

Analyse multiplex des ataxies héréditaires par séquençage long-read

I. Maladies à expansions – problématique

II. Approche long-read

- i. Crispr/Cas9 enrichment
- ii. Adaptive sampling

III. Multiplexage

IV. Conclusion - Perspectives

I. Maladies à expansions

- Répétitions en tandem → 10 à 15% du génome humain
- Polymorphiques, leur instabilité augmente avec leur longueur
- Anticipation → augmentation de la taille d'une génération à l'autre
- Plus d'une cinquantaine de maladies à expansions décrites
 - Plusieurs mécanismes pathogènes rapportés
 - Majoritairement des maladies neurodégénératives
 - > Dont 16 qui correspondent à des ataxies héréditaires dominantes ou récessives

I. Maladies à expansions

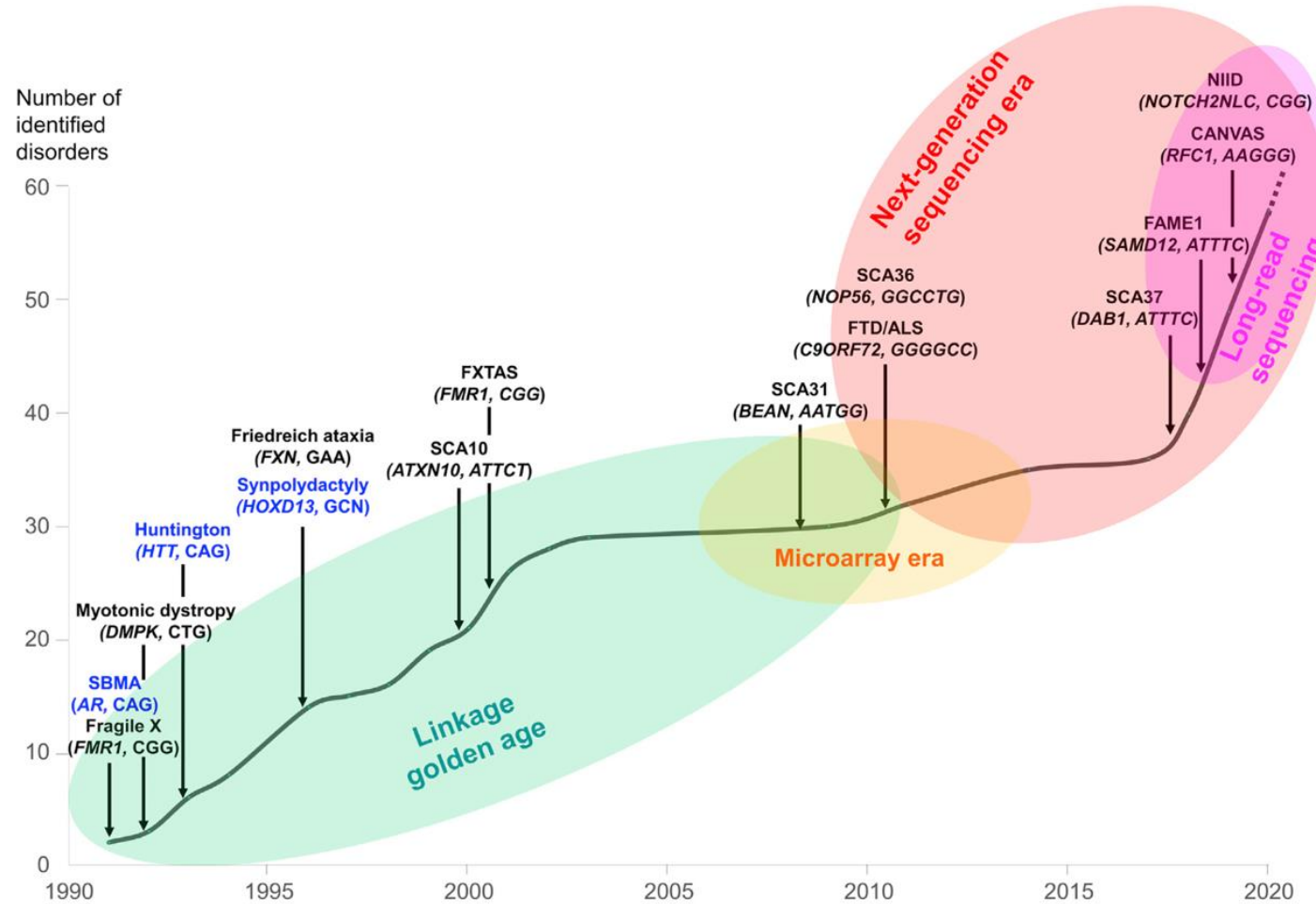


Figure 1. Timeline of repeat expansion discovery in human disorders

Depienne C, Mandel JL. 30 years of repeat expansion disorders: What have we learned and what are the remaining challenges?. *Am J Hum Genet.* 2021;108(5):764-785. doi:10.1016/j.ajhg.2021.03.011

I. Maladies à expansions – problématique



- Nombreuses techniques de routine (PCR, RP-PCR, seq, LR-PCR...)



- Recouvrement des tableaux phénotypiques – analyse séquentielle
 - Surtout pour les gènes d'ataxie (16 maladies par expansion)



- Diagnostic compliqué sur des techniques de séquençage short read
 - Expansion hunter → estimation approximative, interruptions, motifs différents..




→ Approche long reads

- Répond aux limitations du short read tout en multiplexant les loci étudiés

BRAIN
ORIGINAL ARTICLE

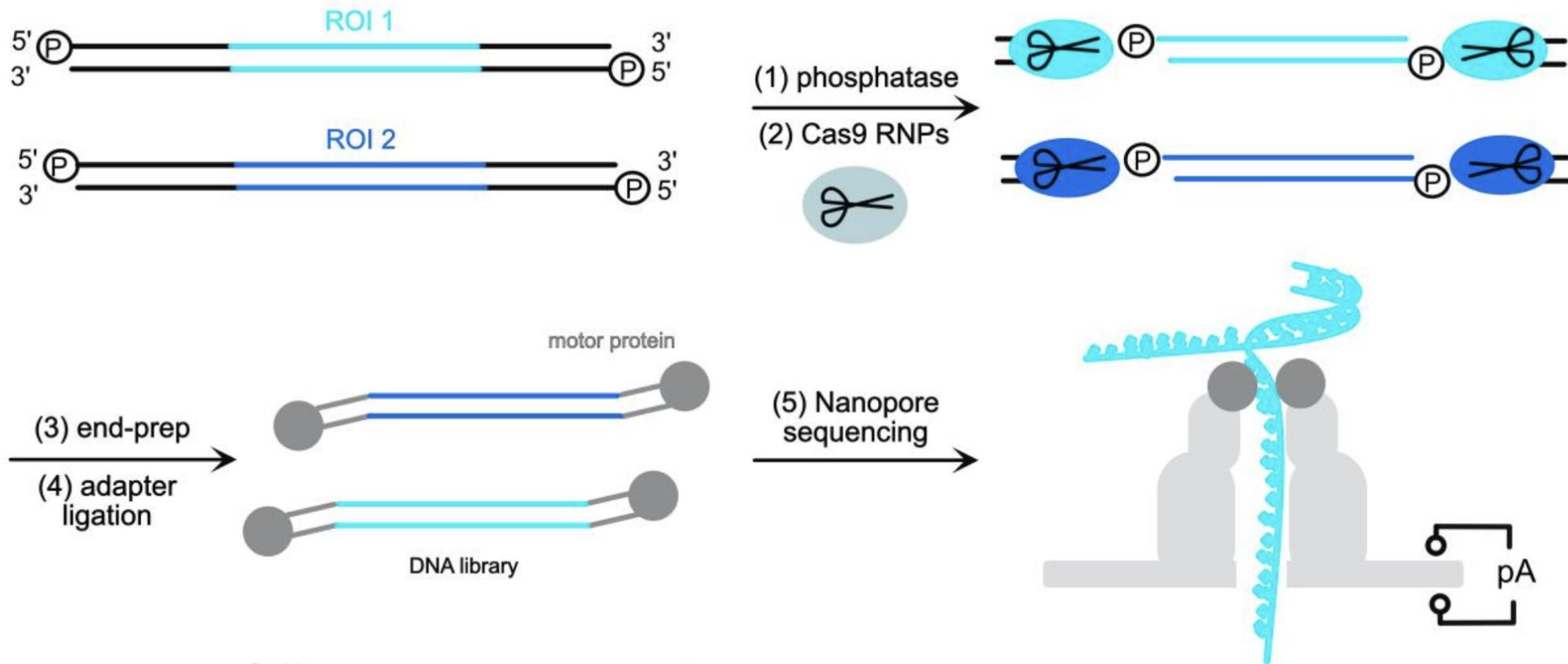


Parallel in-depth analysis of repeat expansions in ataxia patients by long-read sequencing

Hannes Erdmann,^{1,2} Florian Schöberl,³ Mădălina Giurgiu,^{1,4}
Rafaela Magalhaes Leal Silva,¹ Veronika Scholz,¹ Florentine Scharf,¹
Martin Wendlandt,¹ Stephanie Kleinle,¹ Marcus Deschauer,⁵ Georg Nübling,³
Wolfgang Heide,⁶ Sait Seymen Babacan,⁷ Christine Schneider,⁸ Teresa Neuhann,¹
Katrin Hahn,⁹ Benedikt Schoser,² Elke Holinski-Feder,^{1,10} Dieter A. Wolf^{1,11}
and  Angela Abicht^{1,2}

II. Approche long-reads – Crispr/Cas9 enrichment

B



B) Clin-CATS approach for repeat analysis by long-read ONT sequencing and bioinformatics analysis.

II. Approche long-reads – Crispr/Cas9 enrichment

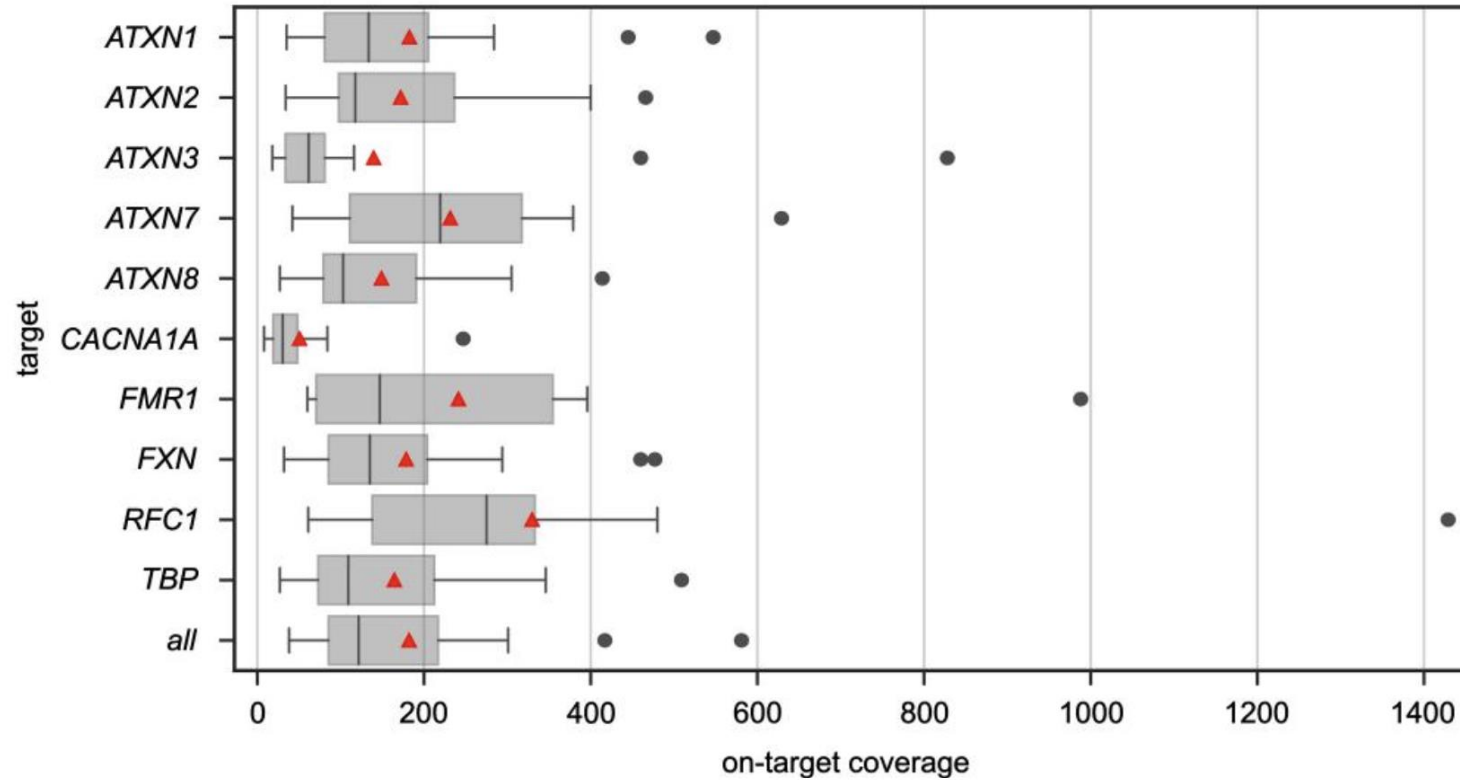


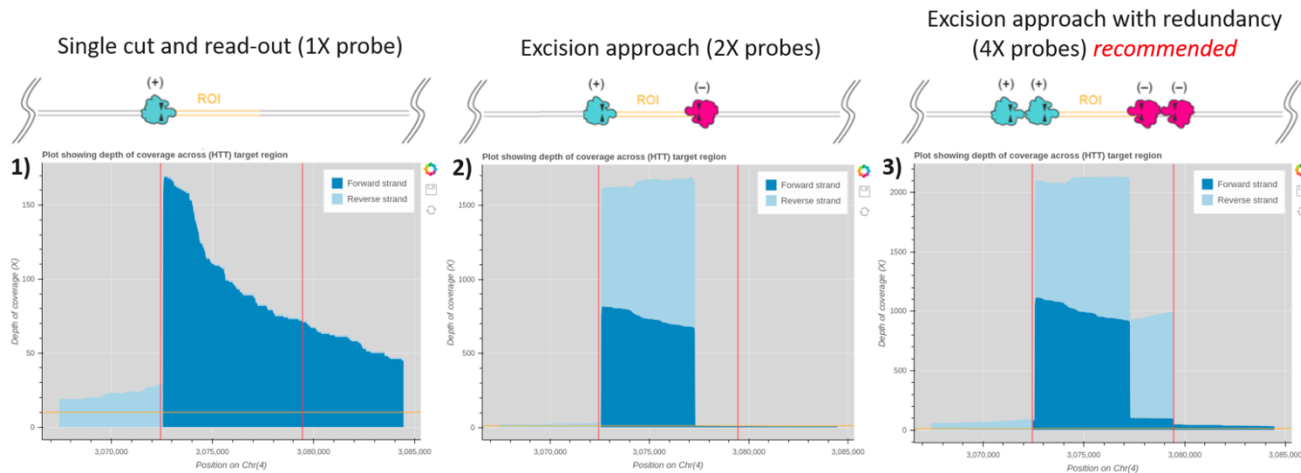
Figure 2 Coverage of the regions of interest. Box plot of the on-target coverage of each locus of interest and overall targets of the 14 samples of the validation cohort.

II. Approche long-reads – Crispr/Cas9 enrichment

- Objectifs :
 - > Ajouter des gènes au panel
 - > (*FGF14, ATXN10, PPP2R2B, NOP56, DAB1, ..., SCA4*)
 - > Multiplexer les échantillons de patients pour les passer sur une même flowcell
 - > Test du programme NanoRepeat pour la lecture des expansions
 - > <https://github.com/WGLab/NanoRepeat>

II. Approche long-reads – Crispr/Cas9 enrichment

Design de guides avec les outils CHOP-CHOP, CRISPOR et IDT



Locus	CRISPR Cas9						
	Guide	Brin	Sens	Sequence 5'→3'	On target Score	Off target Score	Taille site
ATXN1	ATXN1_D1	-	fw	GGTGAAACTTTTATCGTT	49	77	7394bp
	ATXN1_U2	-	rev_c	ATGTAATCGATCTAAGAACC	60	72	
ATXN2	ATXN2_D1	-	fw	GTCGGCTCTGTCTACCGA	34	86	2921bp
	ATXN2_U2	-	rev_c	CCGGTCACCCGGCTCAAGC	30	94	
ATXN3	ATXN3_D1	-	fw	AGCGCATTCCCAATAGACG	53	90	12389bp
	ATXN3_U1	-	rev	GATTACTGCTGAACGCACAT	66	84	
CACNA1A	CACNA1A_D2	-	fw	GGTCCAGTTCTGCGTGAAT	37	77	8407bp
	CACNA1A_U2	-	rev_c	TTGGCACTCGGGCATAGACT	51	83	
ATXN7	ATXN7_D1	+	rev_c	ATCTAGGTTAAACTCCCGC	44	87	1743bp
	ATXN7_U2	+	fw	CGTACTTCGTCGACACC	67	87	
ATXN805/ATXN8	ATXN8_D1	+	rev_c	CATTACAGGTCAGCAAAGA	71	75	8834bp
	ATXN8_U1	+	fw	ATACTGGCCATCGTAATTG	44	83	
TBP	TBP_D1	+	rev_c	GAGGTTACTACTGCATGTTG	68	58	11192bp
	TBP_U1	+	fw	TGAGACGAGTCCAGCGCAA	60	87	
FXN	FXN_D1	+	rev_c	CACCAGTTTCGAGAATCCTG	52	72	9695bp
	FXN_U2	+	fw	CTGTGTAAACCCATACCGG	58	82	
FMR1	FMR1_D2	+	rev_c	ATCAGATCCCAATCTTCTC	72	78	6493bp
	FMR1_U2	+	fw	TTAGGCTTGAGCAACGAAC	67	74	
RFC1	RFC1_D1	-	fw	TTCGTGGAACATCTTGGTA	63	60	8436bp
	RFC1_D2	-	fw	TGATTACAACCATCAAGGAT	100	41	
	RFC1_U1	-	rev_c	TAACTCCAACAACCTCAAC	46	63	
	RFC1_U2	-	rev_c	GCTCAGTCGTTTTAACCCAGG	49	69	
FGF14	FGF14_D1	-	fw	CCGGTAGACCATATTACACA	70	83	14491bp
	FGF14_D2	-	fw	GGTCCTATGGCATTGTGACT	69	67	
	FGF14_U1	-	rev_c	ACTCTCCTAGGCTCCAATAT	82	71	
	FGF14_U2	-	rev_c	CACTCATTGGTTATCTGGAC	76	70	
ATXN10	ATXN10_D1	+	rev_c	TCACTCTATAGAGGTCACCT	80	62	6563bp
	ATXN10_U1	+	fw	TAGTGTGTGTAATAGAGCCC	65	61	
PPP2R2B	PPP2R2B_D1	-	fw	GAGATCTCATGAAGTACCCG	65	89	5877bp
	PPP2R2B_U1	-	rev_c	AGAGGTTGCGCCGCTCGCAA	57	90	
NOP56	NOP56_D1	+	rev_c	CGATGAGACGTGCACCTACC	74	90	4093bp
	NOP56_U2	+	fw	CCAAGCGGAGGAGCAACCG	89	75	
DAB1	DAB1_D1	-	fw	CAACAACGTGACTTCGTAGC	57	93	4246bp
	DAB1_U1	-	rev_c	CGTGACTACTCACTCTACC	61	74	

II. Approche long-reads – Crispr/Cas9 enrichment

Test 1 → CRISPR/Cas9 20231013

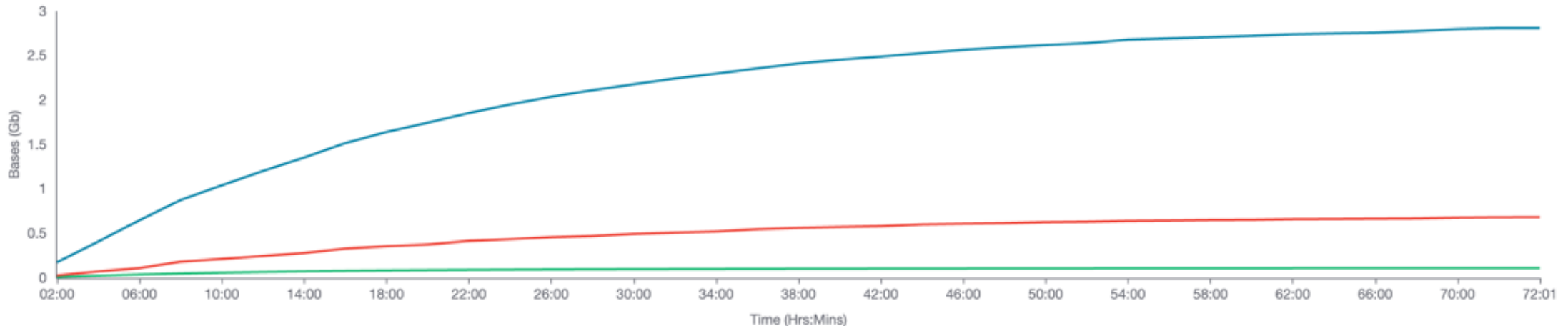
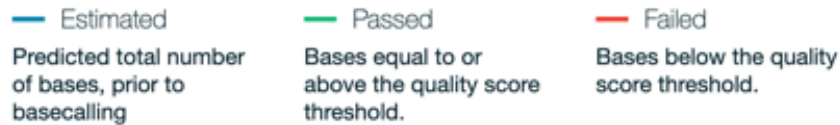
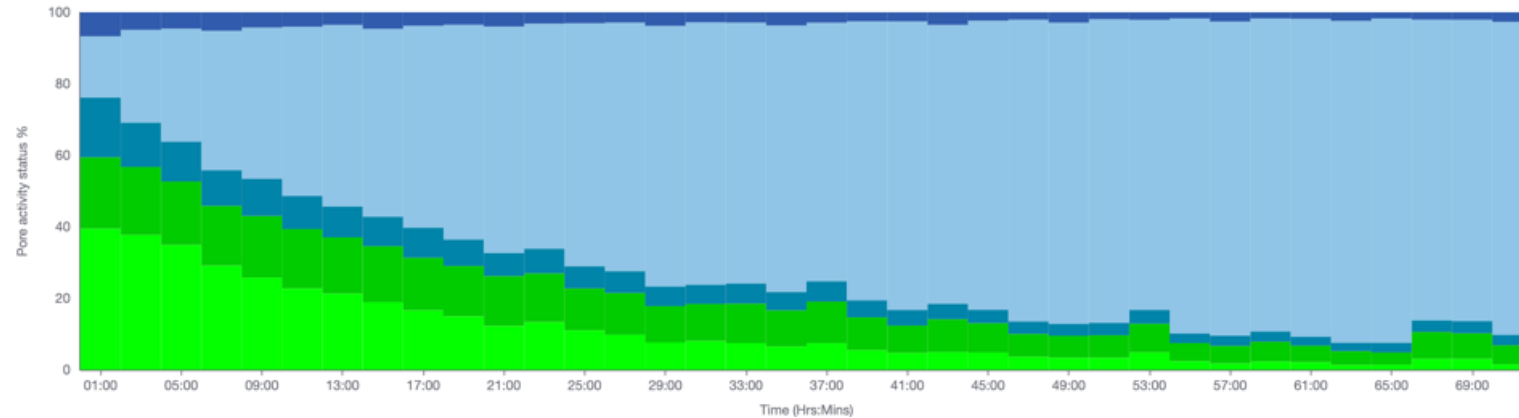
- 1 patient choisit pour son génotype
- 1 flowcell
- 1 librairie
- 72h de run

SCA1 → 29/29
SCA2 → 22/22
SCA3 → 26/68
SCA6 → 11/12
FDRA → 9/E (1200)
CANVAS → N3/N3 (10)

Locus	CRISPR Cas9 20231013		
	≈ X	Rep exp IGV	Rep exp NanoRepeat
<i>ATXN1</i>	58	29_29	29
<i>ATXN2</i>	68	22_22	22
<i>ATXN3</i>	32	26_68	26_68
<i>CACNA1A</i>	26	11_12	11
<i>ATXN7</i>	690	10_10	10
<i>ATXN8OS/ATXN8</i>	33	20_800	20_672/789
<i>TBP</i>	26	40_40	39
<i>FXN</i>	23	9_≈1000	9_146(1)_1034(1)
<i>FMR1</i>	61	24_24	24
<i>RFC1</i>	264	10_10	10
<i>FGF14</i>	43	9_17?	9_16

II. Approche long-reads – Crispr/Cas9 enrichment

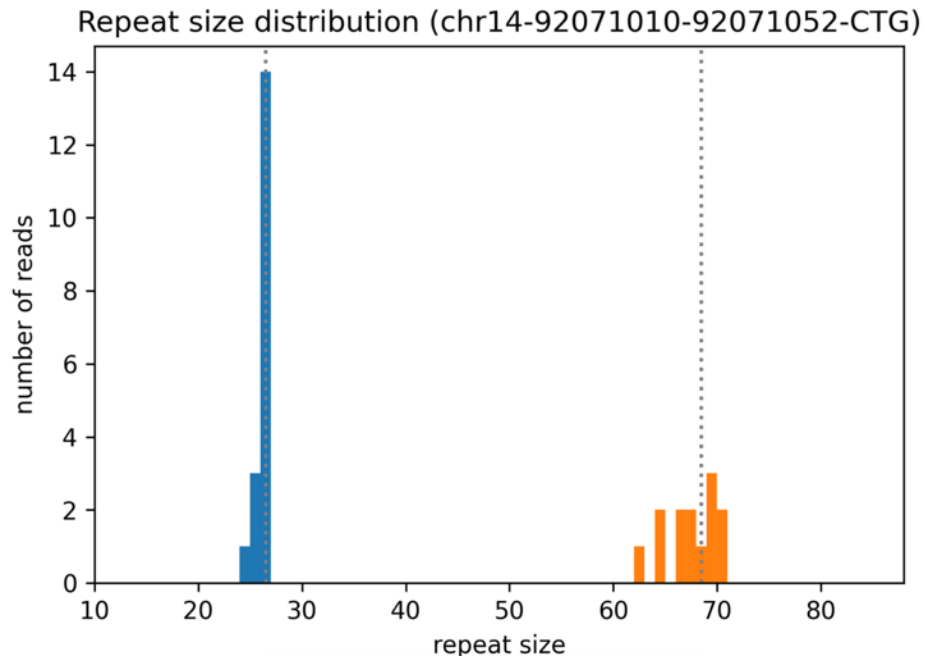
Test 1 → CRISPR/Cas9
20231013



II. Approche long-reads – Crispr/Cas9 enrichment

Test 1 → CRISPR/Cas9
20231013

ATXN3



chr6	16327635	16327722	CTG
chr12	111598950	111599019	CTG
chr14	92071010	92071052	CTG
chr19	13207858	13207897	CTG
chr3	63912685	63912715	CAG
chr3	63912714	63912727	CCG
chr13	70139353	70139384	CTA
chr13	70139383	70139428	CTG
chr6	170561907	170562021	CAG
chr9	69037285	69037304	GAA
chrX	147912050	147912110	CGG
chr4	39348424	39348479	AAAAG
chr13	102161576	102161726	GAA
chr22	45795354	45795424	ATTCT
chr5	146878728	146878758	CAG
chr20	2652733	2652757	GGCCTG
chr1	57367043	57367118	ATTTT

NanoRepeat - <https://github.com/WGLab/NanoRepeat>

II. Approche long-reads - Adaptive sampling

Test 2 → AS 1
20231113

- Arrêt de la commercialisation du kit Cas9 Nanopore
- Approche plus flexible / région plus large; SNP, indel..
- Premier test sur le même patient
- 3 librairies → wash&reload
- 1 librairie → recovery&reload

	AS 1 20231113		
Locus	≈ X	Rep exp IGV	Rep exp NanoRepeat
<i>ATXN1</i>	24	29_29	29
<i>ATXN2</i>	30	22_22	22
<i>ATXN3</i>	23	26_68	26_69
<i>CACNA1A</i>	12	11_12	11
<i>ATXN7</i>	20	10_10	10
<i>ATXN8OS/ATXN8</i>	28	20_800	20_726
<i>TBP</i>	23	40_40	39
<i>FXN</i>	20	9_≈1000	9_760(1)_1145(1)
<i>FMR1</i>	9	24_24	24
<i>RFC1</i>	25	10_10	10
<i>FGF14</i>	32	9_17?	9_17

II. Approche long-reads - Adaptive sampling

Test 2 → AS 1
20231113

Bases

Legend

— Estimated

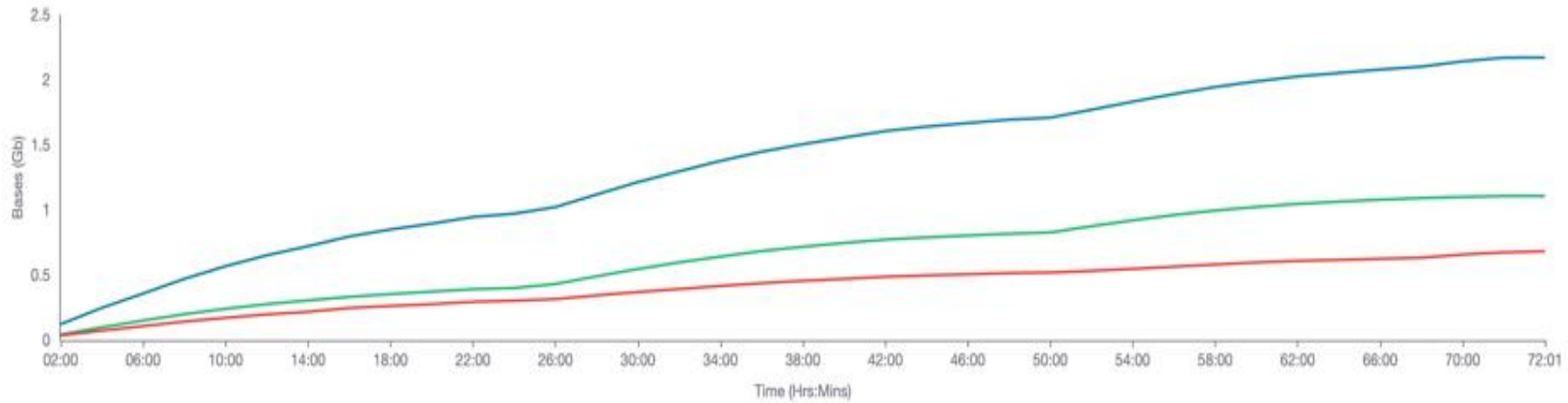
Predicted total number of bases, prior to basecalling

— Passed

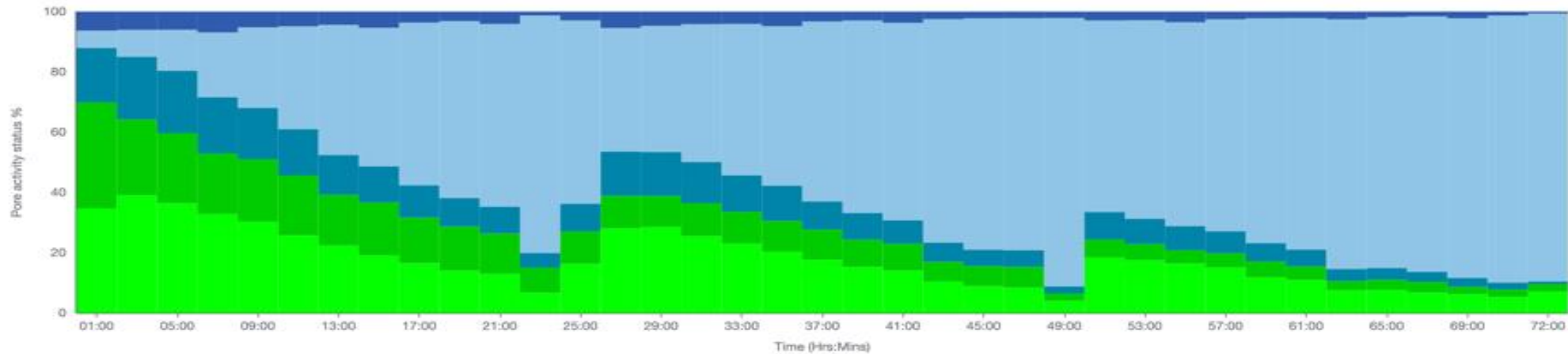
Bases equal to or above the quality score threshold.

— Failed

Bases below the quality score threshold.

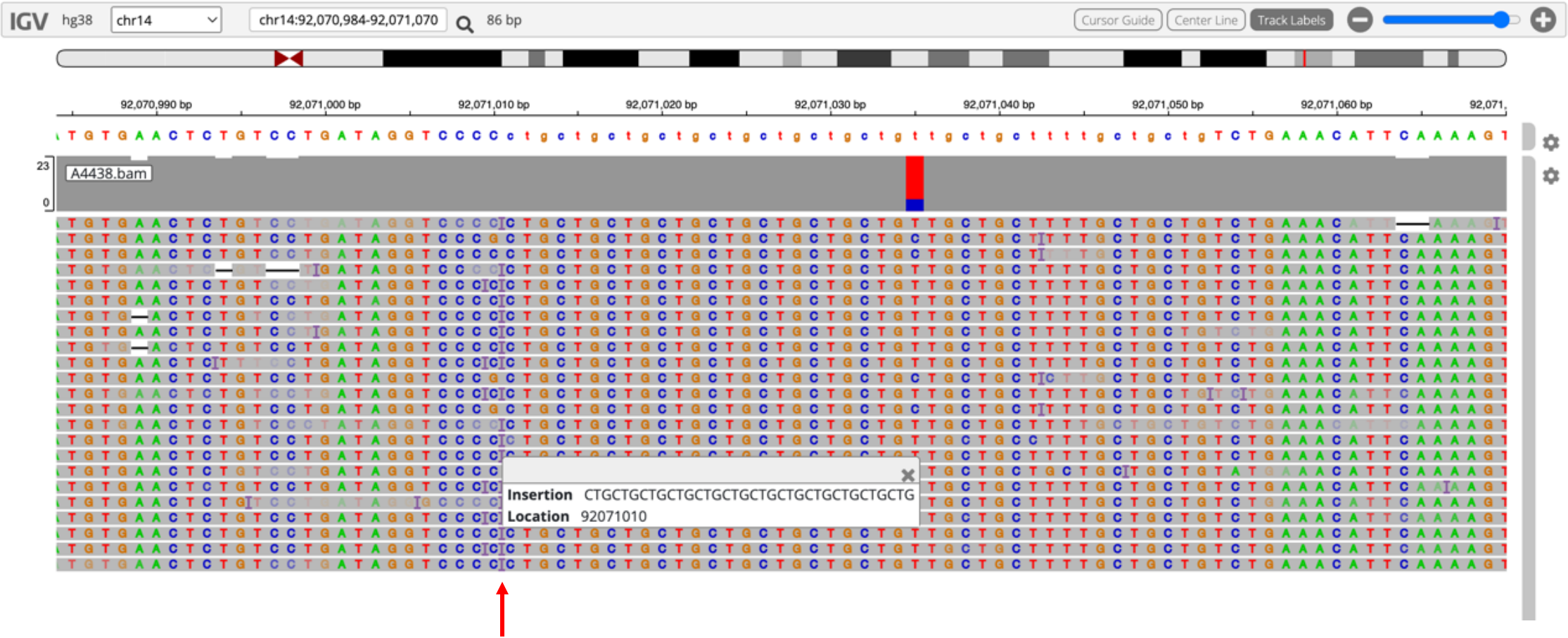


sequencing sequencing unavailable for sequencing for further sequencing



II. Approche long-reads - Adaptive sampling

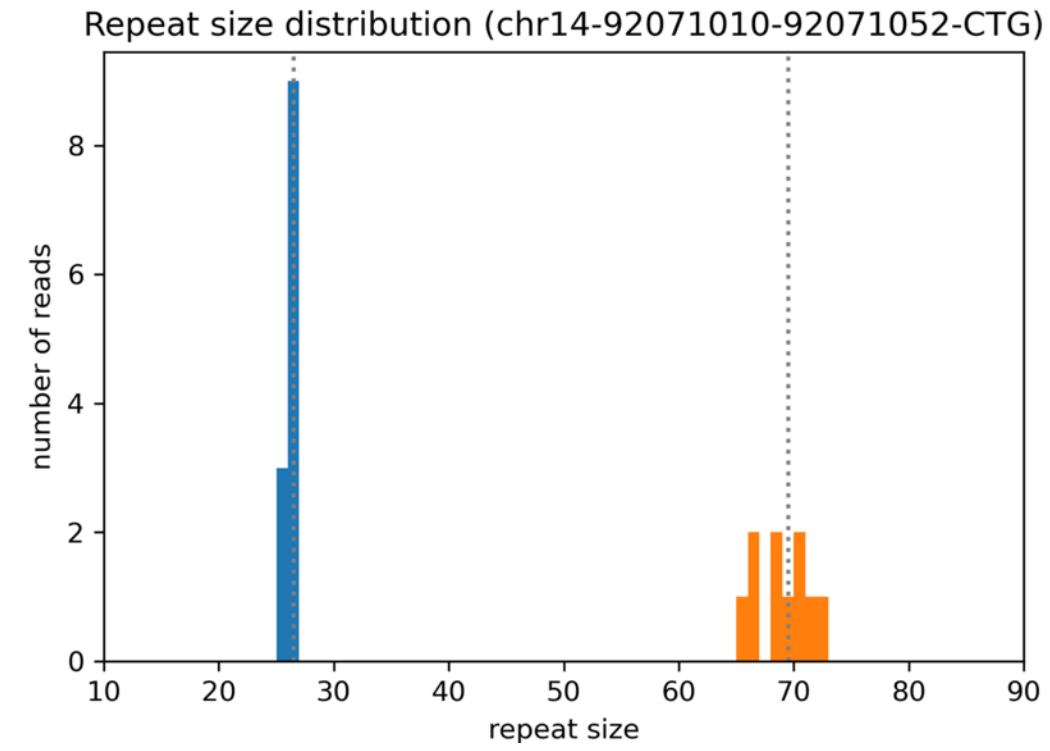
Test 2 → AS 1
20231113



II. Approche long-reads - Adaptive sampling

Test 2 → AS 1
20231113

ATXN3



II. Approche long-reads - Adaptive sampling

Test 2 → AS 1 20231113 / AS 2 20231121 / AS 3 20231130

Locus	AS 1 20231113			AS 2 20231121			AS 3 20231130		
	≈ X	Rep exp IGV	Rep exp NanoRepeat	≈ X	Rep exp IGV	Rep exp NanoRepeat	≈ X	Rep exp IGV	Rep exp NanoRepeat
ATXN1	24	29_29	29	18	29_29	29	17	29_29	29
ATXN2	30	22_22	22	13	22_22	22	10	22_22	22
ATXN3	23	26_68	26_69	19	26_68	26_68(2)_70(3)	9	26_68	26(1)_68(8)
CACNA1A	12	11_12	11	14	11_12	12	13	11_12	11
ATXN7	20	10_10	10	15	10_10	10	9	10_10	10
ATXN8OS/ATXN8	28	20_800	20_726	17	20_800	20_148(3)_743(3)	13	20_800	20(4)_774(5)
TBP	23	40_40	39	22	39_40	39	16	?	39
FXN	20	9_≈1000	9_760(1)_1145(1)	17	9_1100	10(5)_672(1)_1125(5)	9	8_1000	8(4)_88(1)_742(1)_1004(2)
FMR1	9	24_24	24	9	24_24	24	11	24_24	24
RFC1	25	10_10	10	22	10_10	10	16	10_10	10
FGF14	32	9_17?	9_17	15	9_17	9_17	6	9_17?	9(3)_15(2)
ATXN10							10	15_15	15
PPP2R2B							10	10_10	10
NOP56							10	3_3	3
DAB1							14	13_14?	0

Locus	CRISPR Cas9 20231013		
	≈ X	Rep exp IGV	Rep exp NanoRepeat
ATXN1	58	29_29	29
ATXN2	68	22_22	22
ATXN3	32	26_68	26_68
CACNA1A	26	11_12	11
ATXN7	690	10_10	10
ATXN8OS/ATXN8	33	20_800	20_672/789
TBP	26	40_40	39
FXN	23	9_≈1000	9_146(1)_1034(1)
FMR1	61	24_24	24
RFC1	264	10_10	10
FGF14	43	9_17?	9_16

SCA1 → 29/29
SCA2 → 22/22
SCA3 → 26/68
SCA6 → 11/12
FDRA → 9/E (1200)
CANVAS → N3/N3 (10)

III. Multiplexage

- Multiplexage avec un kit de barcoding (native-barcode)
 - x Echec pour l'approche AS
- Décision de repasser sur Cas9, en multiplexant 3 patients
 - ✓ Fusion de deux protocoles

III. Multiplexage

- Multiplexage avec NativeBarcoding sur Cas9 enrichment
 - x Très faible enrichissement
- Nouvelle stratégie :
 - Chargement séquentiel, runs différents sur une flowcell (lavages)
 - En modulant les temps de run selon le taux de décroissance des pores

III. Multiplexage

Test 3 → CRISPR/Cas9 20240328



III. Multiplexage

Test 3 → CRISPR/Cas9 20240328

Patient 1

Patient 2

Patient 3

Locus	CRISPR Cas9 20240328			CRISPR Cas9 20240328			CRISPR Cas9 20240328		
	≈ X	Rep exp IGV	Rep exp NanoRepeat	≈ X	Rep exp IGV	Rep exp NanoRepeat	≈ X	Rep exp IGV	Rep exp NanoRepeat
ATXN1	20	29_29	29	17	29_30	29	79	30_30	30
ATXN2	18	22_22	22	11	22_22	22	62	23_23	23
ATXN3	12	26_68	26_69	18	23_24	23	25	23_30	23_30
CACNA1A	2	11_?	11	4	7_11	7_11	12	13_13	13
ATXN7	89	10_10	10	75	10_13	10_13 (histo)	270	10_13	10_13 (histo)
ATXN8OS/ATXN8	3	20_800	20	5	9_12	8_13	33	10_20	13_20
TBP	1	39_?		3	37	36	15	37	37
FXN	3	9_1000	10_959	5	8_9	9	46	9_9	10
FMR1	17	24_24	24	31	30_30	31	103	23_23	23
RFC1	36	10_10	10	12	AAGGG (~920 rep) / AAAGG	833	98	AAAAG (156 rep) / AAGGG (1000 rep)	105
FGF14	7	9_17	9(1_15(4))	9	9_180	9_193	25	8_40	8_42

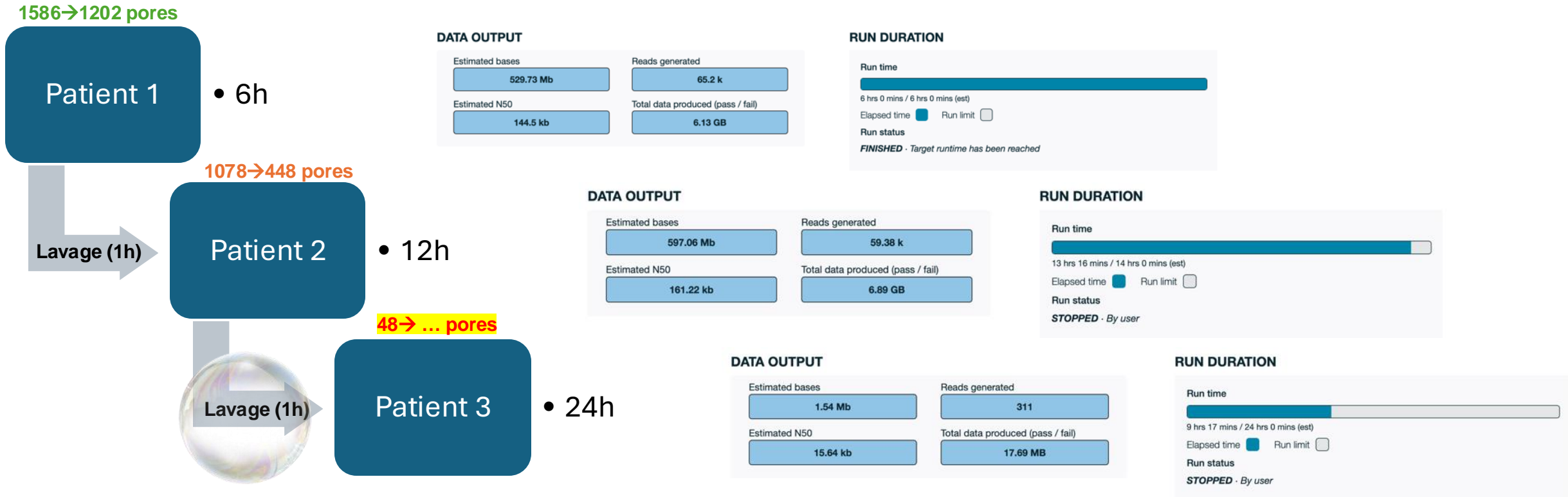
SCA1 → 29/29
 SCA2 → 22/22
 SCA3 → 26/68
 SCA6 → 11/12
 FDRA → 9/E (1200)
 CANVAS → N3/N3 (10)

FDRA → 9/9
 CANVAS → E/N2 (10)

SCA1 → 30/30
 SCA2 → 23/23
 SCA3 → 23/30
 SCA6 → 13/13
 FDRA → 9/9
 CANVAS → E/N3

III. Multiplexage

Test 4 → CRISPR/Cas9 20240409



III. Multiplexage

Test 4 → CRISPR/Cas9 20240409

Patient 1

Patient 2

Patient 3

Locus	CRISPR Cas9 20240409			CRISPR Cas9 20240409			CRISPR Cas9 20240409		
	≈ X	Rep exp IGV	Rep exp NanoRepeat	≈ X	Rep exp IGV	Rep exp NanoRepeat	≈ X	Rep exp IGV	Rep exp NanoRepeat
ATXN1	57	29_29	29	69	29_30	29			
ATXN2	90	22_22	22	82	22_22	22			
ATXN3	19	26_68	26_68	27	23_24	23			
CACNA1A	10	11_12	12	15	7_11	7_11			
ATXN7	334	10_10	10	182	10_13	10_13(histo)			
ATXN8OS/ATXN8	17	20_800	20_773	20	9_12	12_723			
TBP	15	40_40	39	13	37	36			
FXN	18	9_1100	9_991	41	9_9	10			
FMR1	63	24_24	24	156	30_30	30			
RFC1	157	10_10	10	47	AAGGG (~920 rep) / AAAGG	?			
FGF14	45	9_17	9_17	41	9_180	9_198			

SCA1 → 29/29
SCA2 → 22/22
SCA3 → 26/68
SCA6 → 11/12
FXN → 9/E (1200)
RFC1 → N3/N3 (10)

FXN → 9/9
RFC1 → E/N2 (10)

SCA1 → 30/30
SCA2 → 23/23
SCA3 → 23/30
SCA6 → 13/13
FXN → 9/9
RFC1 → E/N3

IV. Conclusion - Perspectives

- Optimiser les temps de runs / la performance du lavage
- 3 patients sur une flowcell de MinION semble jouable
- L'outil nanorepeat semble performant si la profondeur est suffisante
- Ce serait idéal sur une flowcell plus performante (P2solo ?)
 - Qualité des flowcells assez aléatoire – difficile d'anticiper
- Le protocole Cas9 a été mis à jour dans nanopore community et des réactifs de remplacement de NEB sont proposés



Merci pour votre attention

Des questions ?