

# Automation of High-Throughput Whole Exome Library Preparation

Martin Kašný<sup>1</sup>, Irem Mertová<sup>1</sup>, Karolína Stivínová<sup>1</sup>, Alžběta Hamplová<sup>1</sup>, Markéta Kovalová<sup>1</sup>, Eliška Hladíková<sup>1</sup>, Nela Chalupníková<sup>1</sup>, Petr Kvapil<sup>1</sup>

<sup>1</sup> Institute of Applied Biotechnologies, Služeb 3056/4, 108 00 Prague, Czech Republic  
\* Correspondence: kasny@iabio.eu; Tel.: +420 739 394 364

Whole-exome sequencing (WES) is a powerful clinical tool for detecting pathogenic variants in human genes. Manual preparation of NGS libraries at a large scale is problematic in terms of data quality guarantee and it is also a time-consuming process. To enable high quality (HQ) and high throughput (HT) screening, IAB developed automated WES library preparation using CyBio Felix – Automated Liquid Handling Platform.

## Solution & Panel Customization

### Illumina DNA Prep with Enrichment Dx

- EU IVDR 2017/746-compliant solution & FDA-regulated
- Input: 50 – 1000 ng per reaction
- 12 samples per plex (enrichment)

### Twist Alliance VCGS Exome (40.1 Mbp)

- Expert-defined customized and validated inherited and rare disease content (Victorian Clinical Genetics Services (VCGS)).
- Minimized gaps in clinically relevant genes.
- Targets difficult to cover, non-coding regions known to contain pathogenic variants, resulting in high-quality actionable data for clinicians.
- Customizable Twist spike-in panel



**1** Up to 192 DNA samples processed in automated workflow.

**4** Libraries **hybridized** with probes; targeted regions captured and subsequently amplified.

2h 15m manual | 5h 15m automated

**2** DNA samples **fragmented** and tagged in one rapid reaction; amplified; **sequencing indexes** are added in PCR.

1h 15m manual | 4h 30m automated



**5** **Quality control checked** to ensure enriched libraries in plexes met the sequencing criteria.

1h manual Quality control

**3** DNA libraries **purified** and **plexed** into pool of up to 12 libraries.

\*2h manual | 2h automated  
\*Including manual Quality control

**6** Libraries ready for sequencing.

## Results & Performance Evaluation

Parameters (Dragen Coverage QC Metrics)	CUSTOMIZED WES (VCGS Exome + Spike-in Panel + mtDNA Panel)	CARRIER SCREENING REGIONS	HEREDITARY CANCER REGIONS
Average Alignment Coverage over QC Coverage Region	219.48	187.00	194.00
Uniformity of coverage (% > 0.2*mean) over QC Coverage Region	98.67	96.42	99.23
Uniformity of coverage (% > 0.4*mean) over QC Coverage Region	95.26	93.09	96.39
<b>PCT of target with coverage &gt; 100x [%]</b>	<b>92.43</b>	<b>86.87</b>	<b>90.65</b>
PCT of target with coverage > 50x [%]	98.41	95.52	98.81
PCT of target with coverage > 20x [%]	99.33	97.62	99.54

### Coverage & Uniformity Analysis, Performance Evaluation

Selected parameters from coverage analysis of virtual panels performed by DRAGEN v3.10. The parameter "PCT of target with coverage > 100x" reflects the request for 92% of bases covered 100X. The Uniformity of coverage (PCT >0,2\* mean or PCT >0,4\* mean) over QC region shows very good evenness of the coverage depth among different regions.

According to the performance of the tailored WES panel, the ~120M PE reads/sample is a valid target value of average seq. capacity/sample. Complete results in Supplementary Table 2.

## Conclusion

Optimization of the automated whole exome library preparation was successful. Our results show high interplex uniformity and we have exceeded the target of ~ 92% of bases covered 100x. Using Illumina DNA Prep with Enrichment Dx together with Twist Alliance VCGS Exome and a custom spike-in panel, our method covers a wide range of clinically relevant genes. Combined with high uniformity of results, reduced hands-on time and redundancy of this solution, library preparation on CyBio Felix automated liquid handlers is a reliable option for high-throughput WES analysis of 192 samples, not only in clinical laboratories.

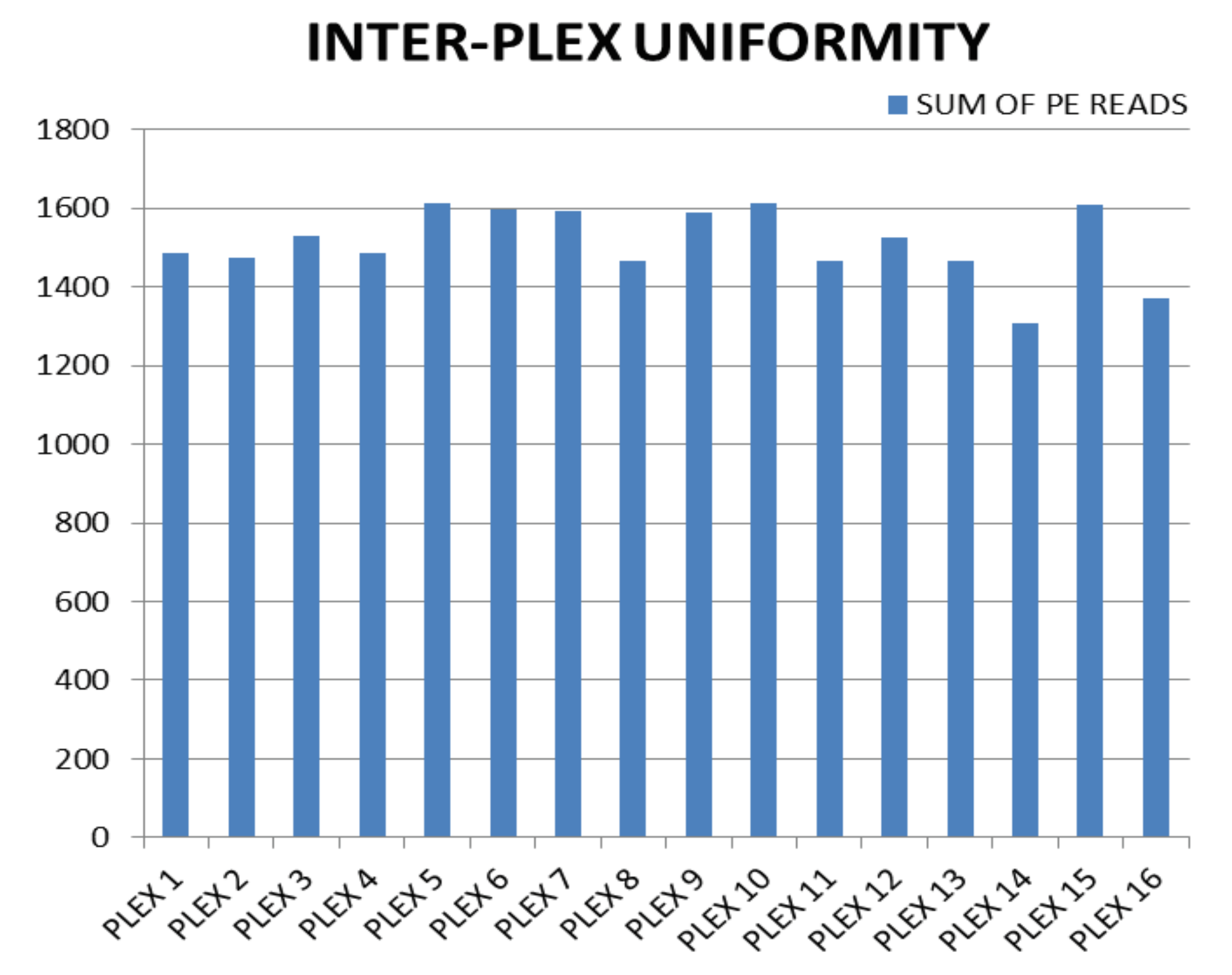
# Automation of High-Throughput Whole Exome Library Preparation

Martin Kašný<sup>1</sup>\*, Irem Mertová<sup>1</sup>, Karolína Stivínová<sup>1</sup>, Alžběta Hamplová<sup>1</sup>, Markéta Kovalová<sup>1</sup>, Eliška Hladíková<sup>1</sup>, Nela Chalupníková<sup>1</sup>, Petr Kvapil<sup>1</sup>

<sup>1</sup> Institute of Applied Biotechnologies, Služeb 3056/4, 108 00 Prague, Czech Republic  
\* Correspondence: kasny@iabio.eu; Tel.: +420 739 394 364

PLEXES	MIN (to AVRG) [%]	MAX (to AVRG) [%]	MAX - MIN DIFFERENCE [%]
Plex 1 IDPE	78	117	39
Plex 2 IDPE	81	115	33
Plex 3 IDPE	84	116	32
Plex 4 IDPE	86	118	32
Plex 5 IDPE	91	110	19
Plex 6 IDPE	85	117	33
Plex 7 IDPE	87	116	29
Plex 8 IDPE	91	119	28
Plex 9 IDPE	74	120	46
Plex 10 IDPE	64	186	122
Plex 11 IDPE	86	117	30
Plex 12 IDPE	82	127	45
Plex 13 IDPE	84	114	30
Plex 14 IDPE	82	118	36
Plex 15 IDPE	89	110	20
Plex 16 IDPE	89	116	27

## Optimization of Inter-Plex Uniformity



Supplementary Table 1. Intra-Plex variability

Intra-Plex uniformity is shown in Supplementary Table 1.

## Coverage & Uniformity Analysis, Performance Evaluation

Selected parameters from coverage analysis of virtual panels performed by DRAGEN v3.10. The parameter "PCT of target with coverage > 100x" reflects the request for 92% of bases covered 100X. The Uniformity of coverage (PCT > 0,2\* mean or PCT > 0,4\* mean) over QC region shows very good evenness of the coverage depth among different regions.

FEMALES 109 SAMPLES ♀	Total input reads	Duplicate Marked Reads [%]	Mapped reads [%]	Reads with MAPQ > 0.4 [%]	Insert Length: Mean	TARGETS PCT of QC Coverage Region with Coverage > 100x [%]	TARGET Average Alignment Coverage over QC Coverage Region	CZE PCT of QC Coverage Region with Coverage > 100x [%]	CZE Average Alignment Coverage over QC Coverage Region	CARR PCT of QC Coverage Region with Coverage > 100x [%]	CARR Average Alignment Coverage over QC Coverage Region
MIN	85 777 098	8,16	99,94	87,73	196	78,53	156,2	72,06	136,3	77,26	140,9
MAY	159 807 198	14,09	99,99	89,74	255	97,09	280,5	96,63	248,9	94,24	260,9
<b>AVERAGE</b>	<b>122 785 435</b>	<b>10,60</b>	<b>99,06</b>	<b>87,92</b>	<b>216</b>	<b>91,39</b>	<b>216,5</b>	<b>89,35</b>	<b>189,9</b>	<b>89,83</b>	<b>195,3</b>
<b>MEDIAN</b>	<b>123 507 394</b>	<b>10,69</b>	<b>99,97</b>	<b>88,75</b>	<b>218</b>	<b>93,03</b>	<b>220,2</b>	<b>91,23</b>	<b>191,6</b>	<b>91,29</b>	<b>195,1</b>
AVERAGE DEVIATION	12 116 297	0,85	0,01	0,35	9	2,52	20,1	3,48	17,8	2,03	18,5

MALES 83 SAMPLES ♂	Total input reads	Duplicate Marked Reads [%]	Mapped reads [%]	Reads with MAPQ > 0.4 [%]	Insert Length: Mean	TARGETS PCT of QC Coverage Region with Coverage > 100x [%]	TARGET Average Alignment Coverage over QC Coverage Region	CZE PCT of QC Coverage Region with Coverage > 100x [%]	CZE Average Alignment Coverage over QC Coverage Region	CARR PCT of QC Coverage Region with Coverage > 100x [%]	CARR Average Alignment Coverage over QC Coverage Region
MIN	100 942 810	8,97	99,93	87,45	197	83,45	175,9	83,26	157,8	61,65	136,0
MAY	248 320 890	16,23	99,98	89,71	238	98,36	397,1	98,66	358,5	94,38	310,5
<b>AVERAGE</b>	<b>124 071 993</b>	<b>10,79</b>	<b>98,77</b>	<b>87,28</b>	<b>215</b>	<b>91,56</b>	<b>218,0</b>	<b>90,16</b>	<b>194,9</b>	<b>76,20</b>	<b>170,4</b>
<b>MEDIAN</b>	<b>123 654 164</b>	<b>10,88</b>	<b>99,97</b>	<b>88,28</b>	<b>218</b>	<b>93,10</b>	<b>215,9</b>	<b>91,57</b>	<b>194,0</b>	<b>76,29</b>	<b>170,9</b>
AVERAGE DEVIATION	11 772 579	0,80	0,01	0,41	8	1,98	18,7	2,45	16,3	4,89	14,6

Supplementary Table 1. Comparison of female and male sample sequencing and mapping metrics

## Performance Evaluation, ♀♂ Comparison

Coverage metrics provided by DRAGEN V3.10. Differences between female and male samples of gDNA can relate to the custom spike-in panel that target various regions in the X and Y chromosome. Essential is the "TARGET PCT of QC coverage region with coverage (100X:inf)" representing the portion (%) of covered bases at targeted regions: CZE – hereditary cancer Czech panel, CARR – Carrier Screening panel.