AMP 2023 Abstract #1577098

Improved SNV and InDel detection in an Exome NGS assay with novel Features

Jieying Chu, Rachel Schell, Faqiang Wu, Michal Krawczyk, Segun C. Jung, Hyunjun Nam, Kenneth B. Thomas, Giovanni Marsico, Fernando J. Lopez-Diaz, Chaugiang Duong, Roya Hariri, Pathum Kanagasunderam, Vincent Funari, Julie Mayer, Shashi Kulkarni, Jiannan Guo, Steven P Rivera

NeoGenomics Laboratories, Aliso Viejo, CA 92656

Introduction: Today, most known disease-causing mutations occur in exons, therefore Whole Exome Sequencing (WES) is an appropriate method to identify the vast majority of disease-causing mutations. WES target regions range from 35-45 Mbp, which makes high coverage expensive. Sequencing clinically relevant targets at deeper coverages for accurate SNV/InDel calling as an enhancement (average depth ≥ 500x, Region 1) to WES, while keeping targets outside of the Region 1 at lower coverage (average coverage ≥ 150x, Region 2) is a good solution. In order to evaluate the feasibility of SNV and InDel calling in such a scenario, we sequenced at 3500x to create a "Truth Set" variant list within the Region 1, and applied it to a hybridization capture WES with Region 1 at 500x, and again to the same samples with WES only (150x). 50ng of DNA was used for detecting somatic SNV/Indel variants in both Region 1 and 2, with as low as 30% tumor purity. Additionally, we incorporated a novel single stranded library prep method that enabled an agnostic workflow for hematological and FFPE sample types.

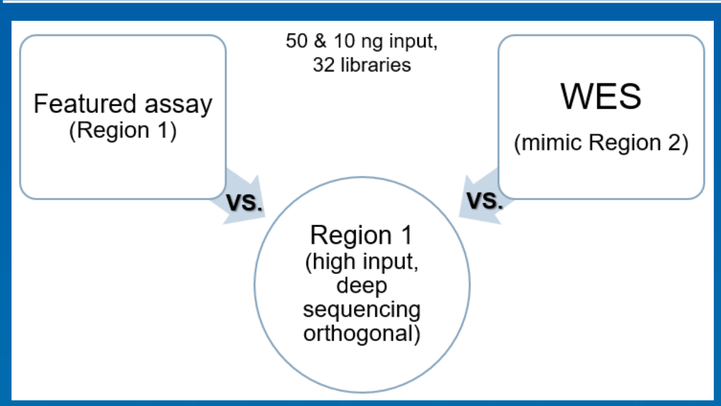
Methods: The Truth Set of SNPs and InDels for eight clinical samples was created by sequencing Region 1 at > 3500x with 200ng DNA. The same samples underwent WES with Region 1 (for Region 1) and WES only (for simulating Region 2) library construction with 50ng DNA and sequenced. These samples served as our 'experimentals' for detection rate of the Truth Set variants. All sequencing data was processed with an in house bioinformatic pipeline for SNV/Indel calling. 500x average coverage for Region 1 and 150x for WES simulating Region 2 were achieved through downsampling and compared to the Truth Set for sensitivity and specificity calculation. For the truth set, the SNV/Indel calling cutoff was AF>= 5% and DP >= 700; whereas for Region 1 it was AF>= 5%, DP >= 100, and AltRead >= 10, and for Region 2 it's AF>= 10% and DP >= 50.

Results: For SNV/Indel calling with 50ng input in Region 1, the assay achieved 98-99% sensitivity and 97-99% specificity. In Region 2, the sensitivity was 94-98%, with the specificity being 97-99%. With as low as 10ng input, Region 1 sensitivity was 97 – 98%, and specificity was 97 – 99%. In Region 2, 10ng input sensitivity was 93 – 98%, and the specificity was 97 – 99%.

Conclusions: This study demonstrated improved SNV and InDel detection using a novel Exome NGS assay with 50ng DNA input (minimum 10ng input). Good performance was achieved in Region 1 for clinically relevant genes at 5% VAF and 10% VAF in the rest of the exome simultaneously.

Jieying.Chu@neogenomics.com

Experiment Design and Sample List



	Body Site	Tumor (%)	TMB	MSI
InD-1	Stomach	60	LOW	MSS
InD-2	Lung, Right Uppe Lobe, Posterior Segment	30	High	MSS
InD-3	Colon, Right Hemicolectomy	70	High	MSI-H
InD-4	Uterus With Bilateral Adnexa	40	High	Indeterminate
InD-6	Colon, Transverse	30	INTERMEDIATE	MSS
InD-9	Colon	60	LOW	MSS
InD-10	Right Lung	80	INTERMEDIATE	MSS
InD-12	Left Parietal-Occipital Mass	90	HIGH	MSS

Good SNV/Indel calling sensitivity and specificity in Region 1 with 500X average coverage

Samples	1100 true set (AF>= 5%, DP >= 700)	Enhanced – 50ng (AF>= 5%, DP >= 100, AltRead >= 10)	Enhanced -10ng (AF>= 5%, DP >= 100, AltRead >= 10)	10ng retention
InD-1	1603	1584 (98.8%)	1566 (97.7%)	98.9%
InD-2	1585	1549 (97.7%)	1542 (97.3%)	99.5%
InD-3	1860	1822 (98%)	1824 (98.1%)	99.1%
InD-4	1819	1792 (98.5%)	1779 (97.8%)	98.9%
InD-6	1580	1540 (97.5%)	1538 (97.3%)	98.9%
InD-9	1614	1592 (98.6%)	1588 (98.4%)	99.6%
InD-10	1670	1644 (98.4%)	1640 (98.2%)	99.3%
InD-12	1679	1644 (97.9%)	1639 (97.6%)	98.5%

Samples	TN_50ng	FP_50ng	50ng_Specificity	TN_10ng	FP_10ng	10ng_Specificity
InD-1	3067	37	99%	3071	52	98%
InD-2	3086	35	99%	3088	45	99%
InD-3	2840	41	99%	2842	42	99%
InD-4	2891	47	98%	2891	46	98%
InD-6	3139	29	99%	3139	53	98%
InD-9	3137	58	98%	3150	85	97%
InD-10	2844	26	99%	2844	23	99%
InD-12	2985	88	97%	2999	70	98%

Red: average coverage < 500x

Good SNV/Indel calling sensitivity and specificity in Region 2 with 150X average coverage

Samples	1100 true set (AF>= 10%, DP >= 700)	Non-Enhanced – 50ng (150X, AF>= 10%, DP >= 50)	Non-Enhanced – 10ng (150X, AF>= 10%, DP >= 50)	10ng retention
InD-1	1402	1371 (97.8%)	1362 (97.1%)	98.8%
InD-2	1370	1327 (96.9%)	1330 (97.1%)	99.1%
InD-3	1635	1597 (97.7%)	1595 (97.6%)	98.9%
InD-4	1588	1555 (97.9%)	1541 (97%)	98.8%
InD-6	1377	1346 (97.7%)	1330 (96.6%)	98%
InD-9	1382	1336 (96.7%)	1317 (95.3%)	97.6%
InD-10	1443	1395 (96.7%)	1384 (95.9%)	98.4%
InD-12	1409	1321 (93.8%)	1310 (93%)	96.6%

Samples	TN_50ng	FP_50ng	50ng_Specificity	TN_10ng	FP_10ng	10ng_Specificity
InD-1	3067	37	99%	3071	52	98%
InD-2	3086	35	99%	3088	45	99%
InD-3	2840	41	99%	2842	42	99%
InD-4	2891	47	98%	2891	46	98%
InD-6	3139	29	99%	3139	53	98%
InD-9	3137	58	98%	3150	85	97%
InD-10	2844	26	99%	2844	23	99%
InD-12	2985	88	97%	2999	70	98%