

Chromosomal copy number variation in *Saccharomyces pastorianus*: evidence for extensive genome dynamics in industrial lager brewing strains.

Van den Broek M¹, Irina Bolat¹, Nijkamp J², Ramos E¹, Luttik MAH¹, Koopman F¹, Pronk JT¹, Geertman JM³, de Ridder D^{2*}, Daran JM¹

Supplemental information

Legend to supplemental figures

Figure S1: Identification of a *Seub*CHRXII DNA non-reciprocal introgression at *Sc*CHRXII in *S. pastorianus* CBS1483 strain. **A-** The graphs represent the ploidy prediction of CBS1483 CHRXII made by Magnolya (Nijkamp et al. 2012). The black box indicates the presence of a contig (Contig00589) aligned on *Seub*CHRXII showing an higher copy number than the rest of the chromosome (three versus two copies). **B-** Annotation of the contig00589, the 6.4-kb fragments harboured three genes YLR411W/*CTR3*, YLR412W/*BER1* and YLR413W/*INA1* encoding a high affinity copper transporter, a protein involved in microtubule-related processes and a putative protein of unknown function respectively. **C-** Localization of contig00589. The 8-kb library reads were mapped to contig00589 using BWA(Li et al. 2010). A total of 8762 reads were aligned, the mate paired reads of the mapped sequences were selected and assembled. Three contigs were assembled. The first contigs was assembled from 3810 reads and matched a scaffold composing *Seub*CHRXII. The two other contigs were assembled from 1119 and 515 reads and showed perfect identity with scaffolds (Scf2 and 3) forming *Sc*CHRXII. The locally assembled contigs matching Scf2 harboured YLR409W and YLR410W whereas the locally assembled contig matching Scf3 harboured ORFs YLR415W and YLR417W. Collectively, these data confirmed the presence in *Sc*CHRXII, the presence of a 6.4 kb *S. eubayanus* DNA introgression.

Figure S2: Characterisation of contig00530 carrying an extra *ScADE1* copy in *S. pastorianus* CBS1483 strain. **A-** The graphs represent the ploidy prediction of CBS1483 CHRI made by

Magnolya (Nijkamp et al. 2012). The end right arm of CHRI showed an extra copy relative to the rest of the chromosome (four versus three). **B-** Annotation of the contig00530, the 9.0-kb fragment harboured four open reading frames YAR014C/BUD14, YAR015W/ADE1, YAR018C/KIN3 and YAR019C/CDC15 encoding the bud site selection protein 14, the phosphoribosylaminoimidazole-succinocarboxamide synthase, a serine/threonine-protein kinase and the cell division control protein 15 respectively. **C-** Localization of contig00530. The 8-kb library reads were mapped to contig00530 using BWA (Li et al. 2010). A total of 12126 reads were aligned, the mate paired reads of the mapped sequences were selected and locally assembled. Two contigs were assembled. The first contig was assembled from 4662 reads and matched a scaffold composing *S. cerevisiae* CHRI as expected. The second contig was assembled from 2351 reads and showed perfect identity with a scaffold forming *S. cerevisiae* CHRX. Analysis of the latter locally assembled contig revealed extremely high homology with multiple Ty sequences and did not allow unambiguous localization of this extra part of *S. cerevisiae* CHRI.

Figure S3: Venn diagram representing the single nucleotide variation found in pairwise comparison between A1 and the reference WS34/70 strain sequence (Nakao et al. 2009) (blue circle), between A2 and WS34/70 (green circle) and A+B11 and WS34/70 (red circle). **A-** Non sense mutations. **B-** Missense mutations.

Figure S4: Sugar consumption and growth profiles of the karyotypic variant A1, A1+B11 and A2 of the *S. pastorianus* WS34/70 strain. **A-** Sugar consumption profile of the *S. pastorianus* A1 strain. The strain was grown in WMM medium supplemented with a complex mixture of sugars from corn syrup containing 9.5 g·l⁻¹ maltotriose (closed square), 27 g·l⁻¹ maltose (closed circle), 4.5 g·l⁻¹ glucose (open circle), 0.75 g·l⁻¹ fructose (open square) as consumable sugars. The strain was grown in shake flask at a temperature of 20°C for 50 hours. maltotriose, maltose, glucose and fructose were analysed using HPLC. **B-** Sugar consumption profile of the *S. pastorianus* A1+B11 strain. **C-** Sugar consumption profile of the *S. pastorianus* A2 strain. **D-** Growth profile of strains A1(closed square), A1+B11 (open circle), A2 (open square). Optical density OD was measured at a wavelength of 660 nm.

References

- Li, H and Durbin, R. 1-3-2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**: 589-595.
- Nakao, Y, Kanamori, T, Itoh, T, Kodama, Y, Rainieri, S, Nakamura, N, Shimonaga, T, Hattori, M, and Ashikari, T. 2009. Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Res* **16**: 115-129.
- Nijkamp, JF, van den Broek, MA, Geertman, JM, Reinders, MJ, Daran, JM, and de, RD. 15-12-2012. De novo detection of copy number variation by co-assembly. *Bioinformatics* **28**: 3195-3202.

Table S1: Description of the libraries used in this study.

	Libraries			
	(Pseudo) Reads	500bp paired-end	3kb mate-pair	8kb mate-pair
Number of reads (10^{+6})	9.26	17.2	40.1	36.3
Read length (bp)	141 \pm	100	50	50
Total sequence (Mb)	1305.9	1718.9	2004.6	1816.8
Coverage (-fold)	52.2	68.8	80.2	72.7

Table S2: List of breakpoints identified in the *Saccharomyces pastorianus* CBS1483 genome.

Chromosome	Location	Scaffold	Primer pair
Break point <i>S. cerevisiae</i> / <i>S. cerevisiae</i>			
<i>ScXV-ScXI</i>	YOR381W/ <i>FRE3</i> -YKL220C/ <i>FRE2</i>	Scaffold21	S50
Break points <i>S. eubayanus</i> / <i>S. cerevisiae</i>			
<i>SeubIII-ScIII</i>	<i>BUD5</i> -MAT α	Scaffold27	S36
<i>SeubVII-ScVII</i>	YGL173C/ <i>XNR1</i>	Scaffold2	S24
<i>SeubVII-ScVII</i>	<i>ZUO1</i>	Scaffold2	S25
<i>SeubXIII-ScXIII</i>	YMR304W-YMR302C	Scaffold12	S30
<i>SeubXVI-ScXVI</i>	YPR159W/ <i>KRE6</i>	Scaffold15 → 39	S40
	YPR190C-YPR191W	Scaffold39	S42
	YPL240C/ <i>HSP82</i>	Scaffold38 → 8	S47
Break points <i>S. cerevisiae</i> / <i>S. eubayanus</i>			
<i>ScX-SeubX</i>	<i>TDH2</i> -ARS1016	Scaffold14	S31
<i>ScXVI-SeubXVI</i>	YPL036C/ <i>PMA2</i>		S34
<i>ScXII-SeubXII</i>			See Figure S2
Break points found in <i>S. pastorianus</i> and already present in <i>S. eubayanus</i> [12]			
<i>SeubII-SeubIV</i>	YBR030W/ <i>RKM3</i> -YDR012W/ <i>RPLAB</i>	Scaffold10	S28
<i>SeubVIII-SeubXV</i>	YHR014W/ <i>SPO13</i> -ARS807	Scaffold16	S32
<i>SeubIV-SeubII</i>	YDR011W/ <i>SNQ2</i> -YBR031W/ <i>RPLA4</i>	Scaffold1	S23
<i>SeubXV-SeubVIII</i>	YOR018W/ <i>ROD1</i> -YHR015W/ <i>MIP6</i>		S27

Table S6: Primers used in this study.

Primers	Sequence 5'→3'	
Chromosome copy number		
CBS1483_Sb1_Contig488_ADE1-Fw	GTACTTGGCCCTGGTTCTG	CNV <i>ScubADE1</i>
CBS1483_Sb1_Contig488_ADE1-Rv	CCTACAAGCTAGGCGAATCC	
CBS1483_Sc1_Contig530_ADE1-Fw	CAGGTACTGTGCATGGTTTG	CNV <i>ScADE1</i>
CBS1483_Sc1_Contig530_ADE1-Rv	AGCCAGTTCGCACTCTAC	
CBS1483_Sb3_Contig460_NFS1-Fw	CAACGGCCTTAACGACATAG	CNV <i>ScubNFS1</i>
CBS1483_Sb3_Contig460_NFS1-Rv	TGGGTAAGGATGATGCTTTG	
CBS1483_Sc3_Contig345_ABP1-Fw	GCCCAAAGACACATAATTGC	CNV <i>ScABP1</i>
CBS1483_Sc3_Contig345_ABP1-Rv	TGATTACGATGCTGCAGAAG	
CBS1483_Sc7_Contig382_EMP24-Fw	CGCCATAAATGTCCCTTCTTC	CNV <i>ScEMP24</i>
CBS1483_Sc7_Contig382_EMP24-Rv	GTCAGCTGGCTACTGGATTG	
CBS1483_Sc7_Contig260_MEP1-Fw	AGCCTCATACGGGGTATGTAG	CNV <i>ScMEP1</i>
CBS1483_Sc7_Contig260_MEP1-Rv	TTGGTATGGATGGCACTACAG	
CBS1483_Sb7_Contig350_SHY1-Fw	AGCAATTTGTCGCCAGTATC	CNV <i>ScubSHY1</i>
CBS1483_Sb7_Contig350_SHY1-Rv	AGCTGACCTATGACCCGATAC	
CBS1483_Sb8_ContigXXX_DUR3-Fw		CNV <i>ScubDUR3</i>
CBS1483_Sb8_ContigXXX_DUR3-Rv		
CBS1483_Sc8_ContigYYY_DUR3-Fw	GATGTTTTAGGCTCACC GGG	CNV <i>ScDUR3</i>
CBS1483_Sc8_ContigYYY_DUR3-Rv	TATCAAGAAAACGGTGCCG	
CBS1483_Sb8_ContigWWW_HIS3-Fw		CNV <i>ScubHIS3</i>
CBS1483_Sb8_ContigWWW_HIS3-Rv		
CBS1483_Sb8_ContigZZZ_MDM20-Fw		CNV <i>ScubMDM20</i>
CBS1483_Sb8_ContigZZZ_MDM20-Rv		
CBS1483_Sc8_ContigVVV_MDM20-Fw	ATGCAAAAACGGGGCAATTGC	CNV <i>ScMDM20</i>
CBS1483_Sc8_ContigVVV_MDM20-Rv	AACTCTAGCAAGAAGTTGTGC	
Verification of the sequence assembly		
S23-CBS1483_Sb4-Sb2_Scf1_SNQ2-Fw	CGTGTACTACCTCTTCCACGTGAGACAGAGTTCTC	Verification Scf1
S23-CBS1483_Sb4-Sb2_Scf1_PDX3-Rv	GCCTTAACGTTCTGTGTGGGTGTGG	
S24-CBS1483_Sc7-Sb7_Scf2_MPI5-Fw	GGATAATTTTGGTAATATATGCGTTACAAACGC	Verification Scf2
S24-CBS1483_Sc7-Sb7_Scf2_SUA5-Rv2	GTTGGAGAGAAGTGGATTGAGAGC	
S25-CBS1483_Sc7-Sb7_Scf2_PXR1-Fw	GCTGTTGATCATTTCATAGCGCAAAAGAAC	Verification Scf2
S25-CBS1483_Sc7-Sb7_Scf2_IMA1-Rv	GTAGATGCCTCTTCCAGAACATTG	
S30-CBS1483_Sb13-Sc13_Scf12_LIP1-Fw	GGACAGAACTTCAACTTGACCTCGG	Verification Scf12
S30-CBS1483_Sb13-Sc13_Scf12_UBP15-Rv	CGCTCCAATTTCTTCAAAGGCACAGAC	
S40-CBS1483_Sc16-Sb16_Scf15_KRE6-Fw	CGGTGAGTACGGTGGCTACTTTC	Connection Scf15-Scf39
S40-CBS1483_Sc16-Sb16_Scf39_SGV1-Rv	CCACTGGGACCACAATCAAGATAC	
S38-CBS1483_Sc12_Scf19_POM33-Fw	CTTATAATGCAATTTAAATGAGGTTGGTCC	Connection Scf38-Scf8
S38-CBS1483_Sc12_Scf33_PAU17-Fw	CATCTCCAGCGCTCTATCTGC	
S42-CBS1483_Sb16_Scf39_SKI3-Fw	GTCCAGCCTGATAGCGGATAAAC	Verification Scf39
S42-CBS1483_Sc16_Scf39_HPA2-Rv	CGTGTATTGGTGTACGGATGAGTC	
S31-CBS1483_Sc10-Sb10_Scf14_POL31-Fw	CCGTCAATTCAGTATCTTTTTCCCTGC	Verification Scf14
S31-CBS1483_Sc10-Sb10_Scf14_GPI14-Rv	CGGAAATTGACGTTAGCAAGGGTG	
S34-CBS1483_Sc15-Sc11_Scf21_MCH2-Fw	GAGATGAGATTGCTGTGCGTGAAG	Verification
S34-CBS1483_Sc15-Sc11_Scf21_RDR1-Rv	CTCATTTAGTTCGGTTAGCACACC	
S28-CBS1483_Sb2-Sb4_Scf10_YPK3-Fw	GTATGATGTTTGTGACAAATGTGGCGAGC	Verification Scf10
S28-CBS1483_Sb2-Sb4_Scf10_KCS1-Rv	GCCATTATCATGGCTAACCATCAGATGT	
S32-CBS1483_Sb8-Sb15_Scf16_DIA4-Fw	CTCCAAACATCAATTTGTCGAGGAAC TAGG	Verification Scf16
S32-CBS1483_Sb8-Sb15_Scf16_HST3-Rv	CCTTCTCTGCATGATGAATCCCAGC	
S27-CBS1483_Sb15-Sb8_Scf6_AUS1-Fw	GCGCAATCTTTTGTGTGGGTTCAAC	Verification Scf6
S27-CBS1483_Sb15-Sb8_Scf6_AUS1-Rv	CATTCTTTTGTGTGGGTTCAACAATG	
S29-CBS1483_Sc7-Sb7_Scf2_MPI5-Fw2	GTGCTAATCAGGGATAATTTTGGTAATTATGC	Connection Scf2-Scf3
S29-CBS1483_Sc7-Sb7_Scf2_SPO74-Rv2	CTTCCATTAGGGATTTTAGAAATCCACTTTTG	
S36-CBS1483_Sb3-Sc3_Scf27_TAF2-Fw	CTTCACGGGCTCGATGACTATAAGG	Verification Scf27
S36-CBS1483_Sb3-Sc3_Scf27_RRP43-Rv	CAATATATATCCGTAGACGGCCTTAATAGTG	
S47-CBS1483_Sb16_Scf8_YAR1-Rv	GTAGAGTTGACCCGCTTTCTTTG	
S47-CBS1483_Sc16_Scf38_IQG1-Fw	CGGGAATAGTGTGAACCTTTCTGG	
S44-CBS1483_Sc7-Sc8_Scf25_ERV29-Fw	CAGAGTATGGTCCTTGCGATGATG	
S44-CBS1483_Sc7-Sb7_Scf2_IMA1-Rv	CCAAGGCCGCAAGAAGAAGAATAAG	
S38-CBS1483_Sc12_Scf19_POM33-Fw	CTTATAATGCAATTTAAATGAGGTTGGTCC	
S49-CBS1483_Sb10-Sc10_Scf24_GPI14-Fw3	CGTGAACATGATTAACAGCAAAG	
S49-CBS1483_Sb10-Sc10_Scf24_SAG1-Rv3	GAGATGTTGCAATATCTGCGG	
S33-CBS1483_Sc12_Scf19_ERP2-Rv	CTTCATGCTCGTCAGTCAAAGTTTTTTC	
S37-CBS1483_Sc12_Scf33_EFB1-Rv	GACATCTAGAGTGACAATGGACTTAGCAG	
S38-CBS1483_Sc12_Scf33_PAU17-Fw	CATCTCCAGCGCTCTATCTGC	
S41-CBS1483_Sb16_Scf8+Scf39_KRE6-Fw	CTACAGTAAATGGTACTTTGCAATTGAG	
S43-CBS1483_Sc3_Scf27_FEN1-Rv	CCGCAAAGGTTAATGAGTATGTTAACG	
S45-CBS1483_Sc8_Scf36_SPO13-Fw	CTTTCAAGTACATTGGATGTTAATTTGTTCCG	
S46-CBS1483_Sc13_Scf7_ATM1-Fw	GACACAGTAAATATAATGGGCGAATGG	

Figure S1

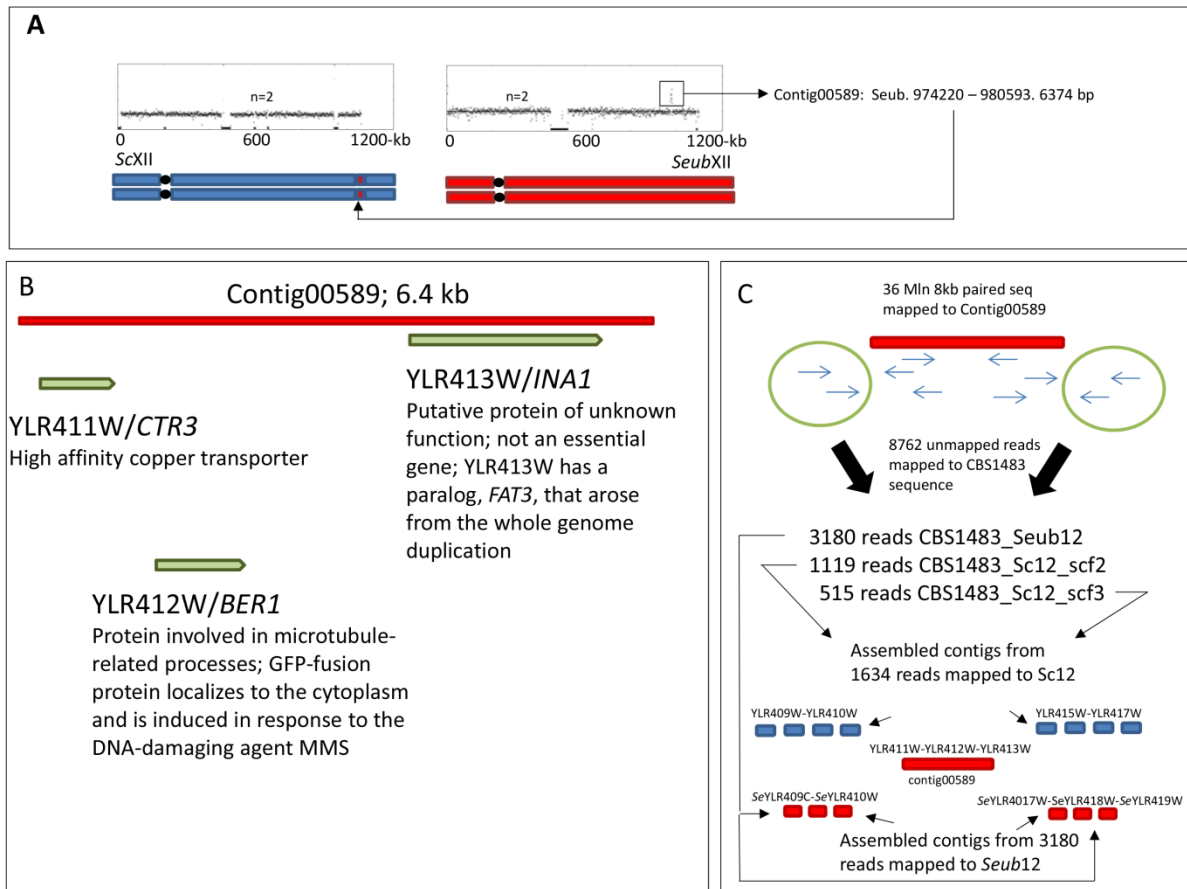


Figure S2

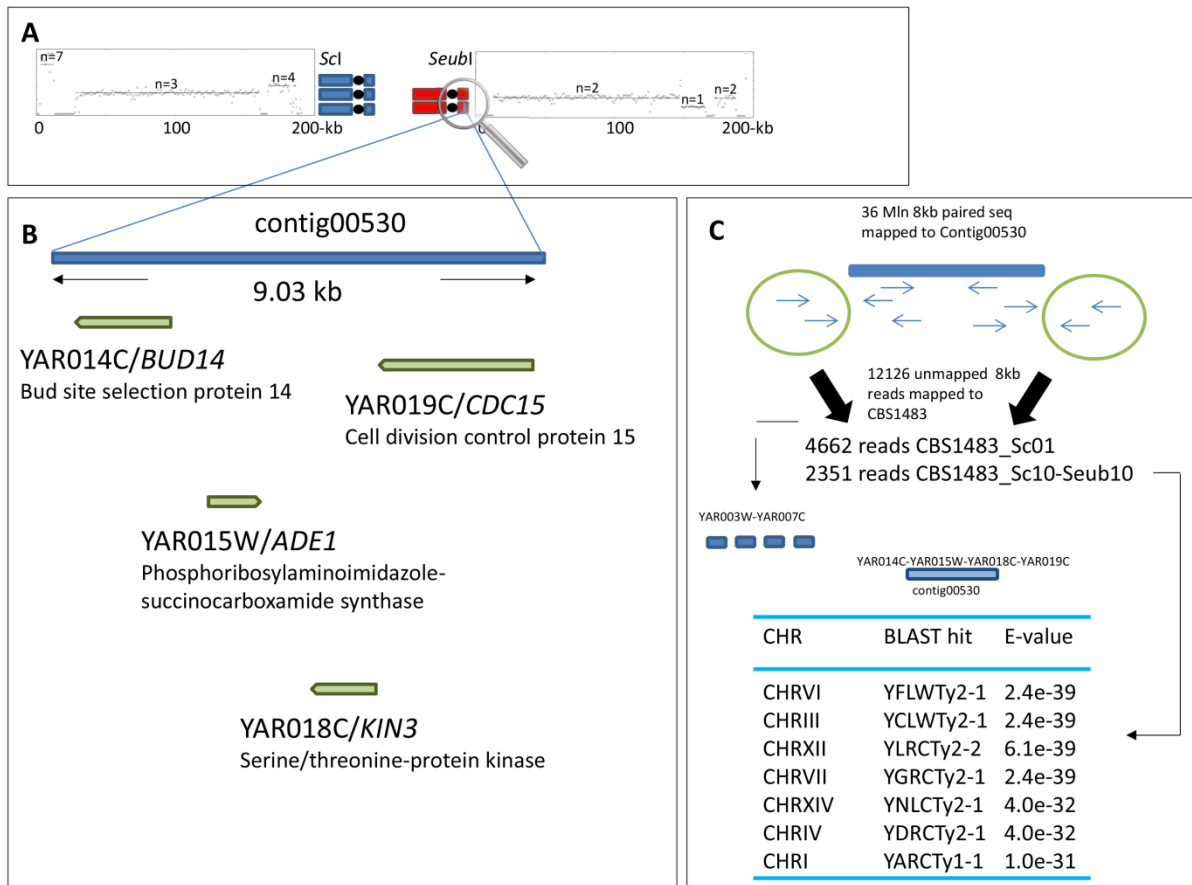


Figure S3

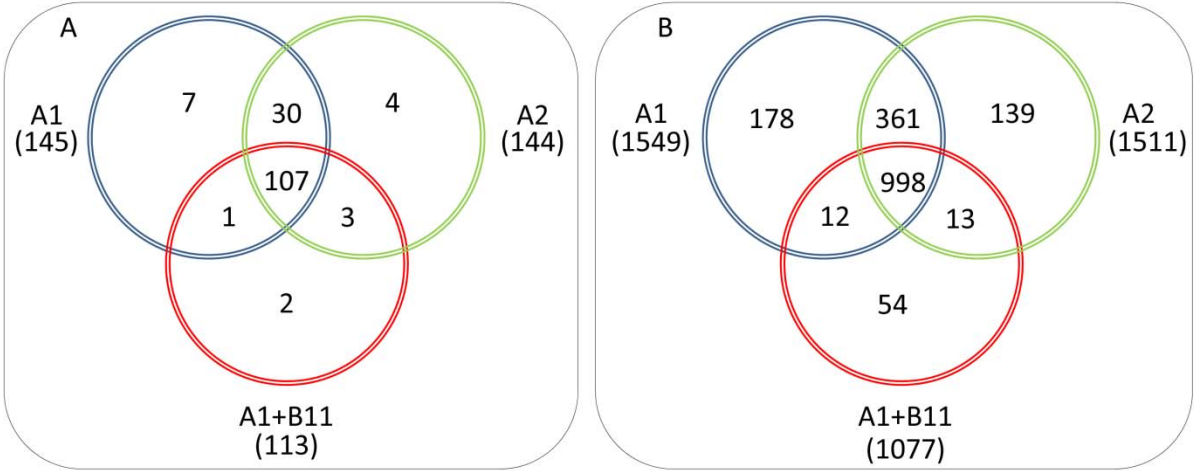


Figure S4

