Automatic treatment of sequences from haploid/diploid DNA, two new tools : SeqQual and Polymorfind

SeqQual:

→Focus on **base quality** for **producing sequence data alignments for population genetic analyses**

Polymorfind:

→ Focus on **SNPs** detection and **indels** in heterozygous species

Présentation par Michaël Mozar et Alix Pernet

UMR 1259 Génétique et Horticulture
Centre INRA d'Angers-Nantes

SeqQual: an automatic pipeline **integrating quality** for identifying SNPs & producing **sequence data files for population genetic analyses**

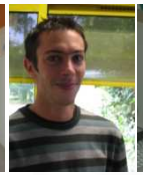
Programmation & expertise
Perl, linux & interfacing with R



Tiange Lang, Jean-Marc Frigerio, Alain Franc

Beta- test & validation





François Hubert, Pierre Abadie, Thibaut Decourcelle, Josquin Tibbits Camille Lepoittevin, Jorge Paiva, El Mujtar Veronica

Sylvain Gaillard expertise Polyscan

Antoine Kremer funding & post-doc recruitment

Coordination, conception et validation du pipeline

Pauline Garnier-Géré

Polymorfind: an automatic pipeline for detecting SNPs and **indels** for heterozygous species

Programmation & expertise
Perl, linux

Sylvain Gaillard, Michaël Mozar

Beta- test & validation

Dereeper Alexis , Nicolas Stéphane

Projects and validation

Fabrice Foucher, Alix Pernet

Collaboration for developing complementary and non-redundant tools

Bioinformatic Tools for Sequence/SNP data

Phred – Base calling and quality determination →

Phred (quality) Score

"A score of 20 corresponds to a error rate of approximately 1 in 100 bases, a score of 30 to 1 in 1000 bases"

Phrap – Sequence Assembly

Polybayes – SNP detection for haploid

Polyphred – SNP detection for haploid and diploid heterozygote nucleotides

Polyscan – SNP detection for haploid, diploid heterozygotes nucleotides and indels

→ Black boxes, large complex outputs, not user-friendly.

Previous SNP analysis pipelines: examples

2004: Loïck Le Dantec *et al*. → Use of phrap and polybayes, **do not consider Phred quality score** *per se*.

2006: Nathalie Pavy *et al*. →**Partial use of polybayes score** only which is an overall probability for SNP detection.

2008: Jifeng Tang *et al*. → **Consider quality score, but only for the whole SNP site**, not for every single nucleotide.

…. And many mores??

→ PROBLEMS with previous available pipelines and motivation for new tools

- Lack of alignment files produced automatically, which would integrate quality check for all nucleotides in the alignment, and also heterozygote IUPAC codes → Good SNP site detected but nucleotides with bad quality score can still be at same position

- → Even bigger problem when aiming at population genetic analyses, not just SNP detection!

- No automatic tools for detecting both SNPs and heterozygote indels in highly polymorphic species, using diploid DNA (ex: polyploid species)

much time spent examining alignments by eyes (CodonCode Aligner / Consed soft. / Genalys)!!!

→ PROBLEMS with previous available pipelines and motivation for new tools

- **Lack of alignment files produced automatically**, which would integrate **quality check** for all nucleotides in the alignment, and also **heterozygote IUPAC codes** → Good SNP site detected but nucleotides with bad quality score can still be at same position

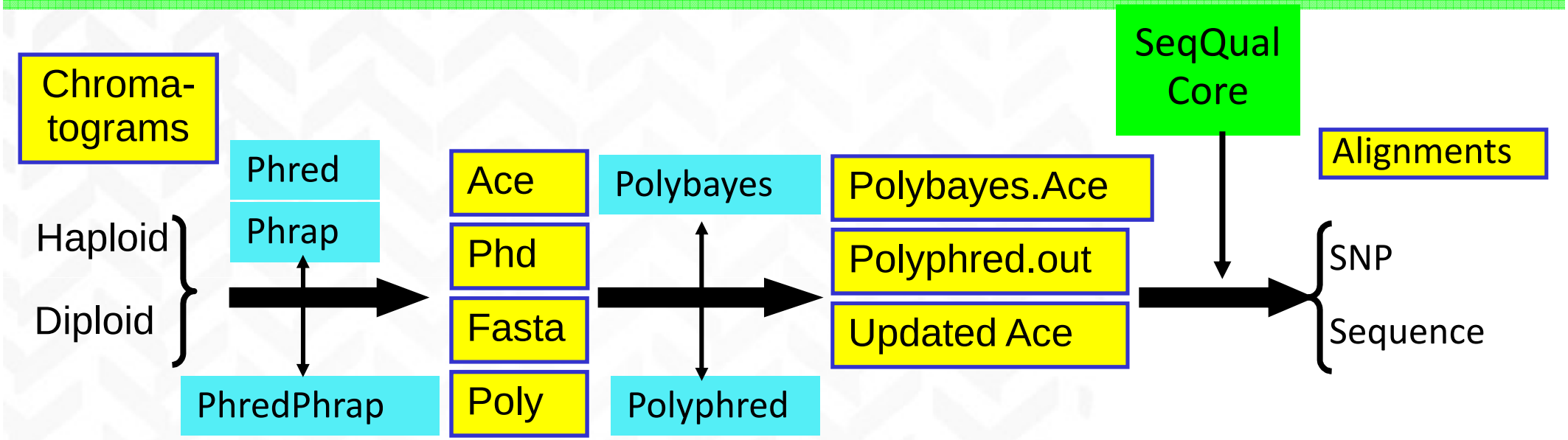
- → Even bigger problem when aiming at **population genetic analyses**, not just SNP detection!

SeqQual

- No automatic tools for **detecting both SNPs and heterozygote indels in highly polymorphic species**, using diploid DNA (ex: polyploid species)

Polymorfind

SeqQual pipeline structure



SeqQual features summary

- automatic launch of all programs for multiple fragments
- ***Given good alignments,*** this allows to get fasta alignments in one click, then to have a quick overview of the sequence data obtained and identify most problems (paralog amplification, bad quality, polymorphisms including indels…) →bad quality bases replaced with « ??? »
- add missing base / truncate ends (user parameters)
- Code heterozygotes & check neighbourhood quality (2N data)
- Optimise alignments/quality integration after quality check
- # file format outputs

SeqQual output treatment

« Given good alignments », what does it mean?

Phrap will align sequences given the default parameters, and the user can play with them to improve the alignments

OUTPUT FILES

Fasta and SNP alignments (IUPAC codes for heterozygotes) - # options

Unaligned fasta → go to your favourite alignment software

Arlequin input files (haplotypic or genotypic format for haplotype reconstruction (phase unknown)) – **possible from your own fasta alignments**

→ use of automatic routine on multiple files allows to get illumina input format for testing them for arrays from fasta outputs

SeqQual interface (in development)

RGui - [R Console]

Fichier Edition Voir Misc Packages Fenêtres Aide

```
> source("SeqQual.interface.r")
> source("SeqQual.interface.r")
[1] File diploid-ab1.txt already
[1] "File  diploid-ab1.txt writte
[1] ----------------------------
[1] perl ~/program/checkdir_Outpu
[1] mkdir Output
[1] cd Output
[1] mkdir aln
[1] mkdir SNP
[1] mkdir unaln
[1] mkdir arlequin
[1] cd ..
[1]
[1] ## diploid
[1] perl ~/program/print_source-
[1] source source-aln.txt
[1]
[1] ## Truncate missing
[1] perl ~/program/print_source-
[1] source source-truncate.txt
[1]
[1] ## Isolated nucleotides
[1] perl ~/program/print_source-
[1] source source-replace.txt
[1]
[1] ## Remove missing / insertio
[1] perl ~/program/print_source-
[1] source source-remove.txt
[1]
[1] perl ~/program/print_source-
[1] source source-take_aln.txt
[1]
[1] ## Arlequin input files
[1] perl ~/program/print_source-arlequin-diploid.pl  myFile  > source-write_arlequin.txt
[1] source source-write_arlequin.txt
[1] perl ~/program/print_source-take_arp-diploid.pl  myFile  > source-take_arp_diploid.txt
[1] source source-take_arp
[1] ==========
```

SeqQual

Input Panel

File name : myFile

Fasta files	○		
Ace files	○	own phd	☑
Chromatograms Haploids	○	own phd	☑
Chromatograms Diploids	●	own phd	☑

Phrap parameters

default_qual	15	max_subclone_size	5000
trim_start	0	trim_score	20
force_level	0	trim_penalty	-2
bypass_level	1	trim_qual	13
maxgap	30	confirm_length	8
repeat_stringency	0.7	confirm_trim	1
node_seg	8	confirm_penalty	-5
node_space	4	confirm_score	30
qual_show	20	indexwordsize	10

Output Panel

SNP output file	yes ●	no ○	if yes: Polybayes posterior score	99
Unaligned fasta file	yes ●	no ○		
Haplotype data (phase unknwon)	yes ●	no ○		
Arlequin input file	yes ●	no ○		
-- id -- with cluster	yes ●	no ○		

If output towards Arlequin with cluster, enter the cluster names :

Enter here myPops

Phred / Polyphred parameters

Polymorphism score	60
Trim_quality_score	20
Phred_quality_score	30
Heterozygote_score.tcl	90

SeqQual parameters / Options

Truncate missing	yes ●	no ○	if yes:	3
Remove missing	yes ●	no ○		
1 - 3 isolated nucleotides	yes ●	no ○	if yes:	3

Launch SeqQual ...

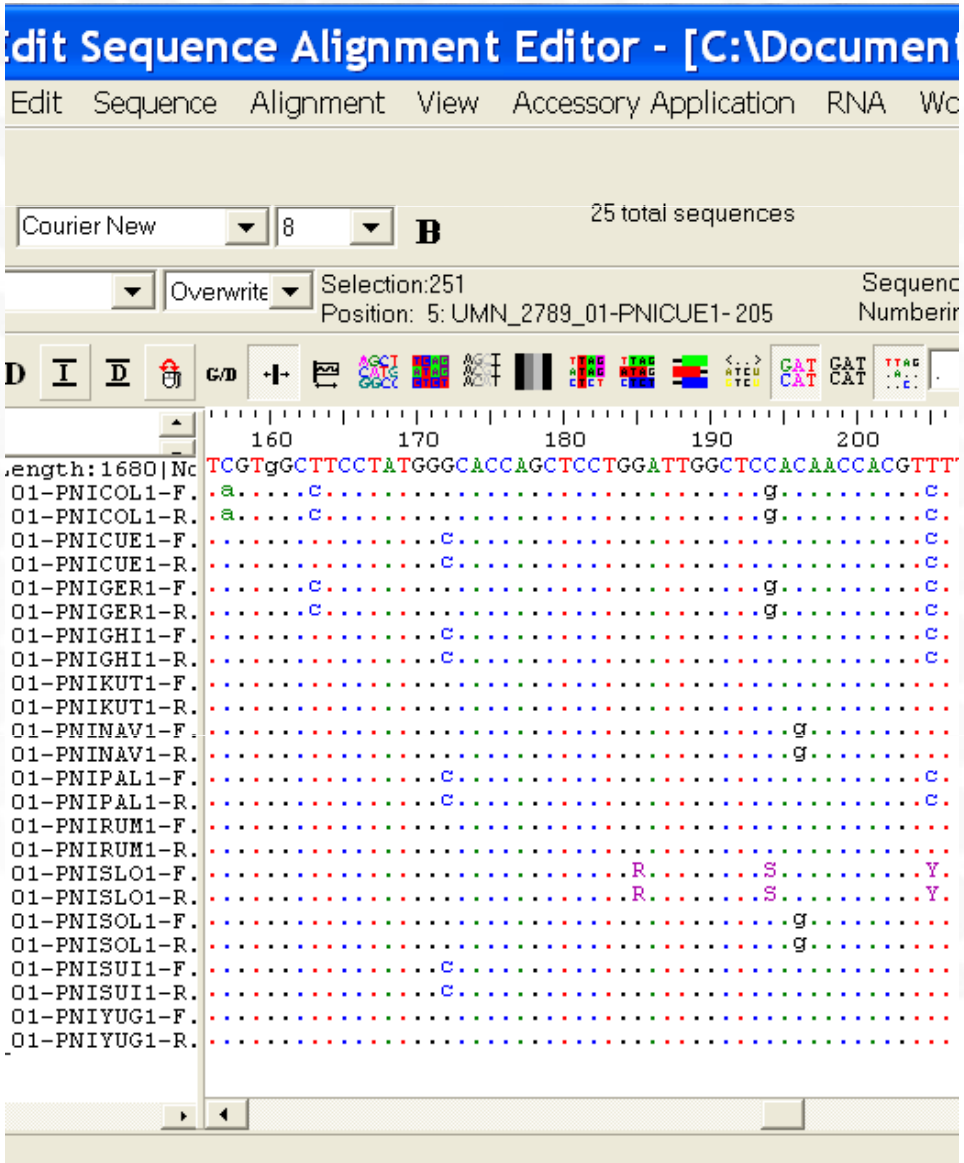
File name: diploid-ab1.txt

Press to launch SeqQual

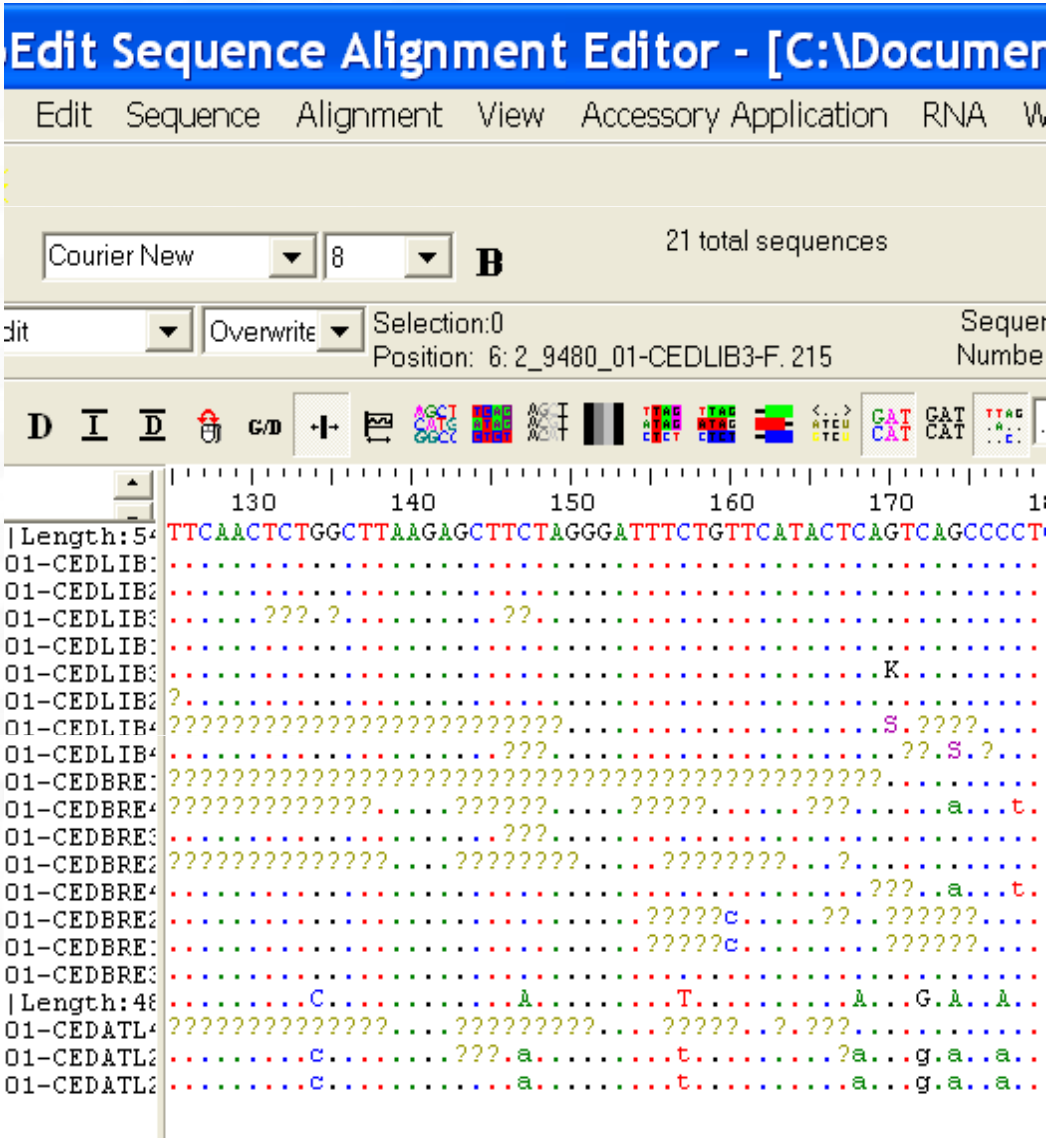
R program → produces the shell file to launch the program with chosen parameters by the user

Examples of SeqQual alignment outputs

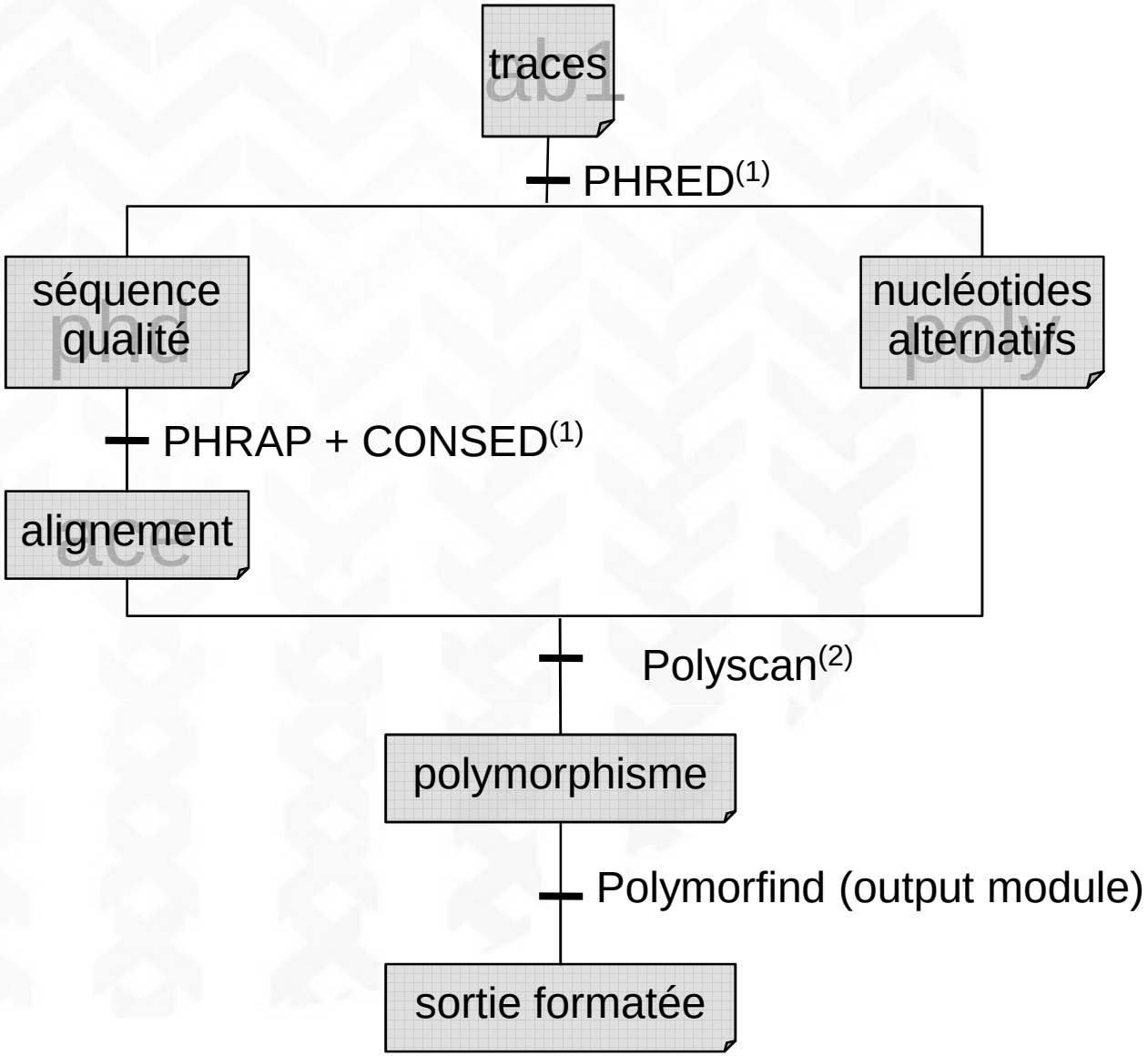
Good quality sequences



Sequences with bad quality bases



Polymorfind : cœur du traitement



Détection des SNP

- Utilisation de 2 passes de Polyscan
  - 1ère : forte stringence élevée
    - bonne qualité de séquence (quality = 30 / 30)
    - score de détection élevé (gtscore = 70 / 99)

    → sélection des positions où le polymorphisme est sûr
  - 2nde : stringence faible
    - qualité de séquence moyenne (quality = 20 / 30)
    - score de détection moyen (gtscore = 40 / 99)

    → détection de tout le polymorphisme aux positions précédemment sélectionnées

Output of SNPs analysis in Polymorfind

homoSNP

heteroSNP

	A	B			E		G
1	SNP						
2							
3	REFERENCE	92	116	122	149	158	177
4	REFERENCE	CC	CC	CC	AA	GG	CC
5							
6	GA20ox__117_RosaSNP-test2_SP6_c1_O17_065.ab1	--	**	**	GG	CC	**
7	GA20ox__119_RosaSNP-test2_SP6_c1_G21_089.ab1	--	TT	--	GG	CC	--
8	GA20ox__AB_RosaSNP-test2_SP6_c1_A13_063.ab1	--	-T(79)	--	-G(51)	C-(59)	--
9	GA20ox__AR_RosaSNP-test2_SP6_c1_A17_079.ab1	--	TT	--	GG	CC	--
10	GA20ox__AS_RosaSNP-test2_SP6_c1_A15_064.ab1	--	TT	--	-G(47)	CC(43)	--
11	GA20ox__BA_RosaSNP-test2_SP6_c1_A21_095.ab1	--	TT	--	GG	CC	--
12	GA20ox__BB_RosaSNP-test2_SP6_c1_K13_053.ab1	--	-T(63)	--	-G(88)	C-(79)	--
13	GA20ox__BE_RosaSNP-test2_SP6_c1_E21_091.ab1	--	TT(53)	--	-G(65)	C-(66)	--
14	GA20ox__BF_RosaSNP-test2_SP6_c1_K21_085.ab1	**	**	**	GG	**	**
15	GA20ox__BJ_RosaSNP-test2_SP6_c1_K17_069.ab1	--	-T(95)	--	-G(55)	C-(56)	--
16	GA20ox__BR_RosaSNP-test2_SP6_c1_C15_062.ab1	--	TT	TT	--	--	--
17	GA20ox__CA_RosaSNP-test2_SP6_c1_C19_078.ab1	GG	--	--	--	--	GG
18	GA20ox__CE_RosaSNP-test2_SP6_c1_O15_050.ab1	--	-T(96)	--	-G(57)	C-(60)	--
19	GA20ox__CS_RosaSNP-test2_SP6_c1_I23_088.ab1	--	-T(67)	--	-G(85)	C-(91)	--
20	GA20ox__DA_RosaSNP-test2_SP6_c1_M21_083.ab1	--	-T(58)	--	--(54)	--	**
21	GA20ox__FC_RosaSNP-test2_SP6_c1_I15_056.ab1	--	TT(61)	--	-G(53)	C-(52)	--
22	GA20ox__FE_RosaSNP-test2_SP6_c1_E13_059.ab1	--	--	--	--	--	--
23	GA20ox__FO_RosaSNP-test2_SP6_c1_E17_075.ab1	--	--	--(78)	--	--	**
24	GA20ox__FP_RosaSNP-test2_SP6_c1_A19_080.ab1	--	TT	--	GG	CC	--
25	GA20ox__GA_RosaSNP-test2_SP6_c1_K19_070.ab1	--	-T(96)	--	-G(63)	C-(63)	--
26	GA20ox__GB_RosaSNP-test2_SP6_c1_O21_081.ab1	--	-T(71)	--	-G(90)	C-(94)	--
27	GA20ox__GG_RosaSNP-test2_SP6_c1_G17_073.ab1	--	TT(47)	--	-G(67)	C-(65)	--
28	GA20ox__H190_RosaSNP-test2_SP6_c1_O19_066.ab1	--	TT	--	GG	CC	--
29	GA20ox__IN_RosaSNP-test2_SP6_c1_G15_058.ab1	--	-T(89)	--	-G(67)	C-(66)	--
30	GA20ox__JU_RosaSNP-test2_SP6_c1_G19_074.ab1	--	-T(70)	--	-G(49)	C-(44)	--
31	GA20ox__LA_RosaSNP-test2_SP6_c1_G23_090.ab1	--	TT(53)	--	-G(59)	C-(53)	--
32	GA20ox__LWP_RosaSNP-test2_SP6_c1_A23_096.ab1	--	TT	--	GG	CC	--
33	GA20ox__MO_RosaSNP-test2_SP6_c1_I13_055.ab1	--	TT	--	GG	CC	--

polymorfind

Output of indels in Polymorfind

37								
38	REFERENCE	546	681	702	746	748	1116	
39								
40	2005-06-17_G01_ELF-8_H190_7_T7_013.ab1	A(homoIns: 56)						
41	2005-06-17_G02_ELF-8_R.w_10_T7_014.ab1	A(homoIns: 61)						
42	FeliciteetPerpetue_ELF8_B10_078.ab1							
43	JubileLoubert_ELF8_B09_077.ab1							
44	LittleWhitePet_ELF8_C10_076.ab1							
45	SandersWhite_ELF8_D10_074.ab1							
46	SemisNepal2_ELF8_E12_088.ab1							
47	ThaliaLoubert_ELF8_E09_071.ab1							
48	TheFairy_ELF8_E10_072.ab1							
49	Webbiana_ELF8_D07_057.ab1		A(homoDel: 51)		TC(homoDel: 59)			
50	arvensis_ELF8_E08_056.ab1							
51	cadic1_ELF8_E11_087.ab1	A(homoIns: 27)	A(homoDel: 62)	TTTGGCTTGAA(homoDel: 62)		GC(homoDel: 62)	T(homoDel: 26)	
52	hirtula_ELF8_G09_067.ab1		A(homoDel: 60)		TC(homoDel: 62)			
53	macounii_ELF8_H09_065.ab1		A(homoDel: 60)		TC(homoDel: 60)			
54	maximowicziana_ELF8_F07_053.ab1							
55	moschataUmbrella_ELF8_F08_054.ab1							
56	multibracteata_ELF8_C08_060.ab1		A(homoDel: 54)		TC(homoDel: 50)			
57	multiflora_ELF8_D11_089.ab1							
58	pablito_ELF8_A11_095.ab1							
59	pendulina_ELF8_F10_070.ab1		A(homoDel: 60)		TC(homoDel: 58)			
60	pteragonis_ELF8_A12_096.ab1							
61	roxburghii_ELF8_A10_080.ab1		A(homoDel: 60)		TC(homoDel: 59)			
62	rugosaSchua_ELF8_B07_061.ab1			TTTGGCTTGAA(homoDel: 62)		GC(homoDel: 62)		
63	rugosaThunb_ELF8_B08_062.ab1			TTTGGCTTGAA(homoDel: 62)		GC(homoDel: 62)		
64	rugosaTroll_ELF8_C07_059.ab1			TTTGGCTTGAA(homoDel: 60)		GC(homoDel: 62)		
65	rugosa_ELF8_A07_063.ab1			TTTGGCTTGAA(homoDel: 62)		GC(homoDel: 62)		
66	setigera_ELF8_D08_058.ab1							
67	ussuriensis_ELF8_C11_091.ab1		A(homoDel: 58)		TC(homoDel: 60)			
68	wichu_ELF8_E07_055.ab1							
69								

And a file with Fasta alignments (Indels) (IUPAC codes for heterozygotes)

SeqQual: ex of validation on 150 < 2000 fragments in *Q. petraea/robur*

Data from the oak resequencing project (funded by EVOLTREE network of excellence – Coord. P. Garnier-Géré /C. Plomion)

DATA: Diploid DNA amplification and sequencing- **4 reads per fragment**

Approach followed

1) All automatic analysis of Fasta outputs from SeqQual (phd score 40, **diploid routine**, **default parameters**)
2) initial screen for at least 1 read with 50 % good quality + visual checks of fasta
3) Visual examination of Chromatograms in CodonCode aligner
4) Record of true SNPs based on visual checks, false positives and negatives, paralog amplification patterns

SeqQual: ex of validation on 150 < 2000 fragments in *Q. petraea/robur*

Data from the oak resequencing project (funded by EVOLTREE network of excellence – Coord. P. Garnier-Géré /C. Plomion)

DATA: Diploid DNA amplification and sequencing- **4 reads per fragment**

Results 150 fragments	NB of reads with 75% phd score quality (max=4)	NB of reads with 50% phd score quality (max=4)	paralog pattern detected < fasta	NB of heteroz. Indels < fasta check	False++	False--	Paralog pattern confirmed?
11	0	1	1 case	2	0	0	1 case
3	0 to 2	2 to 3	yes?	na	0	1 case	no
20	0 to 2	2	no	6	0	2 cases	no
8	0 to 3	2 to 4	yes	na	na	na	yes
27	0 to 3	3	no	4	0	2 cases	no
3	1 to 3	2 to 4	no	4+2 polyA	2 in each (assembly pb)	0	no
70	0 to 4	4	no	4 cases	0	2 cases	no
9	0 to 4	2 to 4	no	6	1 in each	0	no

→If no heterozygote indels or polyA (**97** fragments), **0 false positives** for higher quality reads , 1 case for lower quality (3 fragments), **183 SNPs detected**

→If heterozygote indels, you get a higher rate of false ++ (**7 cases for 26 heterozygotes indels**), which are eliminated by checking the alignments outputs

SeqQual: ex of validation on 150 < 2000 fragments in *Q. petraea/robur*

Data from the oak resequencing project (funded by EVOLTREE network of excellence – Coord. P. Garnier-Géré /C. Plomion)

DATA: Diploid DNA amplification and sequencing- **4 reads per fragment**

Results

150 fragments

Paralog patterns only detected by visual examination of fasta, → now automatically detected with criteria accounting for bother average and maximum lengths of DNA stretch with good quality data

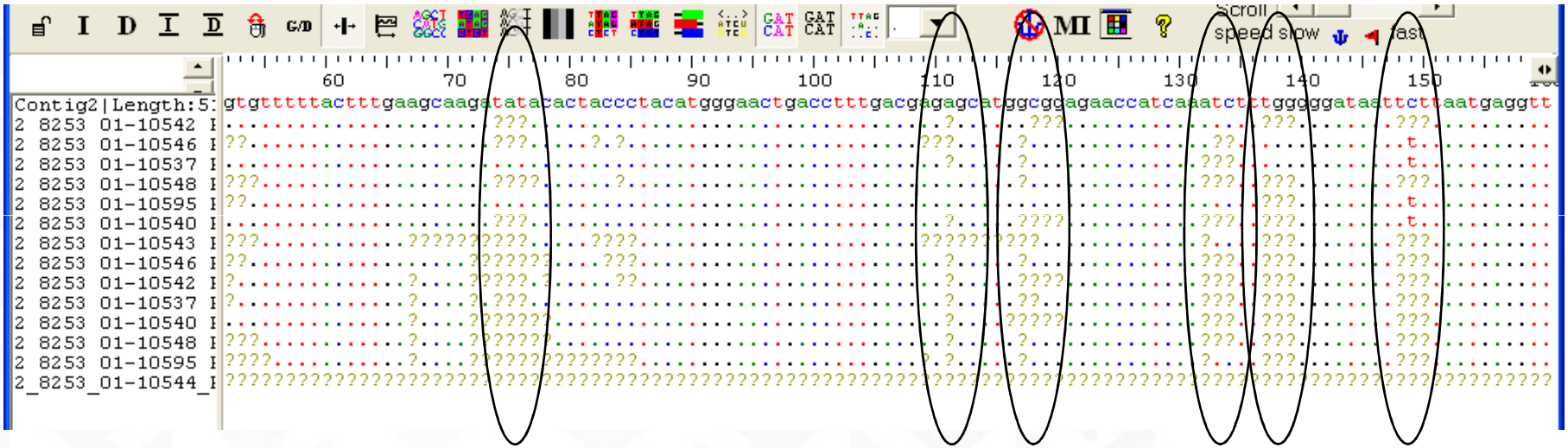
NB with scor (m						Paralog pattern confirmed?
11						1 case
3						no
20						no
8						yes
27						no
3						no
70	0 to 4	4	no	4 cases	0	2 cases	no
9	0 to 4	2 to 4	no	6	1 in each	0	no

→If no heterozygote indels or polyA (**97** fragments), **0 false positives** for higher quality reads , 1 case for lower quality (3 fragments), **183 SNPs detected**

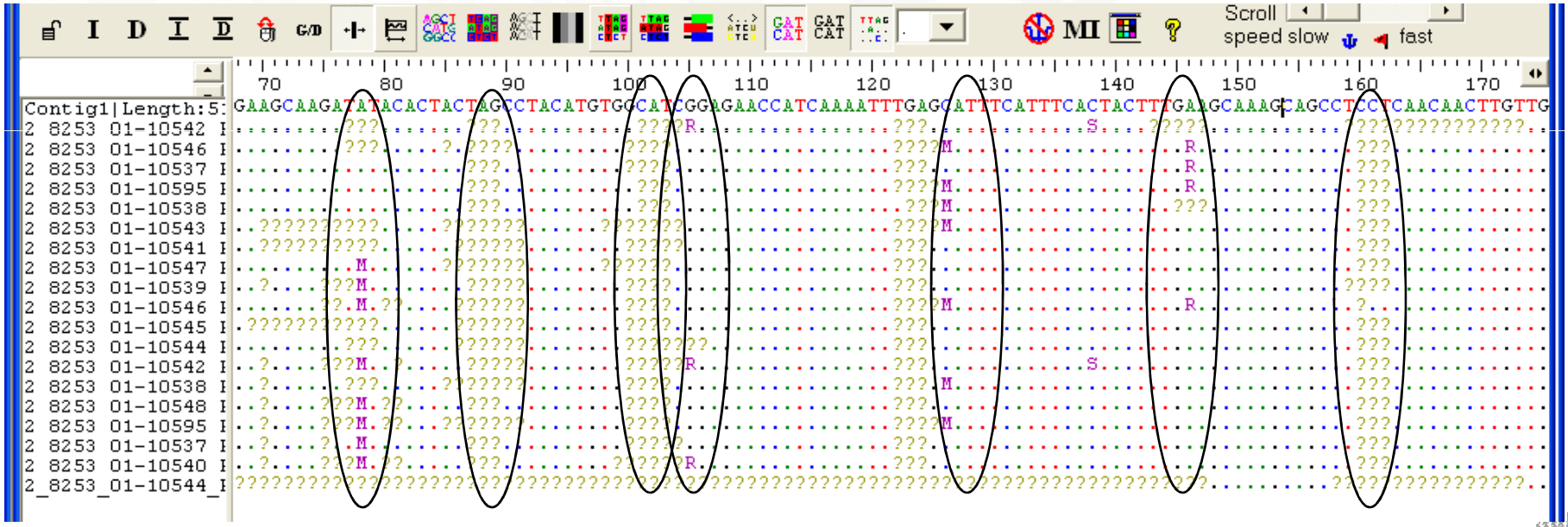
→If heterozygote indels, you get a higher rate of false ++ (**7 cases for 26 heterozygotes indels**), which are eliminated by checking the alignments outputs

Paralog patterns, how does it look like?

From the haploid pipeline option → heterozygotes replaced by ???, regions with mostly ??? Separated by good DNA stretches



From the diploid pipeline option → heterozygotes replaced by ??? Or IUPAC codes



SeqQual scope of applications (user point of view)

Within species sequence data:

few alignments problems / contiguing possible (phrap parameters)

Default sets of parameters

Phred score problems → possibility to import own ABI phd files

Case studies ongoing: **very low error rates** (0-5% / « by eye » checking)

Time saving: 30 to 50 less time-consuming! To be ready for data analyses

Problem of **heterozygote indels**: detectable but to be improved…

More divergent sequence data (orthologs among species as in Barcoding data)

run quality check and get unaligned fasta
or work on assembly parameters

From an *.ace assembly file

allows to get the corresponding fasta alignments integrating quality

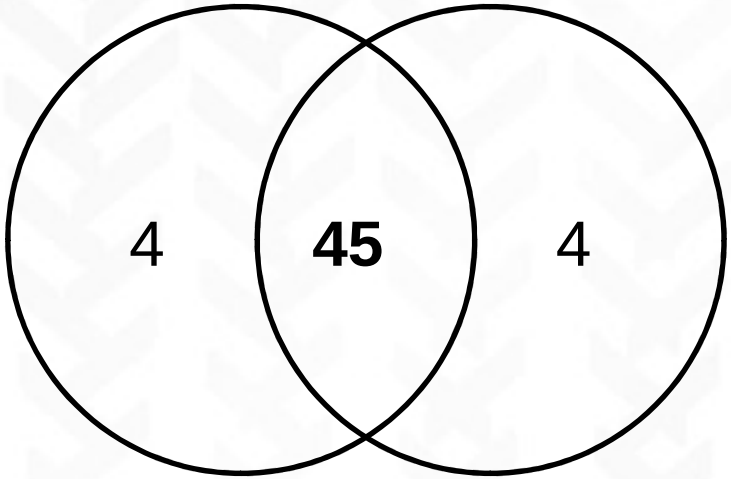
can allow to compute better similarity indices for paralog searches and estimate diversity parameters for confirmed orthologous sequences

Polymorfind : SNP detection

En nombre de positions

Rosier ELF8

29 séquences

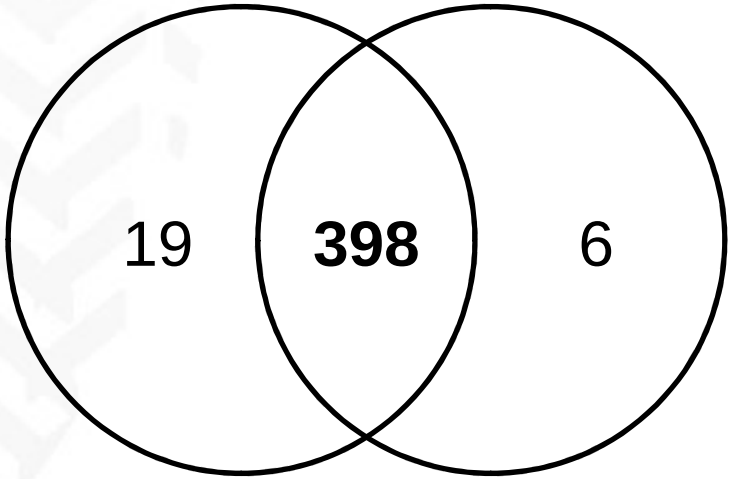


3 SNP récupérés après vérification

En nombre de positions et occurrences confondues

Vigne Wali

146 séquences



Analyse du gène *GA20OX*

Taille de la séquence : 400 pb
Polymorfind : détection de SNP à 13 positions différentes et d'InDel à 1 position
Analyse manuelle : détection de SNP à 1 position supplémentaire

Non détecté par Polymofind

Position SNP	92	116	122	149	158	177	238	246	253	272	304	305	319	358
	Exon						Intron							Ex
Indéterminé	2	3	3	1	2	5	4	2	2	1	2	2	3	
HomoSNP	1	22	1	15	15	1	0	1	1	1	1	0	2	1
heteroSNP	0	14	0	20	20	0	11	0	5	0	0	1	11	4
Synonyme (S) Non Synonyme (NS)	S	S	S	S	S	NS	*	*	*	*	*	*	*	S

Analyse du gène *GA20OX*

Taille de la séquence : 400 pb
Polymorfind : détection de SNP à 13 positions différentes et d'InDel à 1 position
Analyse manuelle : détection de SNP à 1 position supplémentaire

Non détecté par Polymofind

Position SNP	92	116	122	149	158	177	238	246	253	272	304	305	319	358
	Exon						Intron							Ex
Indéterminé	2	3	3	1	2	5	4	2	2	1	2	2	3	
HomoSNP	1	22	1	15	15	1	0	1	1	1	1	0	2	1
heteroSNP	0	14	0	20	20	0	11	0	5	0	0	1	11	4
Synonyme (S) Non Synonyme (NS)	S	S	S	S	S	NS	*	*	*	*	*	*	*	S

De plus un InDel (hetero, T) en position 282, 1 individu

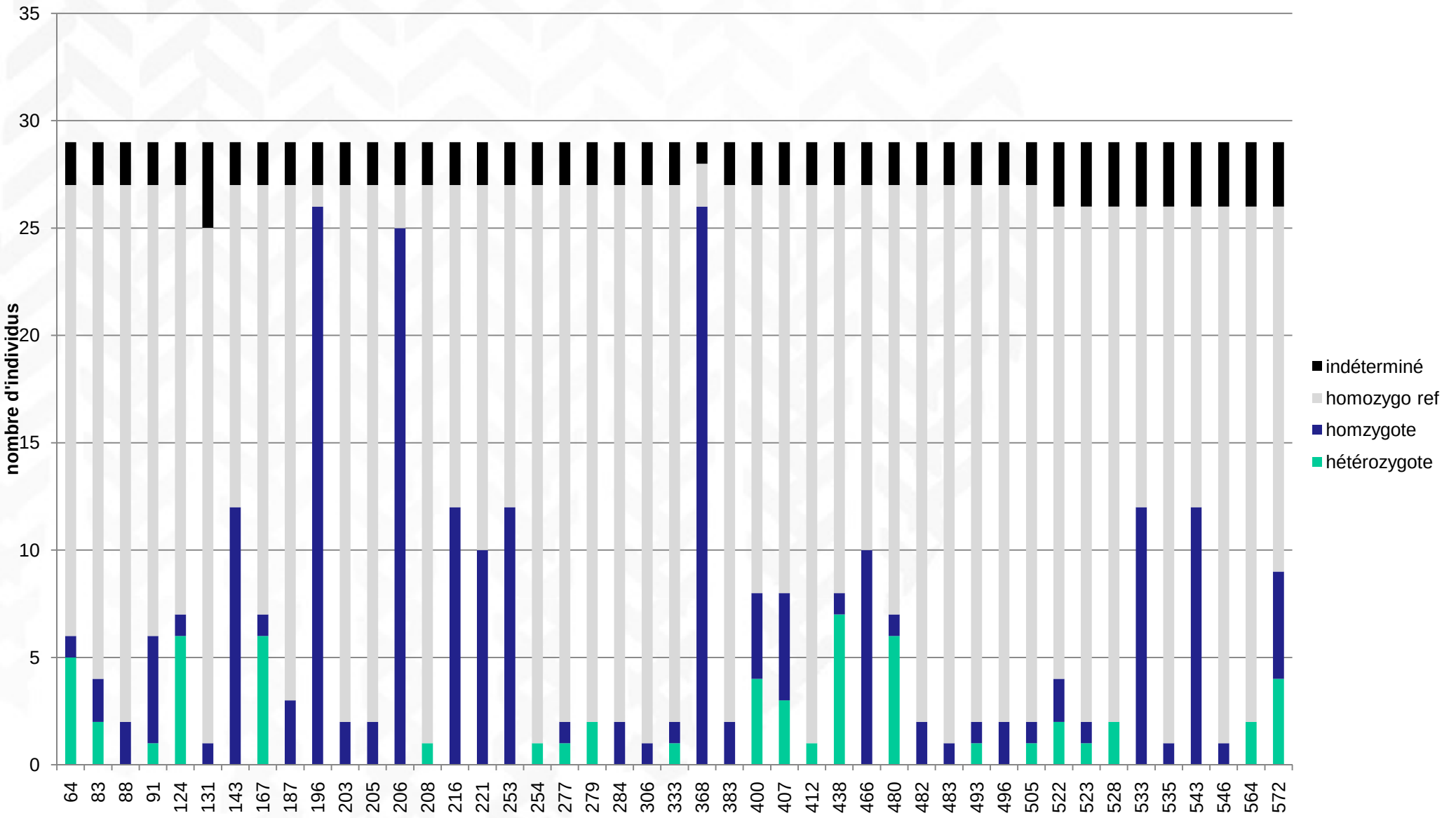
Analyse du gène *GA3OX*

Taille de la séquence : 350 pb
Polymorfind : détection de SNP à 10 positions différentes
Analyse manuelle : détection de SNP à 1 positions supplémentaires

Non détecté par Polymofind

Position SNP	28	78	135	145	174	176	189	210	213	244	303
	Exon										
Indéterminé	11	*	0	0	1	*	2	2	1	1	2
HomoSNP	1	0	0	7	2	1	0	0	2	0	1
heteroSNP	1	1	1	10	1	0	6	2	0	1	0
Synonyme (S) Non Synonyme (NS)	NS	S	S	NS	S	NS	S	NS	S	NS	S

Fréquence des SNP par position pour le gène *ELF8* chez le rosier



Polymorfind : détection des SNP

- Double analyse
  - quasi suppression des faux positifs
  - forte réduction des faux négatifs
    - SNP rares récupérés
- Faible nombre de paramètres utilisés
  - 3 paramètres (quality , gtscore , density of het)
  - 17 paramètres pour Polyscan
- Validé sur plusieurs jeux de données
  - Rosier : INRA Angers
  - Vigne : INRA Montpellier

SeqQual: conclusion & applications

- « huge » amount of time saved! (50 less times to get « perfect » files for population genetic analyses
- Even with only automatic fasta assessment→ **very low rate of false positives, and** even lower rates if examination of SeqQual output alignments (~zero for haploid data)
- Currently used for building 1536plex illumina arrays for maritime pine in EVOLTREE (IA1.2)
- Used also as an necessary step in our assembly strategies for ESTs gene banks from many species (EVOLTREE)

Perspectives/developments

- Integrate/compare with Polyscan instead of polyphred for detecting heterozygotes & Het. indels in diploid sequence data (but inherent limitation to Polyscan (at least 8 reads))
- Adapt the pipeline to input ace-assembly format from sequence data obtained from 454
- Propose more default parameter sets adapted to different case studies

Polymorfind: conclusions

Developpement of an efficient tool for identifying SNPs in PCR products of heterozygous individuals

* Almost no false positives

* Some false negatives (bad quality sequences)

* Very fast to analyse data (1 min versus 1 day)

* Simple use
  - 1 only file with all parameters (default parameters are functional)
  - 1 click = 1 full analysis

* Tool developped in Perl

* Modular code modulaire allowing us to include new tools / fonctionalities

Polymorfind: Perspectives

En cours :

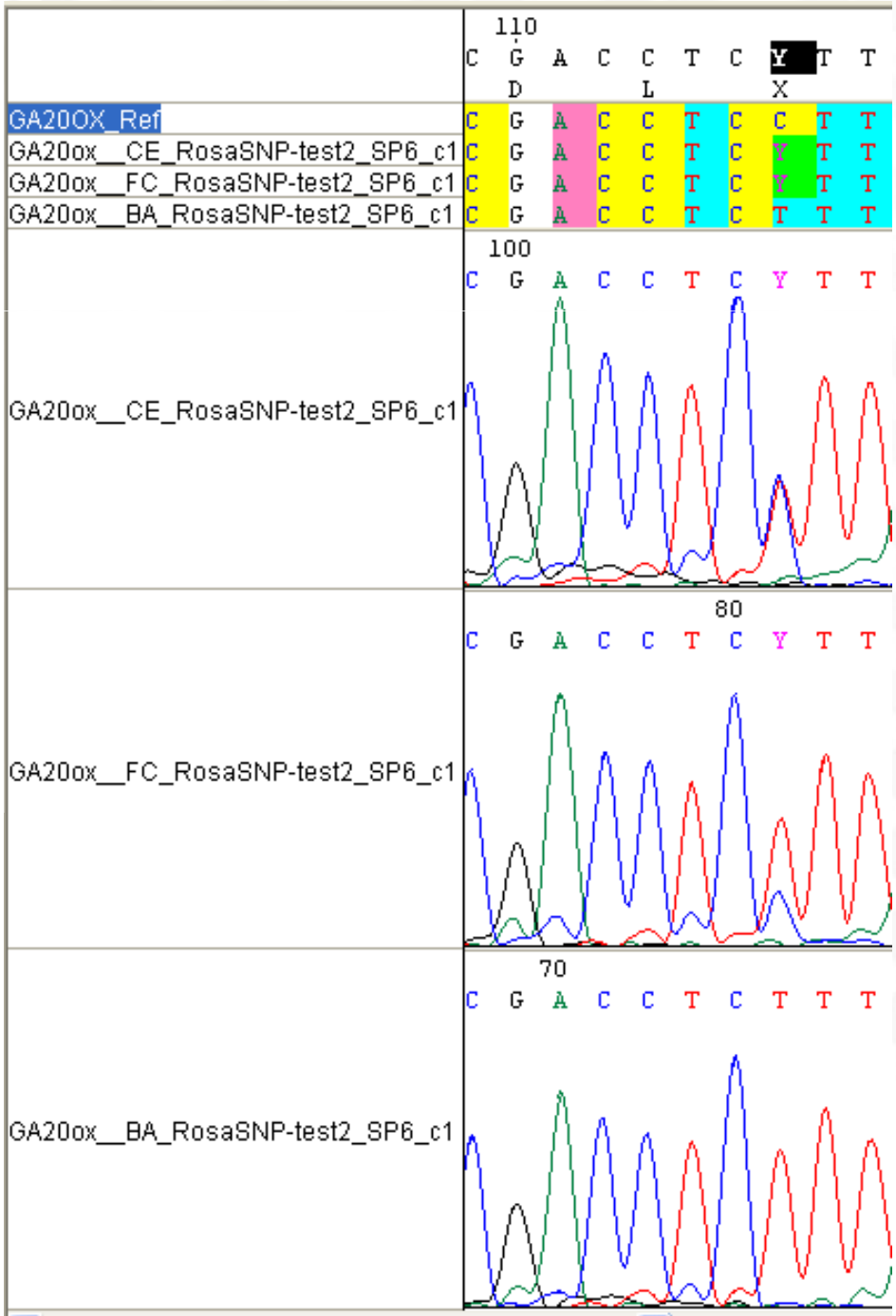
* Indel : nature des indels, dissocier les 2 signaux après un Indel hétérozygote (traitement des zones actuellement inaccessibles)

À venir :

* Polyploïdie : comment prendre en compte dans l'analyse le niveau de ploïdie des individus (SNP à score plus faible, voire non détecté)

* Diversification des sorties :
  - alimentation de bases de données (URGI)
  - post-traitement de l'information par d'autres outils
    - ex : Phase pour l'inférence d'haplotype, DARWin ...

Problèmes des individus polyploïdes (Polymorfind a été défini pour des individus diploïdes)



Cas de 3 individus tétraploïdes en position 116

TC (96)

TT(61) → Fragrant Cloud est un individu tétraploïde

3:1 T/C

TT(100)

Distribution et environnement

- Installation préalable des logiciels gratuits et des librairies Bioperl
- Fonctionne sous environnement UNIX
- Fichier d'aide à l'installation et manuel d'utilisation

SeqQual

Obtention du pipeline (already alpha- and beta- tested):

Demander aux auteurs

•tiange.lang or Pauline Garnier-Géré: pauline@pierroton.inra.fr

•Réception d'info sur les mises-à-jours du pipeline

•Soumission prévue à Bioiformatics

Polymorfind

Obtention du pipeline :

http://genhort.angers.inra.fr/projects/polymorfind

FAQ :

• sylvain.gaillard@angers.inra.fr

Acknowledgements

SeqQual

Test for Arlequin output format:

Florian Alberto

Treesnips EU Project

Digenfor TRILAT project

EVOLTREE network

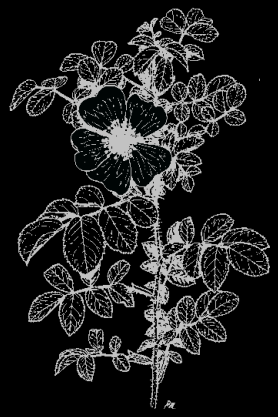
ANR Transbiodiv

UMR BIOGECO:

Christophe Plomion

Rémy Petit

Polymorfind



UMR GenHort

Delerue Thomas, Lucas Virginie, Mozar Michaël

Foucher Fabrice, Gaillard Sylvain, Hibrand-Saint Oyant Laurence, Lalanne David, Pernet Alix

UR EPGV : VaRoFon et RosaSNP

Bérard Aurélie, Brunel Dominique, Chauveau Aurélie, Lepaslier Marie-Christine

INRA Montpellier

Bacilieri Roberto, Dereeper Alexis , Nicolas Stéphane

(1)PHRED-PHRAP-CONSED : Green P, Ewing B, Gordon D. University of Washington.
http://bozeman.mbt.washington.edu

(2) Polyscan : Ken Chen. Washington University in St. Louis.
http://genome.wustl.edu/tools/software/polyscan.cgi


Merci pour votre attention