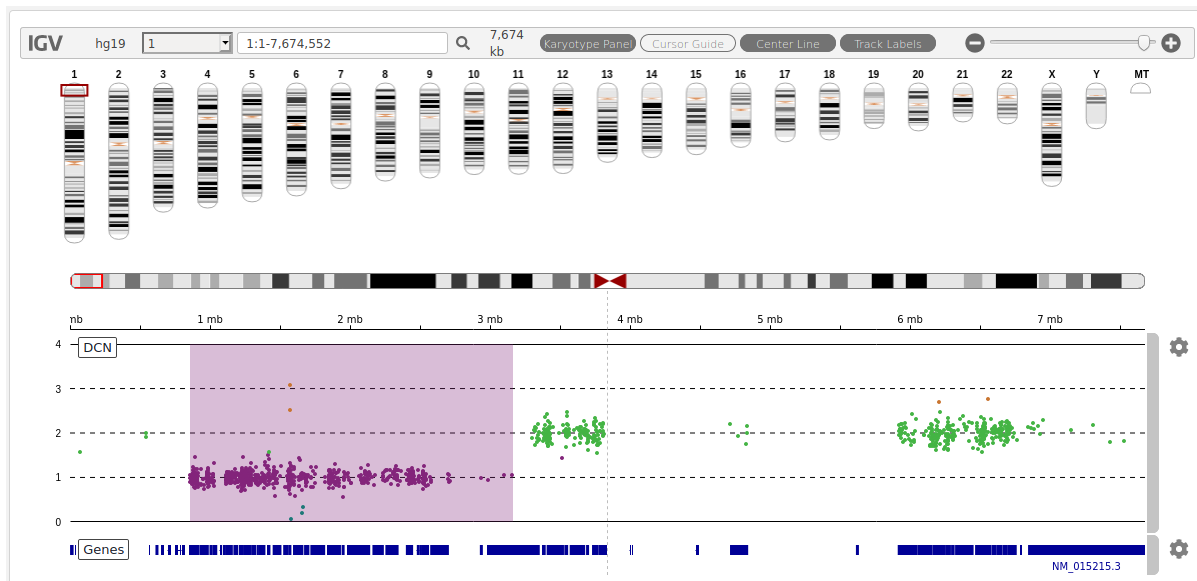
4 Overlapping CNV		100% Max DGV overlap		1.54% Corrected frequency		26 Samples		11 Studies	
ID	Fraction overlap	CNV type	CNV size	Localisation	Frequency	# Samples	# Studies		
gssvG13891	100%	Gain	51.76kb	16:2,138,711-2,190,473	20.24%	17	5		
gssvG13898	88.78%	Gain	25.42kb	16:2,165,630-2,191,053	0.13%	2	2		
gssvL42059	44.04%	Loss	25.71kb	16:2,147,425-2,173,136	100%	2	2		
gssvL42061	32.63%	Loss	10.27kb	16:2,178,475-2,188,740	7.46%	5	2		

Variants
SeqOne-diag 0 / 51
Show all variants

No variant found

Obrázok 26: Prehľad karty súhrnu CNV

- Prehliadač genómu IGV, kde je možné vizualizovať CNV (nie je k dispozícii v analýzach GermVar Tertiary a GermVar Tertiary Family). Na overenie stability signálu je možné zobraziť až päť ďalších pacientov okrem analyzovaného jedinca z konfiguračného menu grafu zobrazujúceho počet kópií (DCN)  .



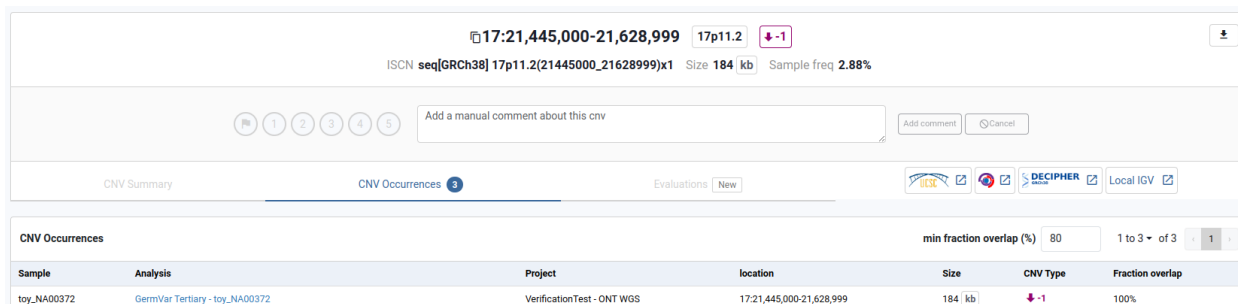
Obrázok 27: Zobrazenie CNV z integrovaného prehliadača IGV

Farebný kód je pre vybraného pacienta nasledovný: modrá (0 až 0,5 kópie), fialová (0,5 až 1,5 kópie), zelená (1,5 až 2,5 kópie), oranžová (2,5 až 3,5 kópie) a červená (>3,5 kópie). Ostatné vzorky sa zobrazujú v šedej farbe.

Výskyt CNV. Táto karta sa používa na identifikáciu ďalších vzoriek obsahujúcich CNV, nezávisle od akéhokoľvek hodnotenia. Pre každý výskyt sa zadajú nasledujúce informácie:

- názov vzorky (Vzorka),
- názov analýzy a odkaz na ňu (Analýza),
- názov projektu (Project),
- genómová poloha CNV (location),
- veľkosť CNV (Veľkosť),
- typ CNV,
- frakčné prekryvanie.

Výsledky sú zoradené v zostupnom poradí podľa percenta prekryvania. Zobrazené CNV je možné filtrovať podľa minimálnej frakcie prekryvania vyplnením pol'a „min fraction overlap (%)“ (minimálna frakcia prekryvania) v hornej časti zoznamu. Predvolené nastavenie tohto filtra je 80 %.



Obrázok 28: Prehľad karty výskytu CNV v analýze CNVExome

Hodnotenia. Klasifikáciu patogenity pre každú CNV môže definovať používateľ, spolu s interpretáciami vo forme komentárov. Každé nové hodnotenie tiež pomáha naplniť osobnú znalostnú bázu laboratória.

UCSC. Externý odkaz na prehliadač genómu [UCSC](#) umožňuje používateľom zobraziť genómovú oblasť CNV. Táto stránka sa otvorí v ďalšej karte prehliadača a obsahuje osobné nastavenia používateľa pre daný prehliadač, čím ponúka personalizované zobrazenie.

CNV-Hub. Externý odkaz na klasifikátor [CNV-Hub](#) vám umožňuje vizualizovať klasifikáciu CNV s ohľadom na viacero klasifikačných nástrojov. Odkaz sa otvorí v ďalšej karte prehliadača a integruje osobné nastavenia prehliadača používateľa, čím ponúka personalizovanú vizualizáciu.

DECIPHER Genome Browser. Externý odkaz na [DECIPHER](#) Genome Browser vám umožňuje zobraziť CNV na genóme. Odkaz otvorí ďalšiu kartu prehliadača a integruje osobné nastavenia prehliadača používateľa, čím ponúka personalizovanú vizualizáciu.

Local IGV. Tento odkaz presmeruje aplikáciu IGV Desktop²⁸ na súradnice CNV, ktoré sa práve interpretujú. Na zobrazenie CNV je potrebné mať vopred načítaný vzorový BAM v aplikácii IGV Desktop. Táto možnosť je kompatibilná iba s verziami IGV Desktop ≥ 17 .

Okrem súborov zarovnaní sú v analýze CNVExome k dispozícii na stiahnutie rôzne súbory v karte *Súbory*:

- Výsledky analýzy CNV vo forme tabuľkového súboru (cnv.tsv)
- Ploidia pre každý chromozóm (contig_ploidy.tsv)
- Predpokladaný počet kópií pre každý exón v manifeste (denoise.tsv)

²⁸ <https://igv.org/doc/desktop/>

Export údajov CNV

Ikona „“ (Exportovať údaje) v pravom hornom rohu umožňuje exportovať súbor TSV obsahujúci nasledujúce údaje o príslušnom CNV:

- jeho hodnotenie VKB a komentár,
- hodnotenie ISCN,
- Klasifikácia ACMG,
- chromosóm, počiatočná a koncová pozícia a cytoband,
- veľkosť a typ CNV,
- počet kópií
- počet kópií
- frekvencia vzorky
- skóre kvality (Phredova stupnica)
- Log2Ratio
- počet použitých sond (Probes)
- počet génov ovplyvnených CNV.

Informácie o príslušných génoch sú uvedené v nasledujúcich stĺpcoch:

- názov génu (Gene_Symbol),
- fenotyp OMIM (OMIM_Phenotypes),
- pravdepodobnosť intolerancie génu na LoF (Pli_Score)
- skóre haploinsuficiencie a triplosenzitivity ClinGen (HI_Score, TS_Score)
- počet exónov prekrývajúcich sa s CNV (Nb_overlap_exon),
- celkový počet exónov v géne (Nb_total_exon).

Ak sa určité popisné polia týkajú len určitých CNV (napríklad v prípade MicroArray CGH), tieto polia majú hodnotu „NA“.

13.1.6. Prehliadač veľkých variantov pre CNVs a SVs (LVV)

Prehliadač veľkých variantov (Large Variant Viewer, LVV) umožňuje používateľovi vizualizovať a analyzovať CNVs a štruktúrne varianty (SVs) na jednej karte. Rozhranie sa skladá z centrálnej tabuľky variantov, panelu filtrov naľavo a bočného panelu, ktorý sa otvorí napravo po výbere variantu, podobne ako na karte variantov (časť Malé varianty).

V kompatibilných analýzach sa prehliadač nachádza na karte s označením "CNVs & SVs" (CNVs a SVs) alebo "SVs", v závislosti od typu údajov:

- Karta CNVs & SVs zobrazuje CNVs (Strata/Zisk), SVs (Delécia/Duplikácia) a udalosti LOH/UPD.
- Staršia karta "CNV": CNVs sa tiež zobrazujú v staršom prehliadači, aby sa zachovala spätná kompatibilita s existujúcimi pracovnými postupmi.
- SVs sú prístupné iba na karte "CNVs & SVs" alebo "SVs".

Dostupnosť údajov v prehliadači veľkých variantov závisí od typu analýzy:

- WGS: Plne kompatibilné (CNVs, SVs, UPDs).
- WES: Plne kompatibilné (CNVs, SVs, UPDs). CNVs sa zobrazia iba v prípade, že sa modul na volanie CNV spustí ako súčasť analýzy.
- Panely: Čiastočne kompatibilné (iba SVs).
- Terciárna analýza (VCF/aCGH):
- aCGH: Zobrazujú sa všetky udalosti počtu kópií.
- VCF: Sú zobrazené CNVs a SVs (symbolické ALT alely). SVs so špecifickými sekvenciami ALT sú zobrazené v prehliadači malých variantov.

Tabuľka výsledkov CNVs & SVs

Podobne ako v tabuľke malých variantov, hlavná tabuľka zobrazuje jeden riadok pre každú udalosť, s nasledujúcimi stĺpcami poskytujúcimi metriky špecifické pre variant:

DiagAI. Priorizačné skóre od 1 do 100, prispôbené pre veľké varianty, ktoré radí varianty podľa pravdepodobnosti, že spôsobujú ochorenie a sú diagnosticky relevantné v kontexte daného prípadu (najmä ak sú poskytnuté termíny HPO). Skóre DiagAI kombinuje viacero komponentov, vrátane: predikcie patogenity prostredníctvom skóre UP², zhody fenotypu prostredníctvom PhenoGenius (riadené HPO) a expertných pravidiel/pravidiel dedičnosti a kvality.

ACMG. Indikátor klasifikácie ACMG založený na veľkých variantoch sa zobrazuje ako klasifikačné skóre 1–5 v tabuľke LVV. V zásuvke s podrobnosťami karta ACMG zobrazuje základné kritériá a živý súhrn klasifikácie od Patogénne po Benígne (hlavne pre CNV Zisk/Strata).

Klasifikácia (ikona vlajky), ktorá umožňuje používateľovi klasifikovať variant vo VKB. Typ CNV/SV.

Typ CNV/SV (zisk alebo strata), veľkosť, umiestnenie a cytopruh.

Oblasti (exónové, intrónové, UTR) a gény prekrývajúce sa s CNV (názvy génov HGNC), prítomnosť v OMIM, podiel kódujúcich génov a podrobnosti o prekrývajúcich sa exónoch. Varianty sú anotované prekrývajúcimi sa génmi a génmi nachádzajúcimi sa do 5 kb od zlomových bodov (5 kb presah).

Gény spojené s variantom sú zoradené podľa klinickej relevancie (napr. morbidný stav v OMIM, percento prekrytia, skóre obmedzenia). Tri najrelevantnejšie gény sa zobrazujú v hlavnej tabuľke. Úplný zoznam je k dispozícii na karte "Genes & Regions" (Gény a regióny) v bočnom paneli ("Details Drawer" (Zásuvka s podrobnosťami)).

RMSK %. Percento udalosti prekrývajúcej sa s anotáciami RepeatMasker (napr. repetitívne oblasti).

ClinVar & Decipher (Klinické databázy). Oba stĺpce zobrazujú počet zhôd z príslušnej klinickej databázy. Najlepšia zhoda je definovaná ako udalosť, kde je prekrytie databázy > 50 %. Klasifikácie ClinVar a Decipher sú mapované na Benígne (B), VUS alebo Patogénne (P).

DGV Gold & GnomAD (Populačné databázy). Tabuľka zobrazuje maximálnu populačnú frekvenciu nájdenú pre variant. Udalosti sa považujú za zhodné, ak je prekrytie variantu > 80 % a typ udalosti je ekvivalentný. Karta Frekvencie v bočnom paneli uvádza všetky prekrývajúce sa udalosti s prekrytím > 1 %.

Phenotypes (Fenotypy). Kliknutím na čísla zobrazené v stĺpcoch zodpovedajúcich výsledkom algoritmu Phenogenius môže používateľ získať prístup k:

- percentu zhody, čo je najvyššie skóre zhody medzi jedným postihnutým génom a termínmi HPO pacienta.
- zlomku zhody, čo je počet termínov HPO pacienta, ktoré sa zhodujú s aspoň jedným postihnutým génom.
- Pri rodinnej analýze sa párovanie vykonáva pomocou termínov HPO probanda.

SampleFreq (Interná frekvencia vzorky). Frekvencia variantu v rámci aktuálneho projektu a subjektu je zobrazená v príslušných stĺpcoch.

- Zhody sú založené na > 80 % prekrytí variantu a rovnakom type udalosti.
- Frekvencia aktualizácie: údaje sú založené na snímke analýz vykonaných pred spustením aktuálnej analýzy. V Details Drawer je k dispozícii manuálne tlačidlo "Refresh" (Obnoviť) na aktualizáciu týchto údajov.

Stĺpce skóre. Používateľ si môže zobraziť viacero skóre:

- UP²: skóre predikcie patogenity, ktoré je súčasťou DiagAI.
- pLI, LOEUF: intolerancia k strate funkcie.
- pTriplo, pHAPLO: Triplosenzitivita a Haploinsuficiencia.
- HI ClinGen, TS ClinGen.

Stĺpce vzorky. Tieto stĺpce zobrazujú metriky udalostí špecifické pre zdroj údajov:

- Pre rodinné analýzy (s výnimkou aCGH) sa metriky vzorky zobrazujú pre všetkých členov rodiny.
- Genotype: pre CNVs, kde vyvolávač (caller) neposkytuje genotyp, systém odvodí genotyp na základe absolútneho počtu kópií vzhľadom na ploidiu:
 - Homozygot (hmz) s plným tmavým bodom
 - Heterozygot (htz) s polovičným tmavým bodom
- VAF / #Reads: frekvencia alely variantu a podporné čítania (k dispozícii pre SVs).
- #Copy: absolútny počet kópií (k dispozícii pre CNVs).
- LogRatio / Probes: k dispozícii pre údaje aCGH.
- QUAL: skóre kvality z vyvolávača variantov alebo vstupného súboru.

Triedenie, filtrovanie, prezeranie a interpretácia

Filtrovanie. Panel filtrov umožňuje používateľovi filtrovať varianty na základe údajov v stĺpcoch, podobne ako systém filtrovania na karte Malé varianty (pozri príslušnú časť).

- Génový panel: filtruje udalosti prekrývajúce sa s génmi v špecifickom zozname génov.
- Kvalita: filtruje na základe stĺpca QUAL. V predvolenom nastavení sú udalosti bez skóre kvality (N/A) skryté, pokiaľ nie je vybratá možnosť "Keep unknown" (Ponechať neznáme).
- Predvoľby: konfigurácie filtrov je možné uložiť ako predvoľby na budúce použitie.
- Pri rodinných analýzach používa filter kvality logiku "OR" (ALEBO). Variant sa zobrazí, ak aspoň jeden člen rodiny spĺňa prahovú hodnotu kvality.

Triedenie. V predvolenom nastavení je tabuľka zoradená podľa DiagAI Score (zostupne). Používateľ môže triediť podľa UP2, Size, Location alebo Sample Freq. Entity, vzostupne alebo zostupne.

Rovnako ako v staršej karte CNV môže používateľ otvoriť okná ploidie a prehliadač CNV v hornej časti tabuľky (pozri časť CNV: Zobrazenie a interpretácia). Ikona správy stĺpcov je tiež k dispozícii úplne vpravo v hlavičke.

Bočný panel CNV/SV (Details Drawer)

Podobne ako na stránke variantov (časť Malé varianty), výberom variantu v Large Variant Viewer sa na pravej strane obrazovky otvorí podrobná stránka. Táto stránka obsahuje niekoľko kariet:

Varianty. Táto súhrnná karta obsahuje nomenklatúru variantu, skóre DiagAI, zoznam podozrivých zložených heterozygotov (Suspected Het Compound), informácie o génoch a regiónoch a odkazy na databázy.

ACMG. Podrobná klasifikácia.

Gény a regióny. úplný zoznam génov a regiónov spojených s variantom, s podrobnými klinickými informáciami, ako je morbidný stav OMIM, percento prekrytia, skóre obmedzenia a zhoda fenotypu.

Klasifikácia. Klasifikujte aktuálny variant vo vašej VKB v hornej časti bočného panelu a uložte ho. Aktuálnu klasifikáciu a výskyty tohto variantu v iných analýzach nájdete na tejto stránke.

Frekvencie. Používateľ má prístup k podrobnostiam o frekvenciách variantov v databázach DGV Gold a GnomAD, ako aj pre subjekt a aktuálny projekt.

Klinické databázy. ClinVar & Decipher.

Viewer. Používateľ si môže vybrať, či chce variant zobrazíť v SeqOne viewer alebo IGV. Na podkarte IGV môže používateľ zobrazíť "Start" (Začiatok) alebo "End" (Koniec) udalosti, alebo "Start to End" (Od začiatku do konca). Na podkarte SeqOne môže používateľ pristupovať k Genome Browser a vizualizovať udalosť v jednej z viacerých dostupných vrstiev: CNV, SV, UPD, Gene, CN (copy number), BAF, DGV.

13.1.7. Signatúry v analýze SomaCGP

Metodika výpočtu skóre MSI

Analýza MSI je individuálna analýza vykonávaná na základe štandardizovanej referencie. Táto referencia pozostáva zo súboru vopred vybraných a zozbieraných negatívnych vzoriek. Vzorka nádoru sa porovnáva s touto referenciou pomocou nástroja *MSIsensor*. Následne sa vypočíta skóre, ktoré predstavuje percentuálny podiel nestabilných mikrosatelitných oblastí pokrytých panelom.

Panel normálov, ktorý v súčasnosti podporuje SomaCGP, zodpovedá panelu Agilent SureSelect Cancer CGP. Použitie inej súpravy alebo vlastného dizajnu odvodeného od CGP si bude vyžadovať vytvorenie špecializovaného panelu normálov.

Metodika výpočtu skóre TMB

TMB sa vypočíta v troch krokoch:

1. Výpočet veľkosti kódujúcich oblastí zahrnutých v paneli
2. Identifikácia somatických variantov
3. Výpočet počtu somatických mutácií na megabázu (mut/Mb)

Identifikácia somatických variantov je založená na súbore filtrov opísaných nižšie:

- Filtry typu a účinku variantu
 - Iba SNV (MNV sú vylúčené)
 - Zachované účinky: zisk stop kodónu, posun čítacieho rámca, variantu splice (akceptor a donor), ablácia transkriptu, fázová delécia/inzercia, missense variantu, synonymná variantu.

Synonymné varianty sú vylúčené z výpočtu nesynonymného skóre TMB.

- Filtry kvality variantov
 - Varianty detegované Mutect2, ktoré prešli filtermi FilterMutectCalls
 - Hĺbka pokrytia (DP) $\geq 50X$
 - Skóre kvality > 30
 - Pomer alelickej podpory (ASR) $> 0,1$
- Filtry somatických variantov
 - Frekvencia alel variantov (VAF) medzi 5 % a 40
 - Frekvencia v bežnej populácii (GnomAD) $\leq 0,0001$

Riziká používania syntetických kontrolných vzoriek TMB

- Neprirodzené rozloženie mutácií:
 - Upravené mutácie nemusia odrážať genomický kontext skutočných nádorov (napr. zhlukovanie mutácií, skreslenie reťazca, kontext sekvencie).
- Chýbajúca heterogenita nádorov:
 - Syntetické vzorky nenapodobňujú klonálnu diverzitu ani čistotu nádoru, čo ovplyvňuje distribúciu frekvencie variantných aliel (VAF).
 - Žiadny šum v pozadí ani artefakty FFPE: skutočné vzorky obsahujú sekvenčné artefakty, poškodenia FFPE a variácie počtu kópií, ktoré v syntetických materiáloch chýbajú.
- Odlišná veľkosť fragmentov a chemické správanie:
 - Syntetické konštrukty sa často správajú odlišne počas prípravy knižnice, hybridného zachytávania a sekvenovania.

Vizualizácia a interpretácia

Karta „Signatures“ (Signatúry) zobrazuje výsledky analýz MSI a TMB v 2 častiach:

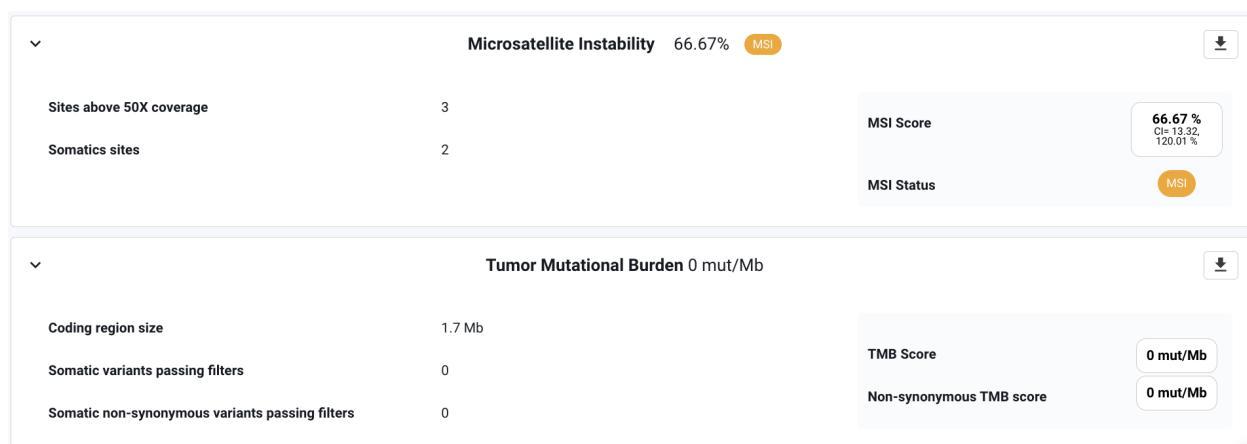
Mikrosatelitová nestabilita. Tento pohľad zobrazuje informácie, ako napríklad:

- počet mikrosatelitných oblastí s pokrytie väčším alebo rovným 50X (miesta s pokrytie nad 50X),
- počet somatických miest,
- skóre MSI v percentách a 95 % interval spoľahlivosti (CI),
- Stav MSI

Mutácia nádoru. Tento pohľad zobrazuje nasledujúce informácie:

- veľkosť kódujúcej oblasti vyjadrená v Mb (veľkosť kódujúcej oblasti),
- počet somatických a nesynonymných variantov, ktoré prešli filrami (Somatické varianty, ktoré prešli filrami a Somatické nesynonymné varianty, ktoré prešli filrami),
- skóre TMB všetkých somatických variantov, ktoré prešli filrami v mut/Mb (skóre TMB),
- skóre TMB nesynonymných variantov, ktoré prešli filrami mut/Mb.

Každý výsledok je možné stiahnuť kliknutím na ikonu „“ (Stiahnuť výsledky) v pravom hornom rohu každej sekcie ako súbor tsv.



Obrázok 29: Zobrazenie karty „Signatures“ (Signatúry) analýzy SomaCGP.

13.1.8. Krátke tandemové repetície (STR)

Vizualizácia krátkych tandemových opakovaní (STR) je k dispozícii pre pracovné súbory analýzy GermVar Tertiary a GermVar Tertiary Family určené na analýzu údajov sekvenovania celého genómu (WGS), ak sú generované sekvenovaním Oxford Nanopore Technologies.

Detekcia krátkych tandemových opakovaní (STR) je tiež k dispozícii v analytických pracovných sadách GermVar a GermVar Family pre WGS z údajov FASTQ generovaných pomocou Illumina, Element Biosciences, MGI a Oxford Nanopore Technologies.

Metodika

Krátke tandemové opakovania z FASTQ

Údaje o krátkych tandemových opakovaniach (STR) sa detegujú pomocou nástroja ExpansionHunter, ktorý spracováva súbor BAM spolu s referenčným genómom a katalógom variantov definujúcim cieľové lokusy. Výsledný VCF je potom anotovaný nástrojom Stranger, ktorý klasifikuje každý lokus ako Normálny, Pre-mutácia alebo Patogénny na základe pozorovaného počtu opakovaní a známych prahových hodnôt ochorenia.

Oba nástroje sú integrované priamo do potrubí GermVar a GermVar Family.

ExpansionHunter: <https://github.com/Illumina/ExpansionHunter>

Stranger: <https://github.com/moonso/stranger>

Krátke tandemové opakovania z VCF

Údaje o tandemových opakovaniach sú extrahované zo súboru VCF (sample.wf_str.vcf.gz) vygenerovaného z pracovného postupu wf-human-variation v2.3.0 aplikácie EPI2ME (Oxford Nanopore Technologies) a poskytnutého používateľom. Keďže tieto údaje boli anotované pre každý lokus nesúci STR z katalógu variácií udržiavaného EPI2ME, náš proces tieto anotácie nemení. Premutovaný alebo patogénny stav každého alela sa určuje podľa počtu sekvenovaných opakovaní a známych prahov premutácie a patogenity.

Zobrazenie a interpretácia

Karta STR vám umožňuje zobrazit' krátke tandemové opakovania vo forme tabuľky, ktorá zobrazuje nasledujúce podrobnosti pre každý genotypizovaný lokus:

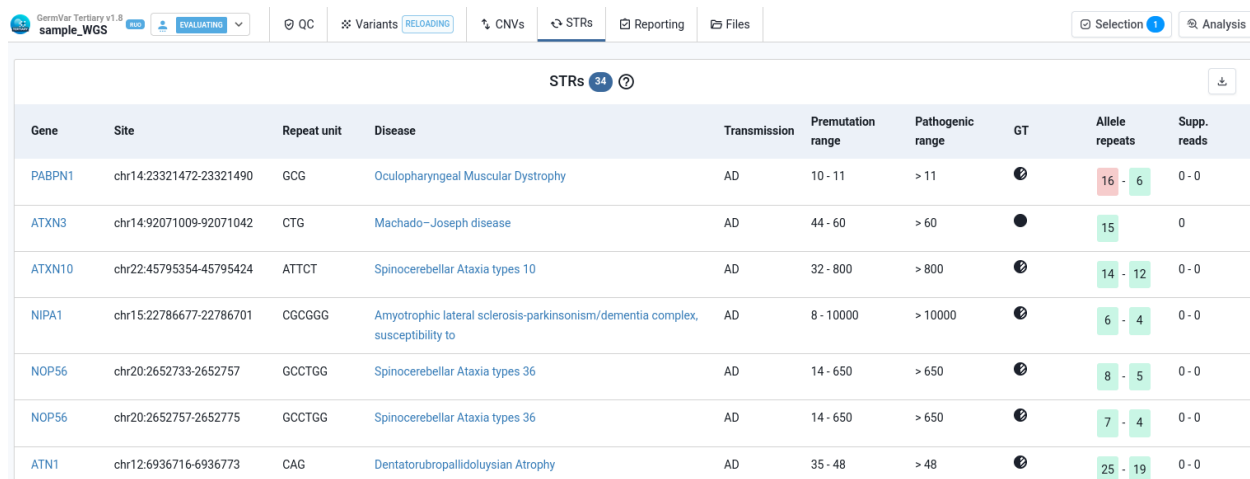
Informácie o lokuse a poznámky

- **Gén:** názov génu (HGNC) a príslušný odkaz.
- **Miesto:** genomické súradnice opakovania.

- **Opakovaná jednotka:** opakovaný nukleotidový motív.
- **Choroba:** súvisiaca choroba podľa databázy OMIM a príslušný odkaz.
- **Prenos:** známy spôsob prenosu.
- **Rozsah premutácie:** rozsah premutácie zodpovedajúci minimálnemu a maximálnemu počtu opakovaní, ktoré určujú premutovaný stav.
- **Patogénny rozsah:** dolná hranica opakovaní určujúca patogénny stav.

Podrobnosti špecifické pre každú vzorku

- **GT:** homozygotný, bialelický alebo heterozygotný genotyp lokusu.
- **Opakovania alely:** počet opakovaní na alelu. Homozygotné lokusy majú jednu hodnotu. Pre bialelické a heterozygotné lokusy je uvedený počet opakovaní pre obe alely. Stav každej alely je určený počtom opakovaní: divoký typ je zelený, premutovaný je oranžový a patogénny je červený.
- **Supp. reads:** počet podporných čítaní pre každú alelu. Homozygotné lokusy nesú jednu hodnotu, zatiaľ čo hodnoty pre 2 alely sú uvedené pre bialelické a heterozygotné lokusy.



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

Obrázok 30: Zobrazenie karty STR v terciárnej analýze GermVar.

Tabuľku je možné stiahnuť ako súbor repeats.tsv prostredníctvom ikony „“ (Stiahnuť súbor) v pravom hornom rohu tabuľky. Tento súbor je k dispozícii aj na karte „Files“ (Súbory) analýzy.

13.1.9. Fúzne gény

13.1.9.1. Fúzie zistené z DNA

Metodika

Modul [GRIDSS](#) (Genome Rearrangement IDentification Software Suite), špecializovaný na detekciu genomických reorganizácií a implementovaný v pracovnej sade SomaVar a SomaVar LF, sa používa na detekciu génových fúzií z sekvenovania DNA, obohatených o zachytenie zamerané na panel génov.

Zobrazenie a interpretácia

Karta Fúzie somatickej analýzy (SomaVar, SomaVar LF, SomaCGP, SomaVar Tertiary) zobrazuje zoznam fúzií, ktoré prešli rôznymi filtermi GRIDSS, vo forme tabuľky s podrobnými informáciami o fúzii a každom zúčastnenom partnerovi:

Informácie o fúzii

- fúzia s génovými symbolmi každého partnera,
- typ udalosti (translokácia, duplikácia, inverzia alebo delécia),
- výskyt fúzie v databáze COSMIC,
- frekvencia alely (VAF),
- počet čítaní podporujúcich fúziu,
- celkové pokrytie,
- filtrov priradených GRIDSS.

Informácie týkajúce sa partnera

- genómové súradnice bodu fúzie (breakend),
- zainteresovaný RefSeq transkript,
- miesto zodpovedajúce bodu fúzie (bod zlomu): 5'UTR, 3'UTR, intron, exon (nekódujúci bod zlomu), CDS (kódujúci bod zlomu) alebo miesto splice,
- číslovanie exónu alebo intrónu zapojeného do fúzie, vzhľadom na celkový počet exónov alebo intrónov.

Ponuka filtrov vám umožňuje

- hľadať konkrétnu fúziu alebo gén zapojený do fúzie prostredníctvom vyhľadávacieho poľa,
- zobrazíť fúzie, ktoré neprešli filtermi GRIDSS (LOW_QUAL), zaškrtnutím políčka „Zobraziť filtrované fúzie“. Toto políčko je štandardne 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

Obrázok 31: Vizualizácia karty Fúzie pre analýzu SomaVar alebo SomaVar LF.

Údaje v tabuľke je možné triediť podľa genomických súradníc (breakend), prítomnosti fúzie v COSMIC (Databázy), alelickej frekvencie (VAF) a počtu čítaní podporujúcich fúziu (Podporné čítania).

Odkaz spojený s genomickými súradnicami každého bodu zlomu otvorí rozdelený pohľad na bod fúzie v IGV. Nástroje podrobne uvádzajú počet rozdelených čítaní a nezhodných párov v stĺpci Podporné čítania a pokrytie v každom bode zlomu (Cov. 1 a Cov. 2) v stĺpci Pokrytie.

Stĺpec Filter zobrazuje 3 filtre kvality používané GRIDSS, ktoré poskytujú informácie o spoľahlivosti identifikácie fúzie:

- PASS: fúzia prešla filtermi kvality. Skóre kvality je väčšie alebo rovné 500 a považuje sa za stredne alebo vysoko spoľahlivé volanie (ak je > 1000).
- LOW_QUAL: fúzia má skóre kvality nižšie ako 500 pre aspoň jeden z dvoch koncov zlomu zapojených do fúzie a je považovaná za volanie s nízkou spoľahlivosťou.
- ASSEMBLY_ONLY: fúzia je volaná na základe zostavenia a nie na základe čítania, ktoré ju podporuje. Zostavenie sa týka rekonštrukcie sekvencií okolo potenciálnych fúzných bodov, ktoré sú potom znovu zarovnané, aby sa identifikovala fúzia.

Ak je pre tú istú fúziu k dispozícii viacero filtrov, znak „+“ zobrazí všetky filtre v stĺpci Filter.

Okrem výsledkov zobrazených v rozhraní je k dispozícii súbor fusions.tsv s podrobnými informáciami o zistených fúziách. K dispozícii je na stiahnutie na karte Súborné (Files) alebo cez

ikonu  v pravom hornom rohu stránky. Obsahuje nasledujúce informácie:

- názov fúzie s génovými symbolmi pre každého partnera (fúzia)
- genomové súradnice každého fúzného bodu (breakend_1 a breakend_2)
- kanonický transkript pre každého partnera (transcript_canonical_1 a 2)
- miesto zodpovedajúce fúznej bodu pre každého partnera (site_1 a site_2)
- číslovanie exónov alebo intrónov podľa príslušného transkriptu RefSeq (exon_1 a 2)
- 60 bp flankujúcich sekvencií obklopujúcich fúzny bod (flanking_region_1 a 2)

- filtry kvality priradené GRIDSS (filter)
- typ udalosti (type)
- počet rozdelených čítaní (split_reads)
- počet nezhodných párov čítaní (discordant_pairs)
- pokrytie pre každého partnera (coverage_1 a coverage_2)
- podiel čítaní podporujúcich fúziu (read_fraction)
- počet čítaní podporujúcich fúziu (supporting_reads)
- celkové pokrytie (coverage)
- Identifikátor fúzie COSMIC (cosmic_id)

Tento súbor obsahuje:

- Fúzie zobrazené v rozhraní, ktoré prešli filtermi kvality (filter = PASS),
- Fúzie, ktoré boli detegované, ale s hodnotením kvality nižším ako 500 pre aspoň jeden z dvoch koncov zlomu zapojených do fúzie (filter = LOW_QUAL).

Hodnotenie kvality pre každý koniec zlomu nájdete v súbore gridss.raw.vcf.gz (stĺpec QUAL), ktorý si môžete stiahnuť na karte Súbory.

13.1.9.2. Fúzie detegované z RNA

Nástroj SomaRNA na analýzu údajov RNA-seq buď z cielených prístupov, ako je zachytávanie a amplikón, čo znamená, že používatelia môžu identifikovať a zobraziť udalosti génovej fúzie na úrovni transkriptu.

Metodika

Na detekciu a zobrazenie fúzií pomocou SomaRNA sa používa kombinácia dvoch nástrojov:

- [STAR](#): nástroj na zarovnávanie údajov RNA-seq určený na zarovnávanie nesúvislých sekvencií priamo s referenčným genómom
- [Arriba](#): detekuje génové fúzie z chimérických zarovnaní

Na základe chimérických zarovnaní z programu STAR aplikuje program Arriba súbor filtrov na elimináciu známych artefaktov a transkriptov pozorovaných v nepatologických kontextoch. Každý potenciálnej fúznej udalosti priraduje skóre spoľahlivosti, ktoré závisí od viacerých kritérií:

- Počet sekvencií podporujúcich fúziu
- Rovnováha medzi rozdelenými čítaniami a nezhodnými pármami
- Vzdialenosť medzi bodmi zlomu, ich blízkosť (intragénová alebo iná)
- Typ udalosti

Fúzie detegované programom Arriba so strednou alebo vysokou spoľahlivosťou sa zobrazujú v SeqOne. Tie s nízkou spoľahlivosťou sú uvedené v súboroch „arriba_discarded“ (pozri nižšie). Spoľahlivosť programu Arriba odzrkadľuje pravdepodobnosť abnormálneho transkriptu, ktorý sa nepozoruje v zdravom tkanive, možnosť základného genomického preskupenia a vylúčenie artefaktov.

Viac podrobností o úrovni spoľahlivosti nájdete v dokumentácii Arriba: <https://github.com/suhrig/arriba/blob/master/documentation/08-Interpretation-of-results.md>

Aby ste pochopili, prečo je fúzia „nízkej spoľahlivosti“, pozrite si stĺpec „filters“ (filtry) (AA) v súbore „discarded“ (vyradené). Uvádza filter, ktorý udalosť vyradil. Podrobnosti o jednotlivých filtroch: <https://github.com/suhrig/arriba/blob/master/documentation/11-Internal-algorithm.md>

Zobrazenie a interpretácia

Na karte Fúzie analýzy SomaRNA získa používateľ zoznam predpovedí fúzií vo forme rôznych kariet.

V hornej časti stránky sú uvedené podrobnosti o fúzii a každom z partnerov, ktorí sa na nej podieľajú:

Všeobecné informácie o fúzii

- počet prekrývajúcich sa spojovacích čítaní, rovný súčtu príslušných rozdelených čítaní každého z génov 5' a 3' zapojených do fúzie,
- počet nezhodných párov,
- skóre spoľahlivosti pridelené Arribou (vysoké, stredné alebo nízke),
- typ udalosti vedúcej k fúzii (translokácia, duplikácia, inverzia alebo delécia),
- vplyv fúzie na čítací rámec,
- prípadne výskyt fúzie v databáze COSMIC.

Informácie o každom fúzovanom partnerovi

- symbol každého génu,
- transkript Refseq alebo, ak nie je k dispozícii, ENSEMBL pre každý gén,
- číslovanie exónov zapojených do fúzie,
- genómové súradnice každého bodu zlomu,
- celkové pokrytie v bode fúzie (Coverage),
- čítaná frakcia podporujúca fúziu.

Rozdelené čítania sú čítania zarovnané prevažne v géne 1 alebo géne 2 a prekrývajúce fúziu. V prípade nezhodných párov čítaní je jedno z čítaní v géne 1 a druhé v géne 2, bez nutného prekrývania fúzie.

Časť čítania je súčet rozdelených čítaní a nezhodných párov čítaní spojených s génom, vydelený celkovým pokrytím. Hodnota časti čítania môže byť väčšia ako 100 %, ak sú rozdelené čítania spojené s génom súčasťou aj nezhodných párov čítaní. Tieto čítania sa potom počítajú dvakrát namiesto jedenkrát.

DHRX - 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

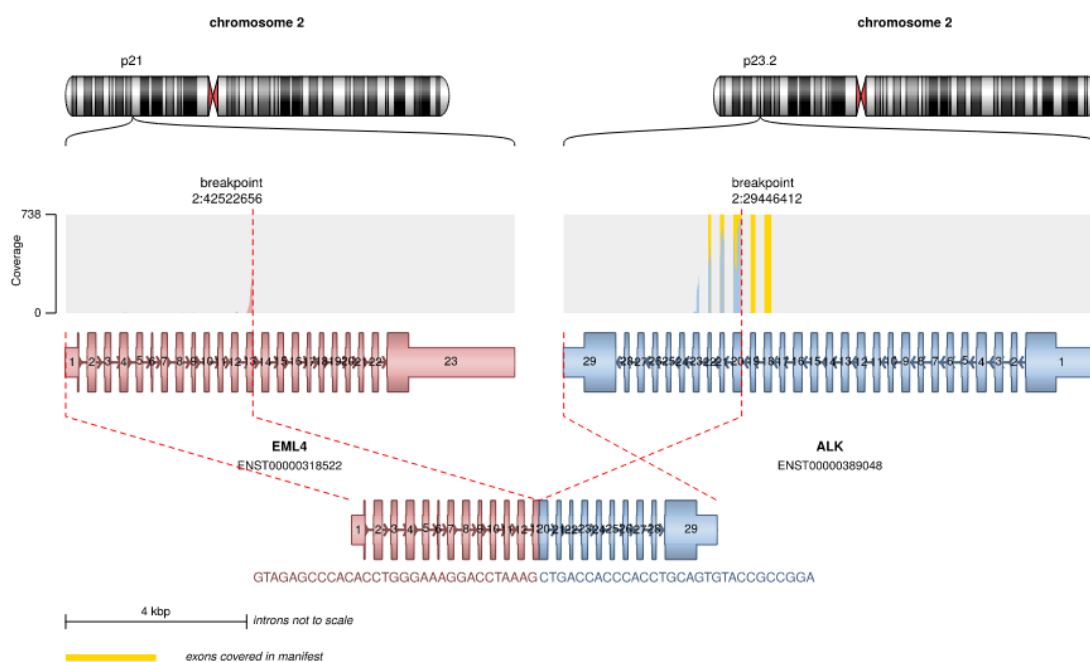
3.9%

6.08%

Obrázok 32: Vizualizácia informácií o fúzii z analýzy SomaRNA.

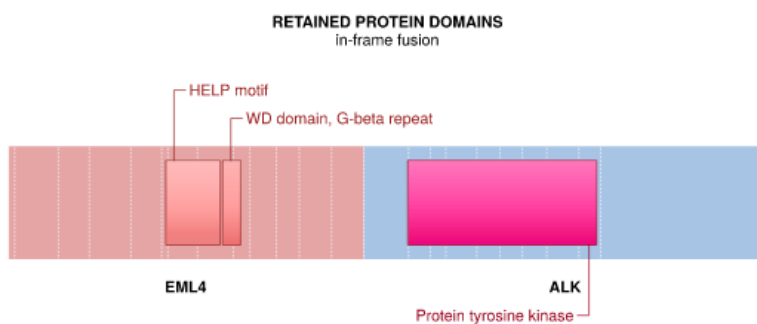
V spodnej časti stránky sú zobrazené grafické znázornenia fúzie, a to na úrovni génu a exónu, ako aj na úrovni proteínu. Táto posledná časť je rozdelená na dve karty:

- Prvá karta predstavuje štruktúru fúzie a zúčastnených partnerov, kde pokrytie exónov zahrnutých v analýze je znázornené ako diagram na žltom pozadí:



Obrázok 33: Pohľad na fúziu medzi dvoma génmi

- Proteínové domény zachované po fúzii sú zvýraznené v druhom zobrazení:



Obrázok 34: Pohľad na polypeptid vzniknutý fúziou dvoch génov

Okrem výsledkov prezentovaných v rozhraní sú k dispozícii na stiahnutie rôzne súbory:

- Fúzie zachované programom Arriba vo formáte TSV (súbor arriba.tsv),
- Fúzie eliminované filtrami Arriba vo formáte TSV (súbor arriba_discarded_standard.tsv),
- fúzie eliminované filtrami Arriba a uvedené v COSMIC vo formáte TSV (súbor arriba_discarded_cosmic.tsv),
- rôzne obrázky zobrazené v rozhraní vo formáte SVG

13.1.10. Klonálnosť

Správa o mutačnom stave IGHV v záložke „Klonálnosť“ analýzy SomaHemato zobrazuje výsledky mutačného stavu.

Sekcia „Výsledky IGHV“ zobrazuje:

- Mutovaný stav IGHV a percento homológie:
 - Mutovaný: ak je homológia <98 % identita s germline
 - Nemutované: ak je homológia ≥98 % identita s germline
 - Hraničný: pre klony v rozmedzí 97–98 % identity
 - Polyclonálny: žiadny významný klon nad 20
 - Neprikladný: ak zlyhá akákoľvek podstatná kontrola kvality
- Určená podmnožina (s klinickým vplyvom) a spoľahlivosť: *2 (vysoké riziko), *8 alebo žiadna
- Varovné upozornenie, ak sa v správe objavujú varovania.

V časti „Klonálnosť“ sú uvedené identifikované klony s ich frekvenciou a produktivitou (produktívne, neproduktívne).

Každý klon má vyhradenú sekciu, v ktorej je uvedený jeho mutačný stav, frekvencia, odkaz na stránku IMGT_quest a ďalšie podrobnosti (homológia, úplnosť) spolu s kontrolami kvality.

13.1.11. Varianty splicingu

Metodika

Sada nástrojov SomaRNA používa sadu nástrojov *Regtools* na detekciu skokov exónov a identifikáciu alternatívnych variantov splice zo sekvenovania RNA, obohatených o zachytenie zamerané na panel génov. RegTools extrahuje spojenia exón-exón zo súboru BAM generovaného nástrojom STAR RNA-seq data aligner a anotuje ich podľa známej štruktúry transkriptu na základe súboru GTF a anotácií Ensembl.

Vizualizácia a interpretácia

Karta Varianty splice analýzy SomaRNA sa zameriava na vizualizáciu variantov splice nasledujúcich klinicky relevantných génov a ich hlavných akčných variantov:

- MET (MET protoonkogén, receptor tyrozínkinázy) a akčná varianta MET Δ 14
- EGFR (receptor epidermálneho rastového faktoru) a akčná varianta spojenia EGFRvIII
- AR (androgénny receptor) a akčná varianta splice AR-V7

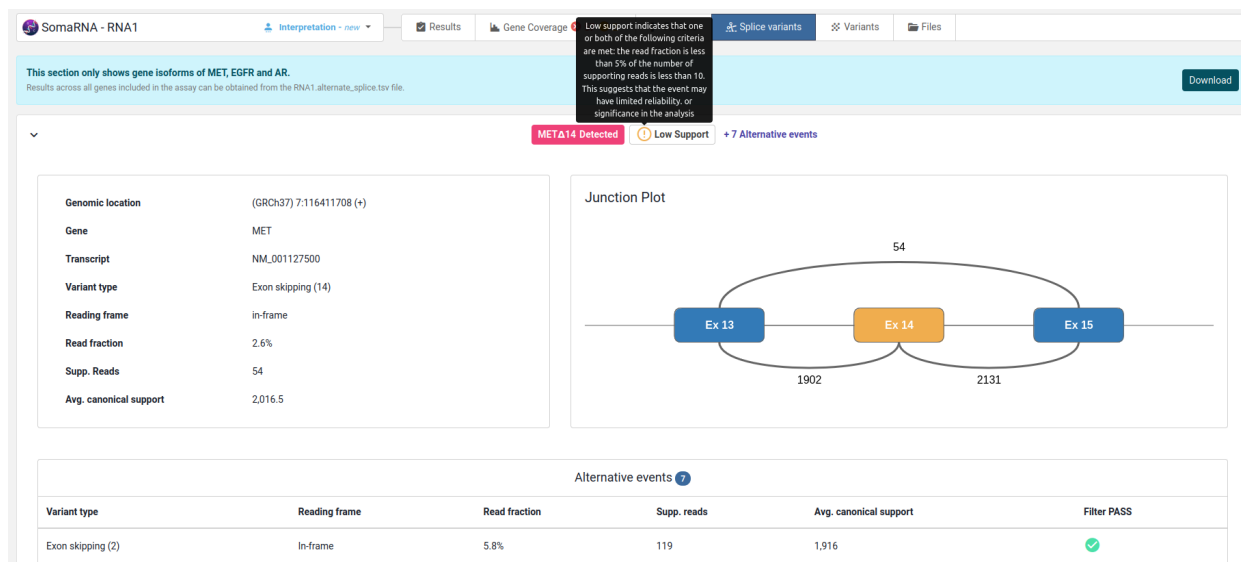
Kompletný zoznam splice variantov detegovaných v analýze SomaRNA, okrem tých, ktoré sú uvedené vyššie, je možné stiahnuť priamo ako súbor TSV prostredníctvom ikony Stiahnuť v rozhraní alebo z karty Súborný v analýze.

Pre každý z vyššie uvedených génov sa zobrazuje stav detekcie (detegovaný alebo nedetekovaný) hlavnej akčnej varianty spolu s počtom ďalších alternatívnych udalostí detegovaných pre daný gén.

Používatelia môžu zobrazit' podrobnosti o hlavných zistených variantoch splicingu otvorením špeciálnej vložky, ktorá je rozdelená do troch častí:

- Vľavo sa zobrazujú informácie o hlavnej použiteľnej variante: genomické súradnice, názov gén, transkript RefSeq, typ zistenej varianty s podrobnosťami o chýbajúcich exónoch, fáza čítania, frakcia čítaní a počet čítaní podporujúcich variantu. Nakoniec sa špecifikujú priemerné kanonické podporné čítania. To zodpovedá priemeru čítaní 5' a 3', ktoré podporujú zaradenie exónu.
- Vpravo je graf sashimi (Junction Plot), ktorý znázorňuje zistené spojenia medzi exónmi, ako aj počet čítaní podporujúcich každé spojenie.
- V spodnej časti sú uvedené informácie o zistených alternatívnych udalostiach, t. j. izoformách iných ako hlavná varianta, ktorá je predmetom záujmu opísaná vyššie: typ zistenej varianty (chýbajúce čísla exónov v zátvorkách), fáza čítania, frakcia a počet

čítaní podporujúcich variantu, priemerný počet čítaní podporujúcich kanonický transkript (Avg. canonical support), stav PASS alebo FAILED podľa filtrov opísaných nižšie.



Obrázok 35: Vizualizácia karty Splice variants analýzy SomaRNA.

Detekcia variantov splicingu v génoch zobrazených na karte nepodlieha aplikácii detekčných filtrov.

Ikona „Low support“ (Nízka podpora) s detekciou MET Δ 14 a EGFRvIII však používateľovi signalizuje, že metriky súvisiace s podporujúcimi čítaniami sú pod nasledujúcimi prahovými hodnotami: podiel čítaní < 5 % alebo počet podporujúcich čítaní < 10.

Stav PASS alebo FAILED iných alternatívnych udalostí tiež používa tieto prahové hodnoty:

- **PASS** : podiel čítaní $\geq$ 5 % a počet čítaní $\geq$ 10.
- **NEÚSPECH** : podiel čítaní < 5 % alebo počet čítaní < 10.

Súbor TSV, ktorý je možné stiahnuť, obsahuje podrobné informácie o všetkých zistených udalostiach spojovania: transkript RefSeq, gén, smer reťazca (+ alebo -), podiel čítaní a počet čítaní podporujúcich variantu, počet chýbajúcich exónov a príslušné čísla exónov, fáza čítania, genomické súradnice spojení. Súbor tiež špecifikuje pod FILTER_PASS (áno/nie), či varianta prekračuje prahové hodnoty uplatňované na podporujúce čítania (podiel čítaní $\geq$ 5 % a počet čítaní $\geq$ 10) a pod Actionable (áno/nie) akčnosť varianty.

13.1.12. Stav MSI/MSS

Analytická sada SomaMSI je k dispozícii pri spustení novej analýzy údajov génového panelu a slúži na určenie stavu MSI alebo MSS vzorky nádoru.

Metodika

Táto neindividuálna kohortová analýza je podobná analýze CNVCapture alebo SomaCNVCapture. Na spustenie analýzy je potrebných najmenej osem vzoriek.

Nástroj generuje referenčnú vzorku zo vzoriek v kohorte, ktoré sú odhadované ako negatívne. Keďže tento stav vzoriek nie je vopred známy, tento odhad je založený na výpočte miery delécií pozorovaných v mikrosatelitných oblastiach: polovica vzoriek v kohorte s najnižšou mierou je preto vybraná na generovanie referenčnej vzorky, pričom každá vzorka prispieva časťou svojich sekvencií.

Každá vzorka nádoru sa následne porovná s referenčnou vzorkou pomocou nástroja MSIsensor. Vypočíta sa skóre, ktoré predstavuje percento študovaných mikrosatelitných oblastí, ktoré vykazujú nestabilitu. Študované oblasti sú mikrosatelitné oblasti pokryté panelom a detegované nástrojom MSIsensor.

Na základe tohto skóre je stav MSI alebo MSS rôznych vzoriek určený prahovou hodnotou, ktorá sa vypočíta takto: $\max(10, \text{medián} + \max(3 \cdot \text{medián} - \text{abs-dev}, 0,3 \cdot \text{medián}))$

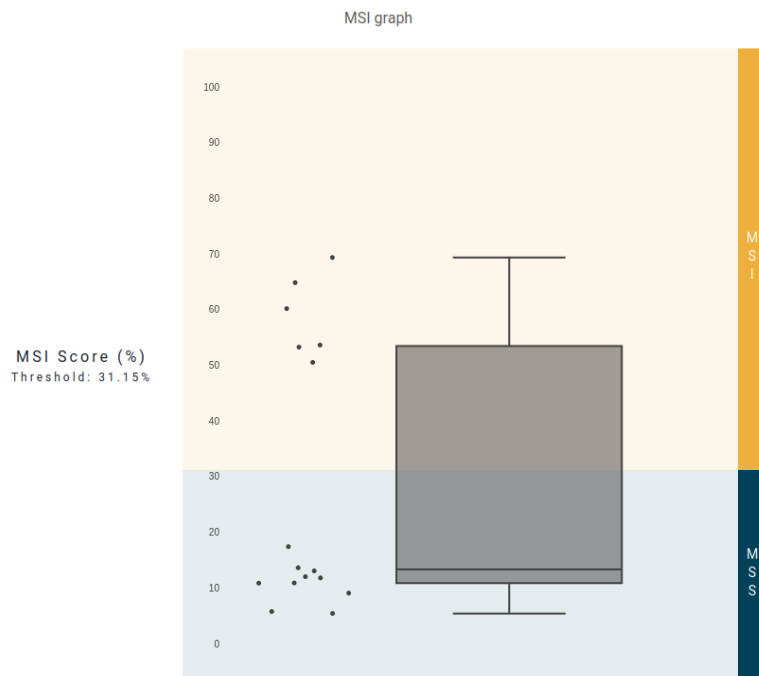
Princíp tohto skóre je nasledovný:

- V analýze prezentujúcej vzorky MSI a MSS bude prahová hodnota spravidla medián + $3 \cdot \text{medián} - \text{abs-dev}$.
- Ak však všetky vzorky majú stav MSS, ak sú skóre veľmi nízke a ak je ich rozloženie extrémne blízke, prahová hodnota sa nastaví na 10, aby sa zabránilo generovaniu falošných pozitív.

Zobrazenie a interpretácia

Po dokončení analýzy je možné zobrazíť výsledky všetkých vzoriek z analýzy SomaMSI. Karta Výsledky je rozdelená do dvoch častí:

Graf MSI: grafické znázornenie výsledkov, kde je každá vzorka zobrazená podľa svojho skóre. Prahová hodnota stanovená vzorkovou kohortou vytvára v grafe dve zóny, ktoré predstavujú vzorky považované za MSI a MSS.



Obrázok 36: Box plot znázorňujúci rozloženie skóre MSI pre vzorkovú kohortu.

Podrobnosti o jednotlivých vzorkách: rozpis skóre a stav priradený každej vzorke sú k dispozícii v pravom paneli.



Obrázok 37: Rozpis skóre a stavu všetkých vzoriek zahrnutých do analýzy SomaMSI.

13.1.13. Stav HRD/HRP

Metodika

Detekcia stavu HRD navrhnutá spoločnosťou SeqOne je založená na analýze údajov sekvenovania genómu s nízkym pokrytím (sWGS), ku ktorým je možné pridať výsledky analýzy génového panelu, ak tento obsahuje gény BRCA1 a BRCA2.

Princíp je nasledovný:

- Kvantifikácia genomovej nestability z údajov sWGS a identifikácia veľkých prestavieb (LGA) ako aj strát rodičovských kópií (LPC), ktoré by odrážali stratu heterozygotnosti.
- Identifikácia variácií počtu kópií (CNV) v kľúčových génoch, ktoré sú opísané ako potenciálne ovplyvňujúce stav HRD.
- Voliteľne detekovať patogénne varianty v génoch BRCA1, BRCA2 a TP53, ak sú výsledky panelu k dispozícii na rozhraní SeqOne pre príslušného pacienta. V tomto prípade sa berú do úvahy typy variantov detekovaných pracovnou sadou SomaVar, ktoré sú SNV, indely a SV (detekované Freebayes, Mutect2 a GRIDSS).

Stav HRD. Pozitívny stav HRD je definovaný aspoň jedným z nasledujúcich kritérií (**len verzie SomaHRD v1.2 a v1.3**): (1) aspoň jedna patogénna mutácia BRCA1/2 a/alebo (2) pravdepodobnosť HRD väčšia alebo rovná 0,5.

Od verzie v1.4 pracovnej sady SomaHRD je stav analýzy definovaný výlučne skóre genomovej nestability (skóre GI). Používatelia môžu vybrať stav HRD (pozitívny, negatívny alebo nepresvedčivý) vedľa stavu GI, aby ho zahrnuli do správy namiesto stavu GI.

Pravdepodobnosť HRD. (**Iba verzie SomaHRD v1.2 a v1.3**) Pravdepodobnosť SeqOne HRD zodpovedá pravdepodobnosti, že nádor je HRD pozitívny na základe charakteristík genomovej nestability LGA a LPC stanovených z údajov sWGS. Veľké genomové zmeny (LGA) sú definované ako počet kópií zlomových bodov, kde zlomový bod zodpovedá zmene počtu kópií medzi dvoma genomovými segmentmi s dĺžkou najmenej 10 Mb a vzdialenosťou najviac 3 Mb. Straty rodičovských kópií (LPC) sú definované ako počet haploidných segmentov s dĺžkou najmenej 10 Mb. Pravdepodobnosť HRD je lineárnou kombináciou dvoch charakteristík genomovej nestability, prevedených pomocou logistickej regresie, aby sa hodnoty umiestnili na pravdepodobnostnú stupnicu medzi 0 a 1.

GI probability (GiP) (Pravdepodobnosť GI) (z SomaHRD v1.4) Pravdepodobnosť genomovej nestability (GI) predstavuje pravdepodobnosť, že nádor vykazuje genomovú nestabilitu, na základe sekvenovania celého genómu (sWGS). Výpočet tejto pravdepodobnosti závisí od dvoch typov charakteristík genomovej nestability LGA a LPC. LGA je definovaná ako počet bodov zlomu počtu kópií. Bod zlomu zodpovedá zmene počtu kópií medzi dvoma dostatočne veľkými genomickými segmentmi v tesnej blízkosti. Skóre LGA nie je jediným meradlom, ale skôr súbor

niekoľkých variácií. Tieto variácie sa počítajú pomocou rôznych parametrov pre minimálnu veľkosť fragmentov a vzdialenosť medzi fragmentmi. Skóre LPC sa naopak vzťahuje na počet haploidných segmentov s dĺžkou najmenej 10 Mb. Pravdepodobnosť genomovej nestability SeqOne nie je priamou funkciou priemerných skóre LGA a LPC. Počíta sa pomocou posilňovacieho modelu, ktorý integruje niekoľko vlastností LGA a LPC, čím sa zlepšuje presnosť predikcie. Konečná pravdepodobnosť je priemerom 100 pravdepodobností (v rozmedzí od 0 do 1) vypočítaných pomocou rôznych parametrov modelu.

GI status. (z **SomaHRD v1.4**) Pozitívny status genomovej nestability (GI status) je definovaný pravdepodobnosťou genomovej nestability väčšej ako 50 %.

Stav BRCA. Na detekciu patogénnych variantov, určenie stavu BRCA nezávisle od pravdepodobnosti HRD alebo pravdepodobnosti GI, sa na varianty BRCA1 a BRCA2 z analýzy SomaVar pacienta uplatňujú nasledujúce kritériá:

- Frekvencia alely $\geq 15\%$ (**verzie SomaHRD v1.2 a v1.3**),
- Frekvencia alely $\geq 10\%$ (**z SomaHRD v1.4**)
- Pokrytie ≥ 35
- Priemerná kvalita každej bázy nesúcej variantu (Phred skóre) ≥ 20
- Alternatívne pozorovanie (počet sekvencií nesúcich mutáciu) ≥ 5
- Klasifikácia 4 (pravdepodobne patogénna) alebo 5 (patogénna) podľa nášho nástroja na klasifikáciu patogenity variantov na základe kritérií ACMG.

Varianty TP53. Na patogénne varianty TP53 z analýzy SomaVar vykonanej na pacientovi sa uplatňujú nasledujúce kritériá:

- Pokrytie ≥ 100
- Priemerná kvalita každej bázy nesúcej variantu (Phred skóre) ≥ 20
- Alternatívne pozorovanie (počet sekvencií nesúcich mutáciu) ≥ 5
- Klasifikácia 4 (pravdepodobne patogénna) alebo 5 (patogénna) podľa nášho nástroja na klasifikáciu patogenity variantov založeného na kritériách ACMG.

Tieto varianty sa v správe analýzy SomaHRD zobrazujú len bez minimálnej prahovej hodnoty frekvencie alely, ale nie sú determinujúcimi faktormi stavu HRD. V správe HRD bude uvedená len patogénna varianta TP53 s najvyššou frekvenciou alely.

Cellularity (Buněčnost). Metadáta bunečnosti (% nádorových buniek vo vzorke) sú nevyhnutné pre spustenie analýzy SomaHRD. Možno ich zadať pri načítaní vzorky alebo pred spustením analýzy úpravou príslušných metadat vzorky (viz **Metadata vzorky**). Pokiaľ je hodnota nižšia než 20 %, výsledok nebude jednoznačný. Stav HRD však môže byť pozitívny, ak je analýza sprevádzaná panelom obsahujúcim BRCA1 a BRCA2, ktorý poskytuje pozitívny stav BRCA.

Upozornenie na frakciu nádoru Postup: Analýza SomaHRD obsahuje varovný mechanizmus, ktorý signalizuje, keď vzorka môže mať nízky obsah nádoru. Je založený na štatistickom meradle nazývanom „štatistika prekrývania“, ktoré hodnotí podobnosť medzi po sebe idúcimi

genomickými segmentmi z hľadiska pomeru počtu kópií. Vzorky s významným prekryvaním majú často nižšie frakcie nádoru, čo sťažuje rozlíšenie genomovej nestability.

Aby sme stanovili spoľahlivú hranicu pre toto varovanie, korelovali sme štatistiku prekryvania s frekvenciou variantných aliel (VAF) mutácií TP53 pomocou klinických údajov. Ak štatistika prekryvania prekročí vopred stanovenú hranicu, spustí sa varovanie. Toto varovanie upozorňuje na potenciálne nezrovnalosti s ručne zadaným percentom obsahu nádoru, čo môže naznačovať, že frakcia nádoru je nižšia ako 20 %. Vzorky s frakciami nádoru pod touto úrovňou majú zvýšené riziko vzniku falošne negatívnych výsledkov. Takéto vzorky však nie sú odmietnuté a výstraha slúži ako usmernenie pre starostlivú interpretáciu výsledkov HRD.

Minimálne pokrytie. Minimálne pokrytie potrebné na vykonanie testu je 0,1X, ale odporúčame minimálne pokrytie 0,5X, ideálne 1X. Keď sa pokrytie zníži na 0,1X, miera chyby klasifikácie sa zvýši o 5 %.

Hoci údaje sWGS mohli byť generované zo súpravy, ktorá obsahuje UMI, tieto sa počas analýzy SomaHRD nezohľadňujú. V pracovnej sade je zahrnutý krok orezávania pre adaptéry aj UMI.

Vizualizácia a interpretácia

Po dokončení analýzy sú výsledky k dispozícii z analýzy SomaHRD. Okrem metrík QC a technickej správy je správa HRD priamo dostupná z karty Správa HRD (**verzie SomaHRD v1.2 a v1.3**).

Od **verzie v1.4 pracovnej sady SomaHRD** sa výsledky analýzy zobrazujú na špeciálnej karte HRD. Každá časť karty odzrkadľuje časť správy HRD opísanej nižšie (súhrn HRD, kvalita vzorky, znaky genomovej nestability, mutácie BRCA1/BRCA2 a TP53, variácie počtu kópií).

Z tejto karty môže používateľ exportovať tieto výsledky ako súbor CSV alebo ako správu HRD vo formáte PDF. V prípade exportu do PDF môže používateľ vybrať jazyk správy HRD zo zoznamu dostupných jazykov: francúzština, španielčina, angličtina, nemčina, taliančina. Predvolený výber je angličtina.

Správa HRD je rozdelená do nasledujúcich častí.

V **hlavičke** sú uvedené metadáta pacienta a analýzy, typ vykonaného testu a stav HRD (**SomaHRD v1.2 a v1.3**) alebo stav GI (**SomaHRD $\geq$ v1.4**).

Kvalita vzorky. V tejto časti sú zhrnuté hlavné ukazovatele kvality analýzy, ako aj minimálne požadované prahové hodnoty.

Kontroly QC sú zobrazené v tabuľke. Pre každý parameter sú k dispozícii nasledujúce informácie:

- **Názov:** názov parametra
- **Hodnota :** hodnota parametra, vrátane :
 - Ikona, ktorá označuje, či je filter v poriadku, varuje alebo zlyhal
 - Hodnota parametra pre vzorku
- **Popis:** Stručný popis parametra

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.

Obrázok 38: Súhrn metrík QC týkajúcich sa vzorky sWGS v SomaHRD v1.2 (hore) a SomaHRD v1.5 (dole).

Kontroly kvality sa uplatňujú na nasledujúce parametre podľa prahových hodnôt uvedených na konci správy HRD (časť **KONTROLY KVALITY VZORKY**):

- percento nádorových buniek, založené výlučne na hodnote poskytnutej používateľom.
- priemerné pokrytie sekvenovania genómu,
- podiel správne zarovnaných čítaní, vypočítaný po kroku orezávania adaptéra (**len verzie SomaHRD v1.2 a v1.3**),
- metrika spoľahlivosti genómovej nestability (LGA a LPC) na základe údajov o pokrytie (**len verzia SomaHRD v1.2**). Táto kontrola kvality sa používa na meranie spoľahlivosti vráteného stavu HRD. Niektoré pozitívne a negatívne výsledky sú označené štítkom „nízka spoľahlivosť“, aby sa identifikovali vzorky s najnižšou spoľahlivosťou stavu.
- interval spoľahlivosti pravdepodobnosti HRD (**len verzia SomaHRD v1.3**). Táto metrika bude mať stav „Varovanie“, ak štandardná odchýlka je väčšia ako 0,07; bude mať stav Neúspešné, ak interval spoľahlivosti pravdepodobnosti HRD zahŕňa hodnotu 0,5.

- interval spoľahlivosti pravdepodobnosti GI (**len verzia SomaHRD v1.4**). Táto metrika bude mať stav „Varovanie“, ak je štandardná odchýlka väčšia ako 0,07; bude mať stav „Neúspešné“, ak interval spoľahlivosti pravdepodobnosti GI zahŕňa hodnotu 0,5.
- disperzia GI (**len verzia SomaHRD v1.5**), ktorá zobrazuje disperziu 100 výpočtov pravdepodobnosti GI. Táto metrika zobrazí stav „Warning“ (Upozornenie) s označením „Moderate“ (Stredná) vedľa hodnoty, ak je rozptyl > 2,6, a stav „Failed“ (Neúspešné) s označením „High“ (Vysoká), ak je > 2,7. V ostatných prípadoch sa vedľa skóre zobrazí označenie „Low“ (Nízka).
- v prípade podozrenia na nízky obsah nádoru (*nízka nádorová frakcia*) na základe sWGS sa zobrazí správa **WARNING** (*varovanie*). Ak tomu tak nie je, na ovládacom paneli sa zobrazí správa **NORMAL** (*normálne*).

Ak niektorá z nasledujúcich metrík QC nedosahuje očakávané prahové hodnoty zlyhania, výsledok analýzy bude nepresvedčivý (NEPRESVEDČIVÝ):

- % nádorových buniek
- pokrytie
- % správne zarovnaných čítaní (**len pre verzie SomaHRD v1.2 a v1.3**)
- GI disperzia > 2,7, **ak je tiež** $0,37 < GiP < 0,59$ (**len pre SomaHRD v1.5**)

V tomto prípade sa metriky pre istotu v genomovej nestabilite a nízkej frakcii nádoru nevypočítajú a v príslušných sekciách sa zobrazí N/A (pre **verziu SomaHRD < v1.5**).

V prípade SomaHRD v1.3, ak je metrika spoľahlivosti genomovej nestability neúspešná, t. j. pod očakávanou hranicou neúspešnosti, výsledok analýzy bude:

- POZITÍVNA NÍZKA SPOLÁHLIVOSŤ, ak je pravdepodobnosť HRD väčšia alebo rovná 50 %.
- NEGATÍVNA NÍZKA SPOLÁHLIVOSŤ, ak je pravdepodobnosť HRD menšia ako 50 %.

V tomto prípade je výsledok analýzy sprevádzaný varovaním.

Ak je však analýza sprevádzaná panelom obsahujúcim BRCA1 a BRCA2 a panel vráti pozitívny stav BRCA, stav HRD bude pozitívny, aj keď metrika QC týkajúca sa vzorky sWGS je pod prahom zlyhania (**len pre verzie SomaHRD v1.2 a v1.3**).

Variácie počtu kópií. Táto časť je venovaná variáciám počtu kópií génov CCNE1 a RAD51 z hľadiska počtu zistených kópií. Tieto sa nezohľadňujú pri výpočte pravdepodobnosti HRD a pravdepodobnosti GI.

Charakteristiky genomovej nestability. Táto časť obsahuje podrobnosti o charakteristikách, ktoré boli vypočítané na určenie genomovej nestability nádoru. Tieto charakteristiky sa používajú na určenie pravdepodobnosti SeqOne HRD (**SomaHRD v1.2 a v1.3**) a pravdepodobnosti GI (**SomaHRD verzia v1.4**). Táto časť má za cieľ vysvetliť, ako sa pravdepodobnosť vypočítava. Pre každú z funkcií sú uvedené nasledujúce informácie:

- **Názov** : Názov funkcie (napr. stav LGA alebo skóre LGA)
- **Hodnota** : Hodnota, ktorá bola vypočítaná pre pacienta
- **Popis**: Stručný popis funkcií

V prípade NEVÝPOVEDNÝCH výsledkov, súvisiacich s metrikami QC pod očakávanými prahovými hodnotami, sa časti AMPLIFICATIONS (Amplifikácie) a GENOMIC INSTABILITY FEATURES (Vlastnosti genómovej nestability) správy zobrazia sivou farbou.

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

Obrázok 39: Vlastnosti pravdepodobnosti HRD súvisiace s genomickou nestabilitou.

Mutácie BRCA1/BRCA2 a TP53. V prípade kompletného testu HRD vrátane analýzy SomaVar poskytujeme dodatočnú časť venovanú patogénnym variantom zisteným v génoch BRCA1, BRCA2 a TP53. Sú uvedené všetky patogénne varianty BRCA1 a BRCA2, ktoré zodpovedajú kritériám opísaným v časti **Metodika**. Uvedená však bude len patogénna alebo pravdepodobne patogénna varianta TP53 s najvyššou frekvenciou alely. Ostatné patogénne varianty TP53 je možné zobraziť v súvisiacej analýze SomaVar. Táto sekcia je skrytá iba pre správy sWGS.

Zistené mutácie sú uvedené v tabuľke s nasledujúcimi stĺpcami:

- **Gén a transkript**: Názov génu HGNC a ID transkriptu RefSeq
- **Varianta**: Nomenklatúry variantov HGVSC a HGVSP
- **Hodnotenie**: Hodnotenie varianty podľa kritérií ACMG (*patogénna* alebo *pravdepodobne patogénna*)
- **VAF**: Alelická frakcia varianty

Ak v prepojenej panelovej analýze neboli identifikované žiadne patogénne mutácie BRCA, v tejto časti správy sa zobrazí nasledujúca správa: „Na génoch BRCA1 a BRCA2 neboli identifikované žiadne patogénne mutácie.“

Ak v prepojenej panelovej analýze neboli identifikované žiadne patogénne mutácie TP53, v tejto časti správy sa zobrazí nasledujúca správa: „Na géne TP53 neboli identifikované žiadne patogénne mutácie.“

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%

Obrázok 40: Patogénne varianty BRCA1/BRCA2 a TP53.

Súhrn HRD. Posledná časť správy je venovaná výsledkom. Cieľom tejto časti je poskytnúť podrobný prehľad výsledkov a zobrazuje nasledujúce informácie:

- **Pravdepodobnosť HRD (SomaHRD v1.2 a v1.3):** Hodnota pravdepodobnosti HRD v rozmedzí od 0 do 1. Stav je pozitívny, ak je hodnota väčšia alebo rovná 0,5.
- **Pravdepodobnosť GI (SomaHRD verzia $\geq$ v1.4):** Hodnota pravdepodobnosti GI v rozmedzí od 0 do 1. Stav je pozitívny, ak je táto hodnota väčšia alebo rovná 0,5.
- **Stav BRCA:** stav stanovený na základe detekcie aspoň jednej patogénnej mutácie BRCA: MUTOVANÝ, DIVOKÝ TYP alebo N/A v prípade, že nebola vykonaná panelová analýza.
- **Stav HRD alebo stav GI:** POZITÍVNY, NEGATÍVNY, NEGATÍVNY S NÍZKOU ISTOTOU, POZITÍVNY S NÍZKOU ISTOTOU alebo NEKONKLÚZIVNE, ako je uvedené v hlavičke správy
- **Metóda:** Metóda (alebo režim), ktorá bola použitá na vytvorenie skóre HRD:
 - Kompletné HRD $\Rightarrow$ „WGS SHALLOW + BRCA1/2 GENE PANEL“
 - Iba Shallow $\Rightarrow$ „SHALLOW WGS ONLY“

V prípade *NEGATÍVNEJ NÍZKEJ ISTOTY, POZITÍVNEJ NÍZKEJ ISTOTY alebo NEZÁVEREČNÝCH* výsledkov sa tieto výsledky zobrazia svetlosivou farbou.

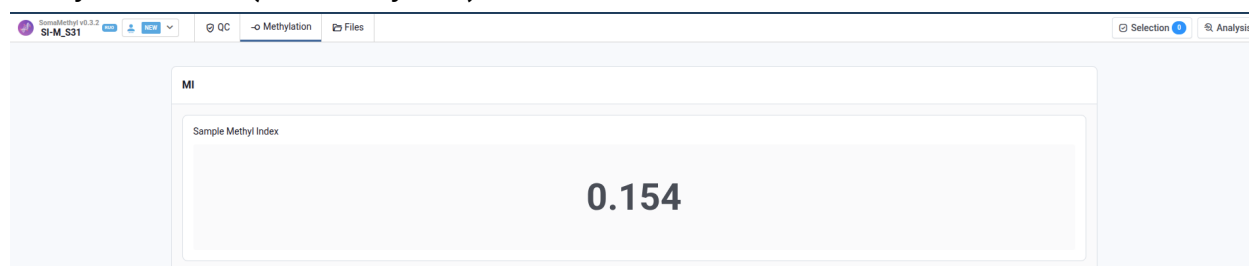
Podrobnosti o rôznych hodnotách zobrazených v správe, ako aj celková metodika testu sú uvedené v častiach „**Poznámky pod čiarou**“ a „Metodika“, ktoré sa nachádzajú za výsledkom.

Okrem správy HRD sú grafy genomovej nestability súvisiace s analýzou HRD k dispozícii na stiahnutie v karte Súborov analýzy SomaHRD. Tieto grafy predstavujú normalizované pokrytie všetkých chromozómov z údajov sWGS.

V prípade *NEKONKLÚZIVNYCH* výsledkov, súvisiacich s metrikami QC pod očakávanými prahovými hodnotami, sa grafy genomovej nestability negenerujú.

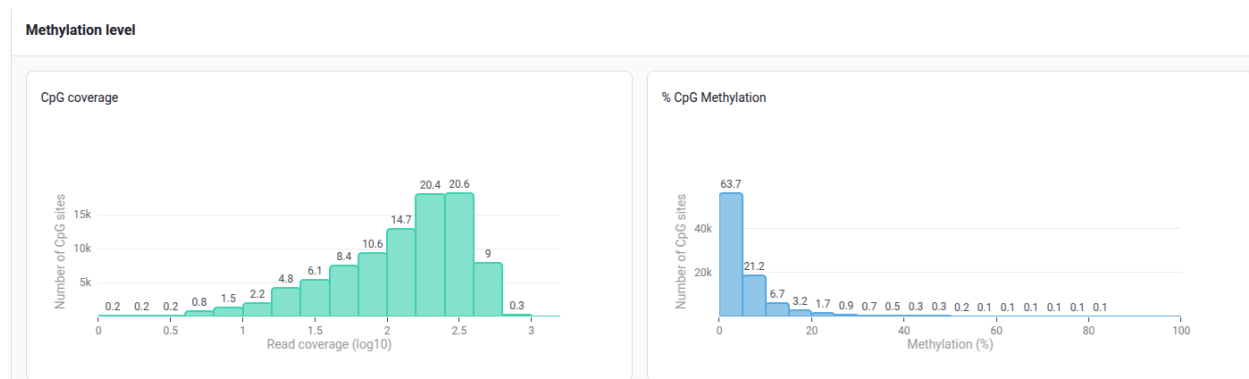
13.1.14. Metylačné dáta pre SomaMethyl

Methylation Index (Index metylácie)



Index metylácie (MI) je jediné číselné skóre, ktoré kvantifikuje celkový stav metylácie biologickej vzorky a sumarizuje signál metylácie spojený s nádorom. Dokáže citlivo detegovať aj nízke hladiny DNA pochádzajúcej z nádoru na pozadí normálnej DNA. V SomaMethyl sa MI počíta v štandardnom režime pomocou proprietárneho algoritmu spoločnosti Agilent (MethylAlgo). Keď je k dispozícii referenčná kohorta, MI vzorky sa môže porovnať s hodnotami MI kohorty a interpretovať pomocou používateľom definovanej prahovej hodnoty. Na vyhodnotenie MI spoločnosť Agilent poskytla titračné vzorky s očakávanými hladinami metylácie. Hodnoty MI vypočítané modulom SomaMethyl v štandardnom režime sa porovnávajú s očakávanými hodnotami MI poskytnutými spoločnosťou Agilent a výkon sa sumarizuje pomocou váženej metriky zhody definovanej ako $1 - wMAPE$ (Vážená priemerná absolútna percentuálna chyba).

Methylation level (Úroveň metylácie)



CpG coverage (histogram) (Pokrytie CpG (histogram))

Distribúcia hĺbky čítania naprieč miestami CpG v cieľových oblastiach (manifest).

Os X (pokrytie čítania log10): pokrytie na miesto CpG (počet čítaní podporujúcich volanie metylácie na danom CpG, t. j. počet metylovaných + nemetylovaných).

Os Y (# miest CpG): počet miest CpG v manifeste, ktoré majú toto pokrytie.

Môžete skontrolovať, či sú cieľové CpG dostatočne pokryté a či je pokrytie dostatočne rovnomerné na spoľahlivú interpretáciu percent metylácie.

% CpG methylation (histogram) (% metylácie CpG (histogram))

Distribúcia úrovni metylácie naprieč miestami CpG v cieľových oblastiach (manifest).

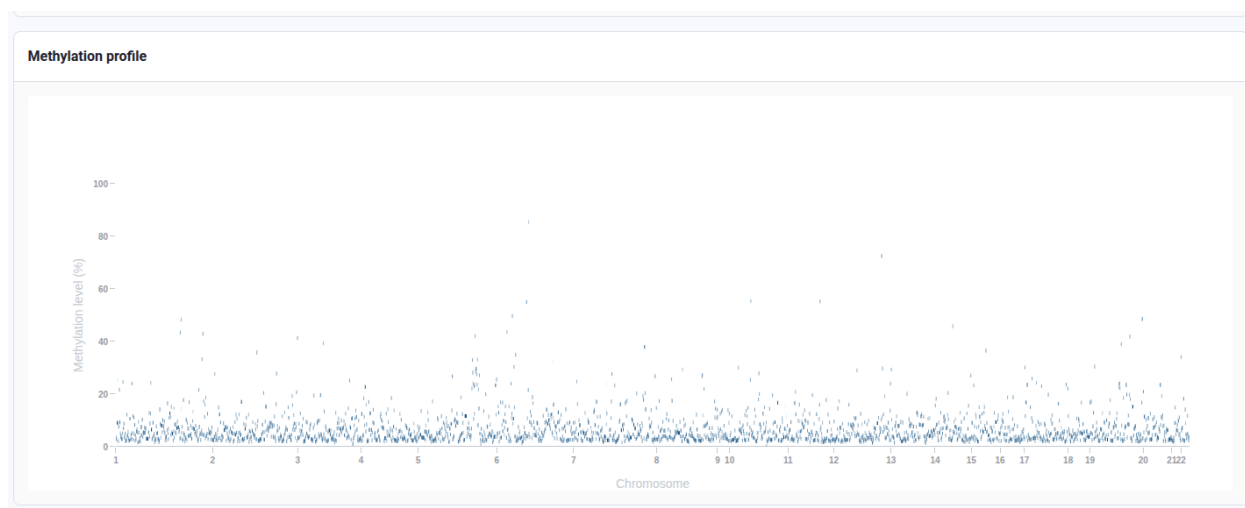
Os X (% metylácie): pre každé miesto CpG zlomok metylovaných čítaní zo všetkých čítaní pokrývajúcich toto CpG.

Os Y (# miest CpG): počet miest CpG v manifeste vykazujúcich túto úroveň metylácie.

Poskytuje pohľad na úrovni vzorky na stavy metylácie CpG naprieč cieľovými oblasťami (často ukazuje vrcholy v blízkosti nízkej a vysokej metylácie v závislosti od zachytených oblastí a biológie vzorky).

Miesta CpG s veľmi nízkym pokrytím poskytujú menej spoľahlivé percentá metylácie, pretože malý počet čítaní môže % vychýliť nahor alebo nadol. Z tohto dôvodu by sa mal histogram % metylácie CpG vždy interpretovať spoločne s histogramom pokrytia CpG (a ďalšími metrikami kontroly kvality).

Methylation profile (Profil metylácie)



Graf profilu metylácie ukazuje signál metylácie CpG vzorky naprieč genómom, zobrazený podľa chromozómu. Každý bod zodpovedá priemernému percentu metylácie CpG vypočítanému v po sebe idúcich genomických oknách (ako vo výstupe `.mcpG.track.bedGraph`). Tento pohľad pomáha vizualizovať široké vzory metylácie a identifikovať oblasti s vyššou alebo nižšou metyláciou v porovnaní so zvyškom genómu.

Methylation Table (Tabuľka metylácie)

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 ▶					

Tabuľka metylácie uvádza genomické oblasti (chromozóm:začiatok-koniec), v ktorých SomaMethyl sumarizuje metyláciu CpG (pevné okná alebo zlúčené segmenty, v závislosti od kroku). Pre každú oblasť METH. LEVEL hlási agregovanú úroveň metylácie vypočítanú z volaní metylácie CpG v tomto intervale a CpG NBR. udáva, koľko miest CpG prispelo k výpočtu. Každý riadok je potom anotovaný pomocou CpG REGIONS (ostrov, breh, šelf, kde je to vhodné) a funkčnými anotáciami, vrátane prekrývajúcich sa regulačných prvkov GENOMIC FEATURE (promótor, zosilňovač, tlmič) a informácií REFSEQ CDS (kódujúca sekvencia).

14. Správa

14.1.1. Výstupné súbory

Na karte Súbory sú uvedené súbory vytvorené v rôznych fázach worksetu. Každý súbor je možné stiahnuť kliknutím na ikonu „“ (Stiahnuť súbor) .

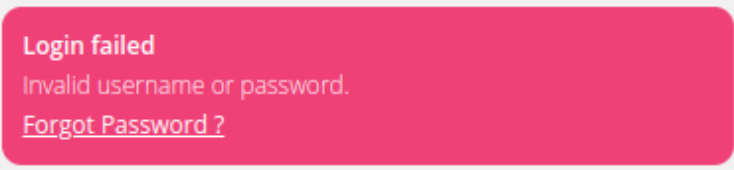
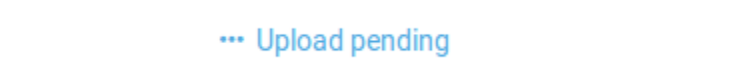
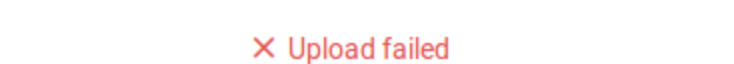
Niektoré súbory nemajú ako predponu názov vzorky. Ak ich chcete stiahnuť s touto predponou, musíte aktivovať možnosť „Prefix with sample name“ (Predpona s názvom vzorky) v pravom hornom rohu stránky.

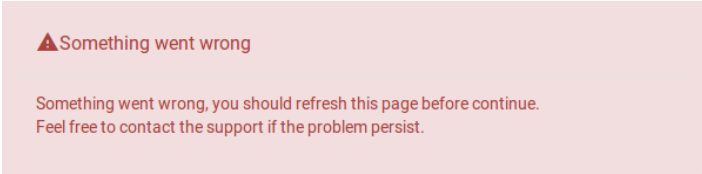
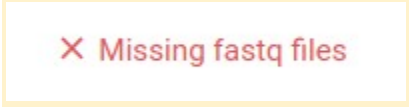
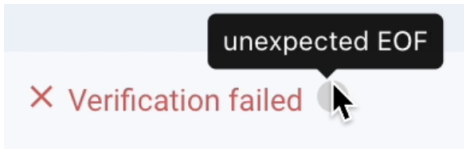
Súbory sú k dispozícii na stiahnutie z rôznych častí rozhrania:

- Súbory FastQ sú k dispozícii na karte Vzorky projektu po výbere príslušnej vzorky a následnom kliknutí na ikonu „“ (Stiahnuť súbory).
- Manifesty (súbory BED) sú k dispozícii v *nastaveniach* používateľa prostredníctvom ikony v pravom hornom rohu riadiaceho panela a potom na karte *Manifests (Manifesty)*.
- Analytické správy vo formáte PDF sú k dispozícii na karte „*Reporting*“ (Správy) po výbere analýzy.
- Súbory vytvorené počas analýzy (BAM, VCF atď.) a rôzne správy vo formáte HTML (kvalita, technické atď.) sú k dispozícii na karte *Files (Súbory)* pre každú analýzu.

15. Riešenie problémov a správy o chybách

Na upozornenie používateľov na chybu sa používa niekoľko správ

Chybová správa	Popis
 Login failed Invalid username or password. Forgot Password ?	Heslo a/alebo prihlasovacie meno je nesprávne. Používatelia sú vyzvaní, aby znovu zadali svoje prihlasovacie meno a heslo alebo klikli na „Zabudli ste heslo?“, aby si heslo obnovili.
 ... Upload pending	Nahratie vzorky je v procese. Používateľom sa odporúča počkať, nezatvárať okno prehliadača a neprerušovať pripojenie k internetu.
 ✗ Upload failed	Vzorku sa nepodarilo nahráť. Táto chyba sa môže vyskytnúť, ak bol prehliadač uzavretý alebo ak došlo k výpadku internetového pripojenia počas nahrávania. Používatelia sú vyzvaní, aby sa pokúsili nahráť súbor (súbory) znovu.
	Upozorňuje na priebeh aktuálneho nahrávania. Používateľom sa odporúča počkať, nezatvárať okno prehliadača a neprerušovať internetové pripojenie.
	Načítavanie stránky. Používatelia sú vyzvaní, aby počkali.

 Bioinformatics - <i>pending</i>	Analýza prebieha. Používatelia sú vyzvaní, aby počkali.
 Failed - <i>error</i>	Analýza sa nepodarila. Používatelia sú vyzvaní, aby skontrolovali nastavenia a údaje použité na spustenie analýzy a potom analýzu znovu spustili.
	Došlo k chybe. Používateľ je vyzvaný, aby klikol na klávesu „Esc“. Ak problém pretrváva, kontaktujte podporu na adrese support@seqone.com
	V tejto vzorke chýbajú príslušné súbory FastQ, čo znamená, že pre túto vzorku nie je možné spustiť žiadne nové analýzy.
	Správa „Unexpected EOF“ (End of File) zodpovedá skrátenému súboru. Používateľ je vyzvaný, aby skontroloval údaje a znovu načítal príslušné súbory.

Používatelia môžu kontaktovať tím podpory kliknutím na logo  v pravom hornom rohu používateľského panela a potom na ikonu  **Support** .

16. Záruka

Podmienky záruky nájdete vo vašej zmluve o údržbe.

17. Právne upozornenie

17.1. Podmienky používania

Informácie o používateľskej licencií sú dostupné po kliknutí na logo používateľského účtu a následne na položku „Podmienky používania“. Používatelia môžu zobrazíť a stiahnuť informácie vo formáte PDF z vyskakovacieho okna kliknutím na „OTVORIŤ PDF“.

17.2. Obmedzenie zodpovednosti

SeqOne odmieta akúkoľvek zodpovednosť za akékoľvek zneužitie alebo neprofesionálne používanie platformy SeqOne.

SeqOne je nástroj na podporu klinického rozhodovania a nezodpovedá za diagnózy odvodené z získaných výsledkov.

Preto musia používatelia pri používaní softvéru a jeho výsledkov postupovať s náležitou opatrnosťou. Diagnózy by mali byť potvrdené rôznymi klinickými a lekáorskými vyšetreniami.

17.3. Autorské práva a ochranné známky spoločnosti

Všetky názvy spoločností a produktov uvedené v tejto príručke môžu byť ochrannými známkami ich príslušných vlastníkov.