

Государственное научное учреждение
«Институт генетики и цитологии Национальной академии наук Беларуси»

Молекулярная генетика и биоинформатика

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОГРАММНОГО
ОБЕСПЕЧЕНИЯ ДЛЯ ОБРАБОТКИ ДАННЫХ
ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ**

Левданский Олег Дмитриевич
заведующий сектором биоинформатики
кандидат биологических наук
o.liaudanski@igc.by

Молчун Анна Михайловна
a.molchun@igc.by

Минск, 2026

Высокопроизводительное секвенирование





De novo сборка

Поиск SNP

CNV анализ

Аннотация

Метагеномный анализ

...

Strategy	Variant callers
Alignment and pre-processing	
Read alignment	BWA-MEM [25], Bowtie 2 [26], minimap2 [27], Novoalign
Marking duplicates	Picard tools [28], Sambamba [29], SAMBLASTER [30]
BAM file creation	Samtools [31], GATK [19]
Sequencing metrics	BEDTools [32], Picard tools [28], QualiMap 2 [33]
Sample quality control	KING [34], VerifyBamID [35]
Variant calling	
Inherited SNVs/indels	FreeBayes [36], GATK HaplotypeCaller [19], Platypus [20], Samtools/BCFtools [37]
Somatic mutations	deepSNV [38], MuSE [39], MuTect2 [40], SomaticSniper [41], Strelka2 [42], VarDict [43], VarScan2 [44]
Copy number variants	cn.MOPS [45], CONTRA [46], CoNVEX [47], ExomeCNV [48], ExomeDepth [49],XHMM [50]
Structural variants	DELLY [51], Lumpy [52], Manta [53], Pindel [54], SVMerge [55]
Gene fusions (RNA-seq)	fusionCatcher [56], fusionMap [57], mapSplice [58], SOAPfuse [59], STAR-Fusion [60], TopHat-Fusion [61]
Variant review/storage	
Visualization and review	Artemis [62], Integrative Genomics Viewer [63]
VCF/BCF file manipulation	BCFtools [37]

Сборка геномов *de novo*

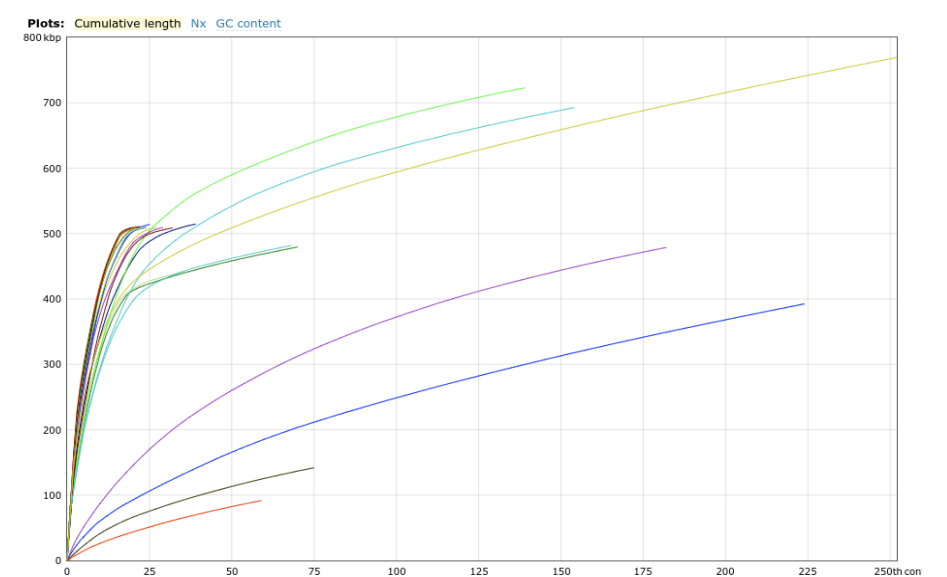
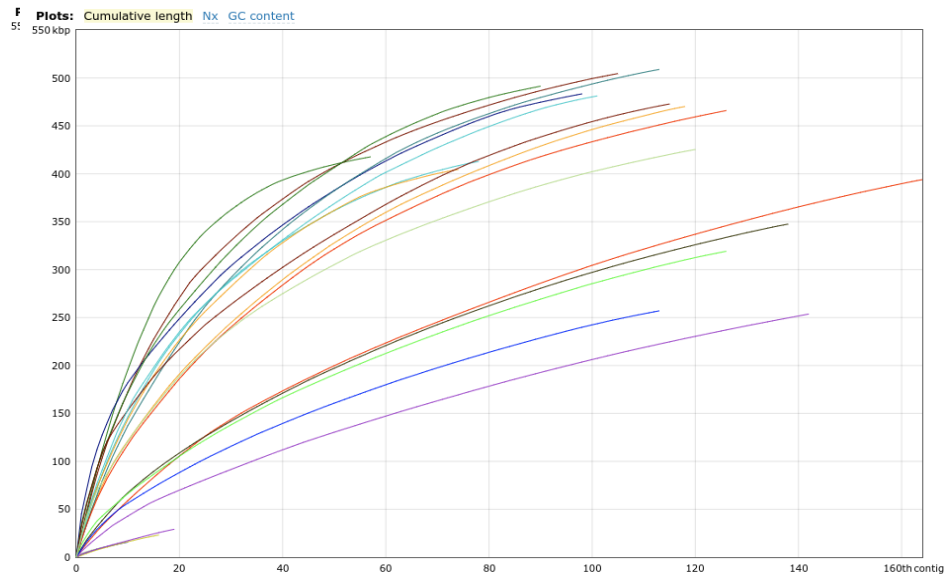
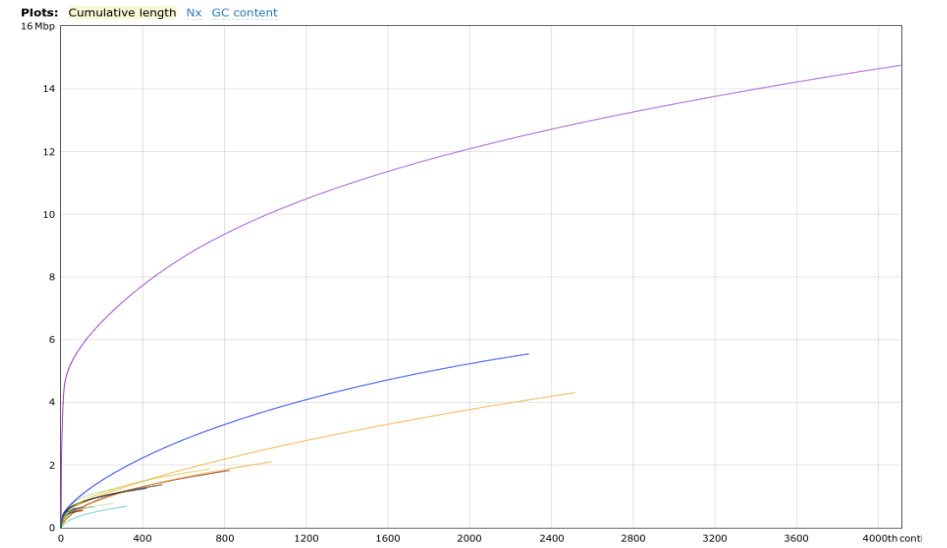
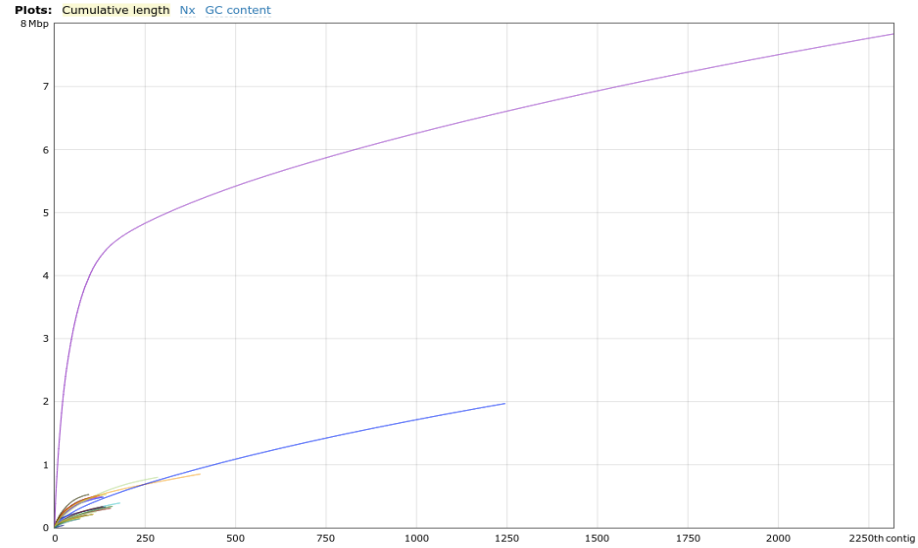
- 47 образцов хлоропластной и митохондриальной ДНК, полученные на приборе MiSeq (Illumina) с использованием набора реагентов Nextera XT DNA Library Preparation Kit (Illumina).
- FastQC (v0.11.9)
- Trimmomatic (v0.39)
- ABySS (v2.3.9)
- SPAdes (v3.13.1)
- NOVOPlasty (v4.3.5)
- QUAST (v5.2.0)

Сборка геномов *de novo*

Распределение суммарной длины контигов

AbySS

SPAdes

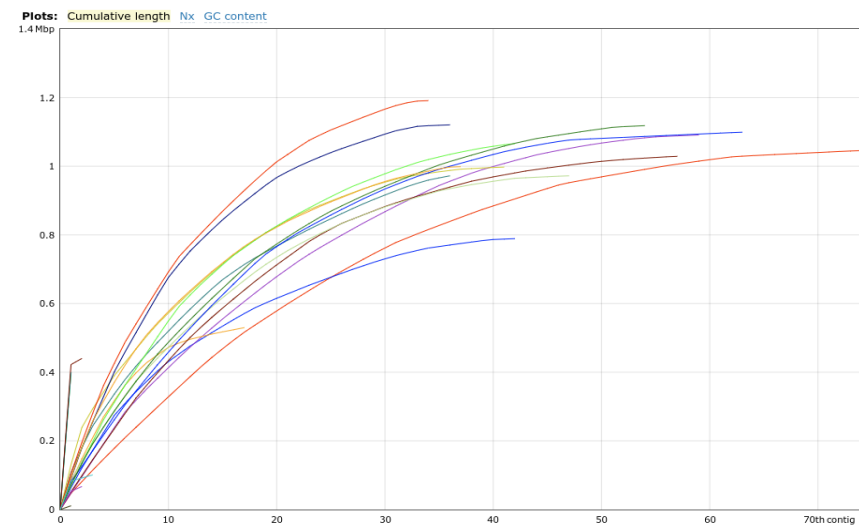
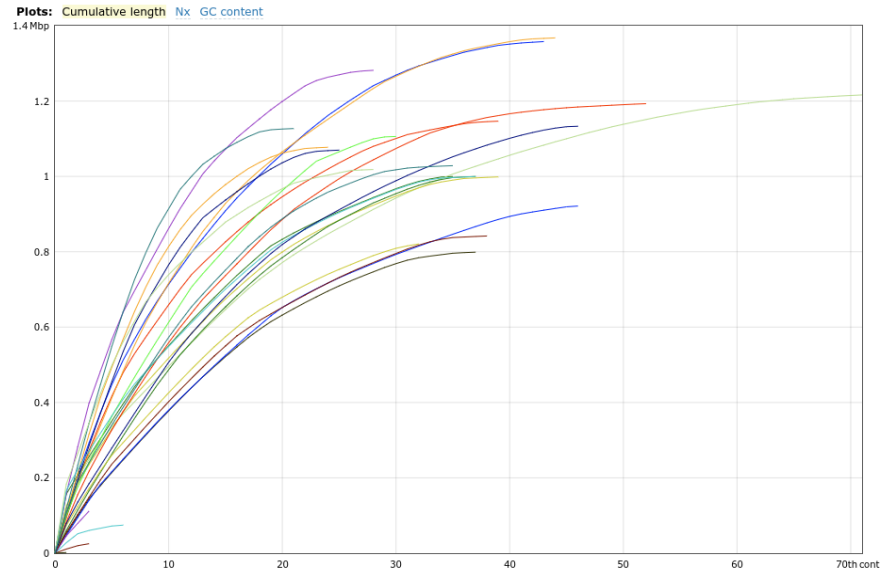


Сборка геномов *de novo*

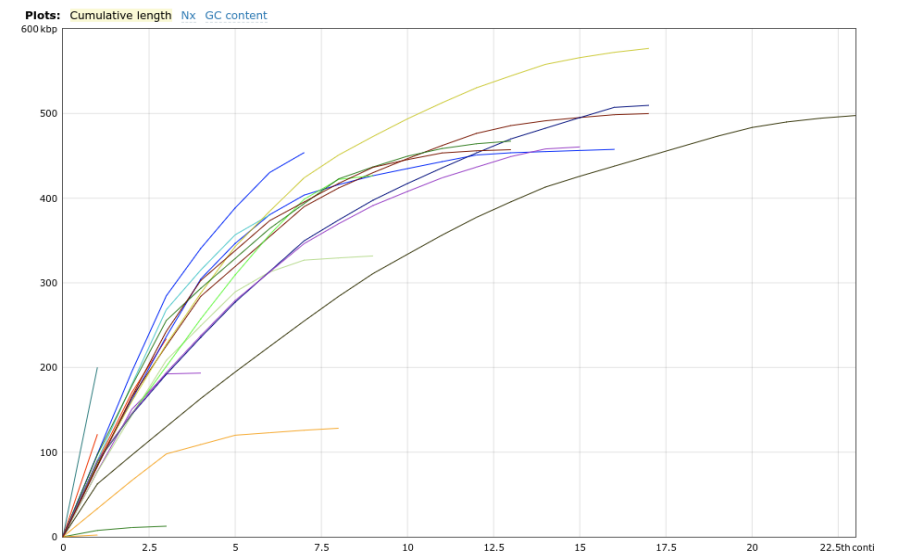
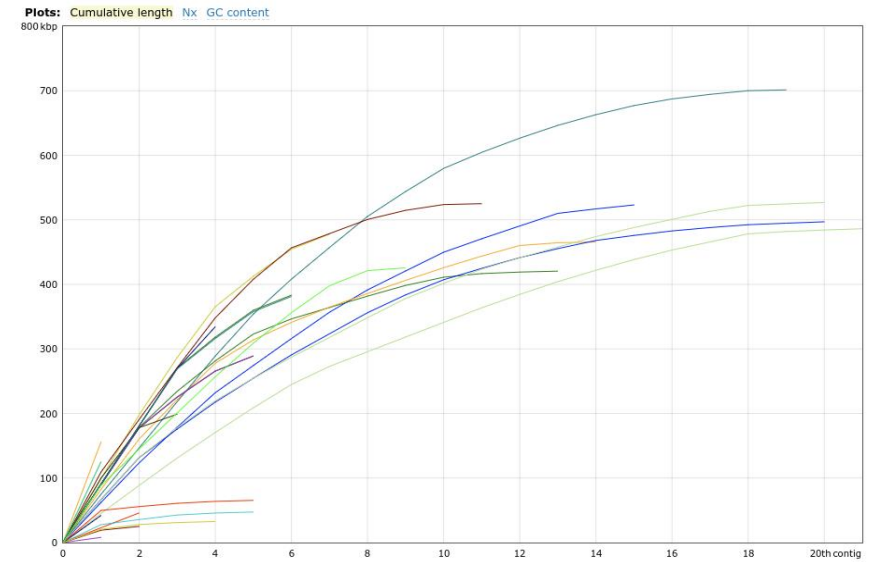
Распределение суммарной длины контигов

NOVOPlasty

mtDNA



hpDNA



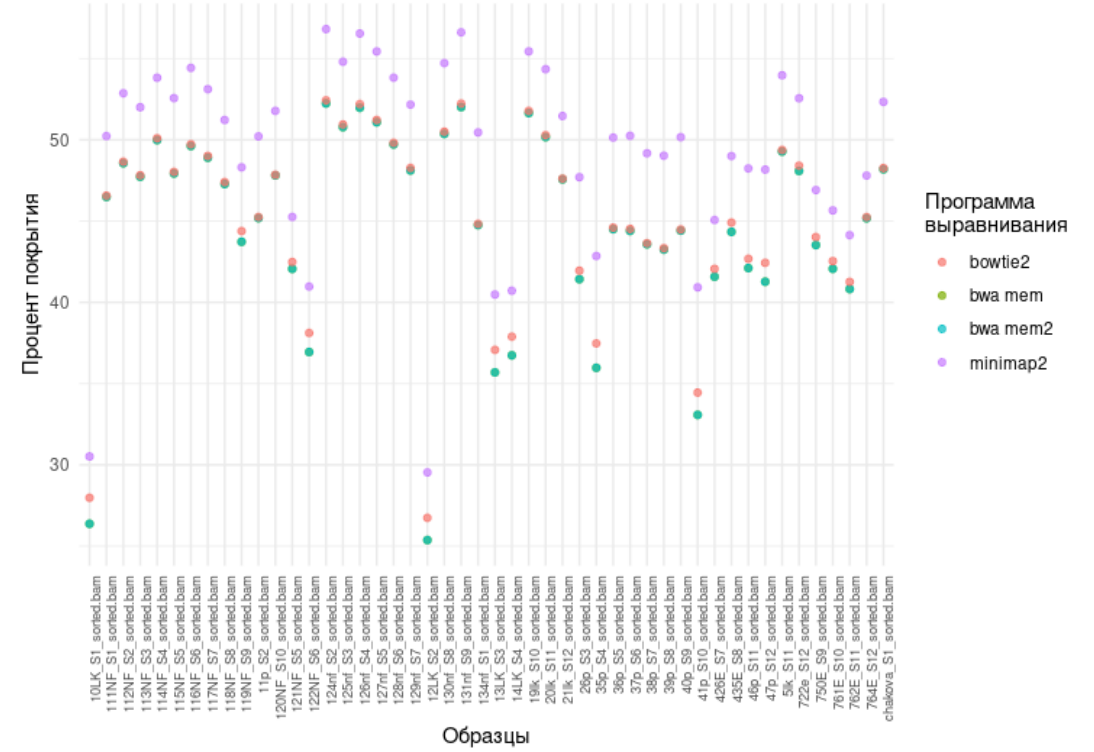
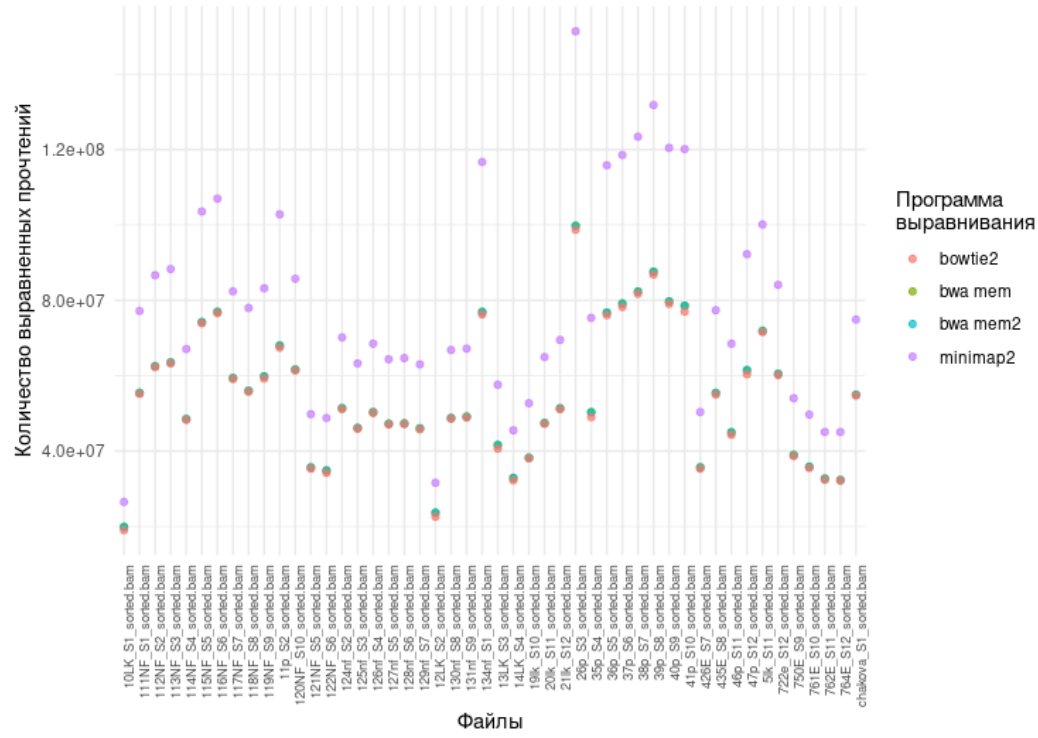
Сборка геномов *de novo*

Название программы и количество прочтений		Кол-во контигов	Наибольший контиг (тыс. п.н.)	Суммарная длина (тыс. п.н.)	N50 (тыс. п.н.)	Время (мин.)
SPAdes	Высокое количество прочтений	242,9±293,1	61,2±24,0	837,9±517,8	16,6±10,1	6,72±11,38
	Низкое количество прочтений	51±55,9	60,7±32,9	431,5±156,9	23,8±14,2	0,78±0,33
ABuSS	Высокое количество прочтений	105,4±47,7	14,1±8,3	307,9±174,2	3,5±1,8	1,75±1,57
	Низкое количество прочтений	88,4±48,9	16,4±12,4	319,5±179,8	4,6±3,3	0,74±0,12
NOVOPlasty (митохондр.)	Высокое количество прочтений	26,1±13,6	92,8±49,7	867,2±426,8	42,5±21,7	7 568,3± 3 610,0
	Низкое количество прочтений	26,8±17,9	131,7±134,5	805,1±384,0	100,0±147,1	7 101,9± 4 671,3
NOVOPlasty (хлоропл.)	Высокое количество прочтений	6,1±5,8	75,0±39,6	266,4±211,5	67,0±39,5	1 706,4± 1 041,8
	Низкое количество прочтений	8,8±6,5	86,6±40,1	335,9±162,9	71,4±44,1	1 570,4± 953,8

Выравнивание + Вызов вариантов

- 64 образца данных секвенирования экзомных последовательностей человека, полученные на приборе NextSeq (Illumina).
- FastQC (v0.11.9)
- Trimmomatic (v0.39)
- Bwa-mem (v0.7.17)
- Bwa-mem2 (v2.2)
- Bowtie2 (v2.5.4)
- Minimap2 (v2.26)
- GATK HaplotypeCaller (v4.1.4.1)
- FreeBayes (v1.3.10)
- Samtools / Bcftools (v1.22)

Выравнивание



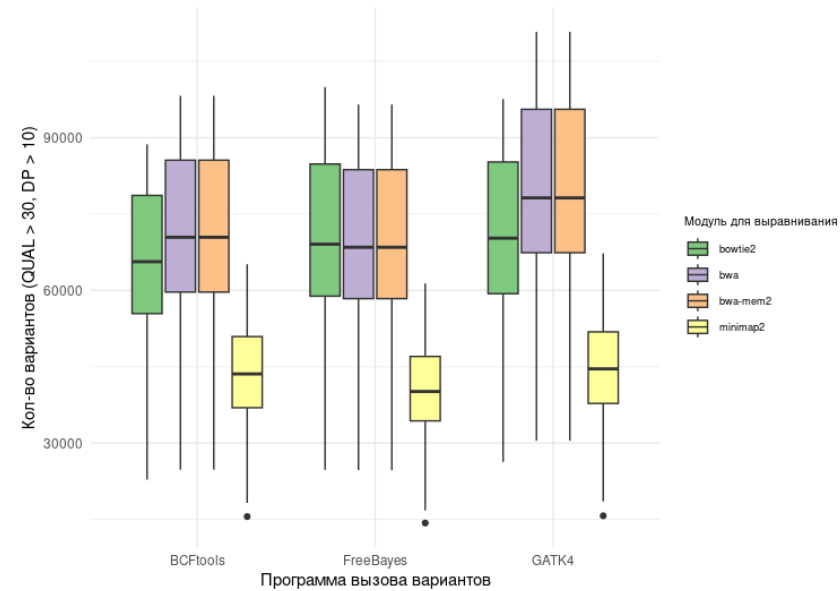
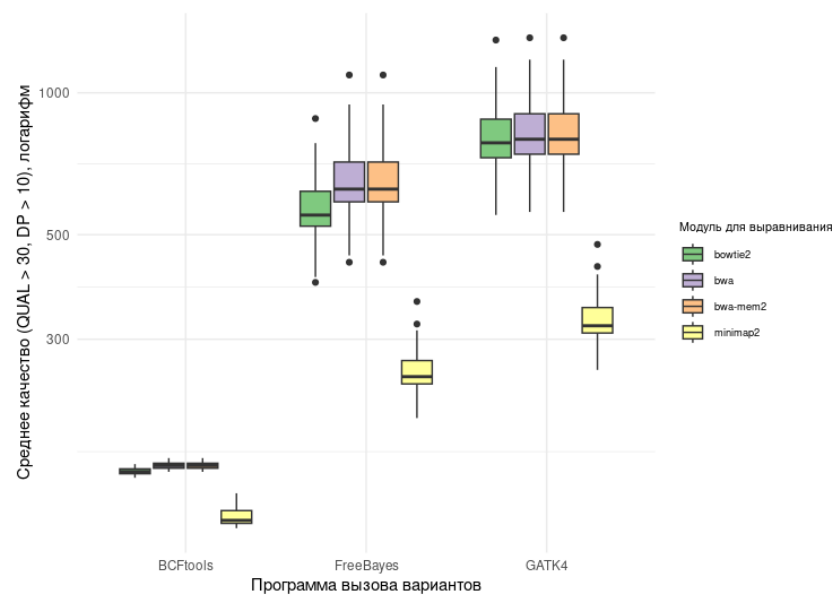
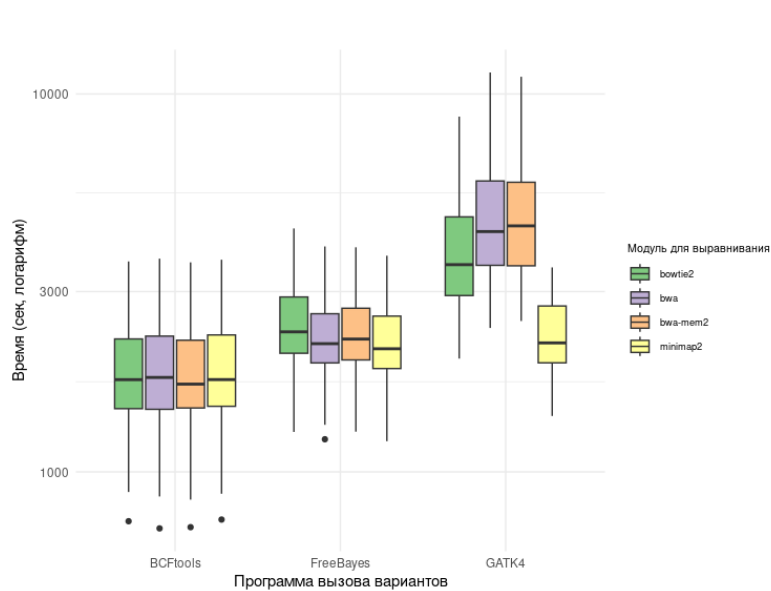
Выравнивание

Программа	Время (сек)	Глубина покрытия	Процент покрытия	Кол-во прочтений (млн.)
bowtie2	347,7±106,5	20,7±4,2	88,7±9,3	54,4±17,6
bwa mem	237,4±66,4	20,5±4,2	88,8±9,0	54,9±17,7
bwa mem2	216,0±58,3	20,5±4,2	88,8±9,0	54,9±17,7
minimap2	242,6±47,2	22,5±4,9	88,7±9,2	78,2±27,7

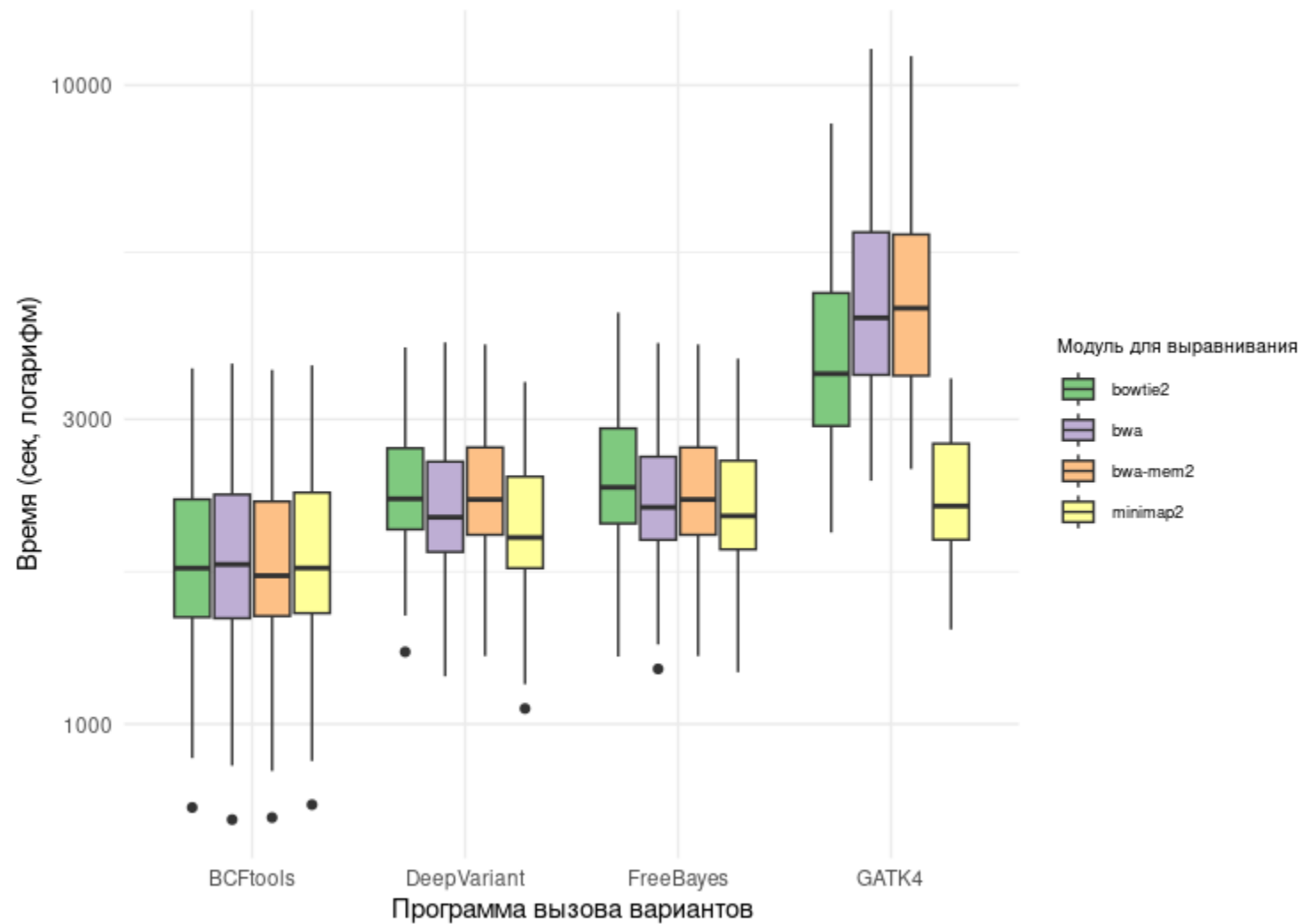
Вызов вариантов

Программа вызова вариантов	Программа для выравнивания	Время (сек.)	Количество (тыс.)	Качество
BCFtools	bowtie2	1866,6±620,6	65,5±15,3	157,4±2,6
BCFtools	bwa	1884,0±640,5	71,3±16,9	162,0±2,8
BCFtools	bwa-mem2	1848,2±630,9	71,3±16,9	162,0080833
BCFtools	minimap2	1933,3±683,0	44,1±11,5	126,3±6,2
FreeBayes	bowtie2	2485,4±673,8	70,4±17,3	563,8±88,9
FreeBayes	bwa	2308,7±599,0	70,0±16,5	646,8±119,8
FreeBayes	bwa-mem2	2350,2±590,9	70,0±16,5	646,8120625
FreeBayes	minimap2	2261,6±594,9	40,9±11,0	257,4±31,3
GATK4	bowtie2	3941,9±1510,6	71,1±16,4	796,3±135,5
GATK4	bwa	4961,6±2102,1	80,2±18,4	814,5±143,4
GATK4	bwa-mem2	4920,4±2041,2	80,2±18,4	814,5344583
GATK4	minimap2	2320,6±517,9	45,3±12,1	333,2±44,6

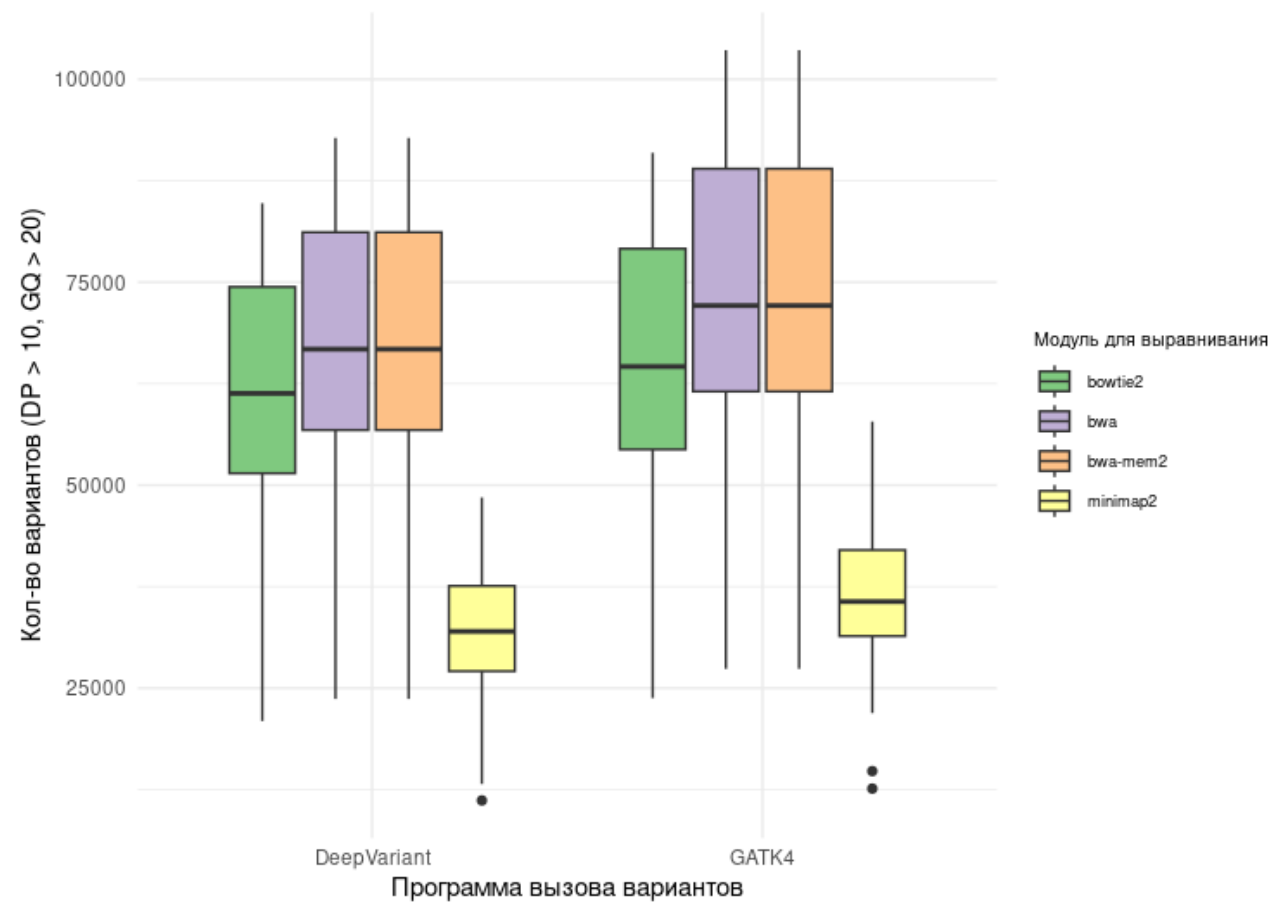
Вызов вариантов



DeepVariant



DeepVariant



БЛАГОДАРНОСТИ

Михаленко Елена Петровна

Щаюк Анна Николаевна

Мазур Оксана Чеславовна

Малышева Ольга Михайловна

Синявская Марина Георгиевна

Александрович Валерия Вадимовна

Спасибо за внимание!