## Mapping with Record When Joinmap can't handle your data

## Presenter: Dr. Allen Van Deynze, UC Davis

## Host: Heather Merk

## http://www.extension.org/plant\_breeding\_genomics

Slides available at http://www.extension.org/pages/63330 Sign up for PBG news at http://pbgworks.org



Please fill out the survey evaluation: (You will be contacted via email)

Watch past webinars and sign up for future webinars! http://www.extension.org/pages/60426



## Mapping with RECORD When Joinmap can't handle your data



Allen Van Deynze and Hamid Ashrafi University of California, Davis SOLCAP workshop April. 12, 2012,



# Acknowledgements

- Richard Michelmore
- Alex Kozik
- Maria Truco
- Sebastian Reyes Chin-Wo
- Theresa Hill
- Kevin Stoffel



# MadMapper: Suite of Python scripts for quality control of genetic markers, group analysis and inference of linear order of markers on linkage groups

Alexander Kozik, UC Davis Genome Center http://cgpdb.ucdavis.edu/XLinkage/MadMapper/

## **Record:** Marker order and distances

http://www.plantbreeding.wur.nl/UK/software\_record.html

# **CheckMatrix:** Visualization and validation of genetic maps using two-dimensional heat-plots.



## <u>MadMapper - group analysis and quality control of genetic markers</u> <u>STEP 1</u>

1. Group (cluster) all markers from locus file and assign markers to distinct linkage groups.

2. Select markers for linkage group 8 (as an example) and extract from that group reliable markers with high quality scores.

**LG8.loc** locus file with raw marker scores.

3. Then we will go to Part 2 (Win Record) and will try to infer linear order of selected markers on Pepper linkage group 8.



×	👌 Co	nvert 🔻 🛃	Select																				
x	G <mark>oo</mark> g	gle madn	mapper					- 3	Search 🔹 M	ore »												0 5	Sign In 🔌 🔹
x	Y	• <b>Q</b> Ya	ahoo! Sear	ch	SE	ARCH 🔻 🔶 🛅			W 🕅 🛛	<b>***</b>	J 🖗 • 🔽	3 • 🤇	<u>6</u> - 🚹	• 🗾 • 🤞	• 🛄 •							[	■ \$ \$
× Fi	nd: 🕻	Criterion				Previous N	lext 🛛 🃝 Opt	tions 🔻															
popt nind	= R = 1	01									$\checkmark$	- [	Line	s 🔨	~								E
nloc		357 1	<b>_</b>	3 4	5	6	7	8	9 1	0	11 :	12	13	14 15	16	17	18	19	20	21	22	23	24
; RS10 ve00	1	CS/N1900 B	(4) _ A	CS/N1901 - B A B	(5) C	S/N1903 (13)  3 -			CS/N1905  A A	(17)	CS/N4664 - 1		в	CS/N1906 (  B B			1907 (25) B B		008 (29) - A	CS/N19 - A		CS/N19 B B	
PVV4 nga5 ACC2 0846	9		- А А	A – A B A B A B	- B B	3 –	- A -	B B B	 A A A A		A I	- B B B	В	 B B B B B B	– A A	- В А	- B B	- A - A	A A A	- А А	- А А	B B B	B B B
A32 agp1 ve00	6 2	B B B	A A A	A – A B A B	- B B		-	- B B	 - A A A		- 1 A 1 A -	в В -	- A	B – B – B –	A A A	- - A	- B B		- A A	- A A	- A A	- - B	- B -
SEP4 A1 ATEA g471	T1	B B B	– A A	- B A - A B A B	- - B B	- – - – 3 –	- - A	- - B	A – – – A A A A		- 1 A 1	B B B B	A	 B - B B B B	A A A	-	- - B B	- - - A	А - А А	- - A A	А - А А	- B B	В - - В
m488 ve00 ve00	3 4	B B B	А А А	A B A B A B	B B B	3 –	-	B B B	A A A A A A		A I A I A I	B B B	A A A	B B B B B B	A A A	-	B B B	A - -	A A A	А А А	А А А	B B B	B - -
agp1 mi37 po GT45	2	B	Μ	arker	S B		A A - A	В - - В	- A A A A - A -		A I A ·	B B - B	A A	B – B B B – B B	A A A A	-	B B B B		А А А -	А А А А	А А А А	- B B B	В - - -
PAI1 apx1 cSOD	A	-	- A B	 A B A - A B			- A A	- B B	- A A A A A		A I A I	B B B	-	 B B B B	A A A	-	- B B	- A -	A A A	- A -	A A A	B B B	- B -
mi10 phyA m241 nga6	A	B B B	A A A	A B A B A B - B	B B B	- 3	А А А А	B B B	- A A A A - A A		 A 1	B - B B	- A	B B  B A B A	A A A	-	В - В В	- A A	А - А А	А - А А	А - А А	B B B	B B B B
mi44 hma2 ve00 CDs1	5	B B B	A A -	A B A - A B A B	B - B B		A - -	В - В В	A A  A A A A		A I	B B B B		B A B - B A B A	A A A	- - - B	В - В В	- - -	A - A	А - А А	A - A A	В - В В	В - В
pC1 NCC1 ATTS		B -	A A B	A B A - A A	B - B	- 	- A -	B - -	A B - B B A		A I A I	B B A		B A  A B	A A A	B B A	В - В	- - B	A A -	A A A	A A A	B B A	B - B
agp2 CDs7 ve00	0e 6	B B -	A A -	A B A B A -	B B B	3 – 3 –	A - -	B B B	- B A B A B		A I A I	B B B	A A	B – B A B A	A A A	B B B	B B B	- - -	A A A	A - A	A A A	- B B	B - B
pC2 0818 EG17 g378	G9	-	Б - В В	A B  A B B B	B B B			В - В В	A B  A B A B		 A 1	B - B B	- A	в А  В А В А	А - А А	В - В В	В - В В	- B B	А - А А	А - А А	А - А А	ь - В В	В - В В
ve00 mi34 mi11 Y3	8 3	-	B B B	B B B B B B	B B B	3 – 3 – 3 –	A A	B B B	A B A B A B		A I A I	B B B B	A A	B A B A B A B -	A A A	B B B	B B B	- -	A A A	A A A	A A A	B B B	B B B
ve03 wak1 pFNR	6	B	- B B	B B B - B B	- B - B		- - A	В - -	A B  A B		A 1 - 1	B B B	A -	B A B - A -	A A A	В - В	- - B		А - А	_ _ А	A - -	В - В	- В В
CAT3 pC3 mi20	3	B B B	B B B	B B B B B B	B B		A - A	B B	A B A B A B		A I A I	B B B	A	B – B – B A	A A A	B B	A B B		B A A	А А А	- A A	B B	B B B
JD11 g382 DHS3 m235	9		B B B	 B B B - B B	- B - B			- В - В	 A B  A B		A 1 - 1	B B B		в – В А В – В А	A A B A	– – A	- B - B	- - B	- A - A	- В - В	- A - A	- B - A	- B - B
ve03 mi26	7	- B	B	 В В	B	- 3	- A	B B	A – A B		A I	B B	A	B A B A	A A	A A	B	-	A A	A A	A A	A B	- B -

**First Step - Group Analysis -** We will run the script with the following options/arguments:

Prompt> python Python\_MadMapper\_V248\_RECBIT\_012.py Testdata.loc Testdataout 0.2 100 25 frame.IDs 0.33 50 NOTRIO 3

# Prompt> python Python\_MadMapper\_V248\_RECBIT\_012.py

PROGRAM USAGE: MAD MAPPER TAKES 10 ARGUMENTS/OPTIONS IN THE FOLLOWING ORDER: (1)input\_file[LOC\_DATA/MARKER SCORES] (2)output\_file[NAME] (3)rec\_cut[0.2] (4)bit\_cut[100] (5)data\_cut[25] (6)group\_file[OPTIONAL] (7)allele\_dist[0.33] (8)missing\_data[50] (9)trio\_analysis[TRIO/NOTRIO] (10)double\_cross[3]



## Results of Step 1

Output will be represented by 76 files [Note, that at this time the script was running with "NOTRIO" option]

At this step we are interested in <u>TestDataout.x tree clust</u> file only. This file has the information about marker clustering/grouping. From this file we will select markers belonging to the linkage group 8.

We will analyze <u>TestDataout.x</u> tree clust file using MS Excel to find and highlight (select) a set of markers belonging to the linkage group 8. <u>TestDataout.x</u> tree clust.xls - tree clustering file in MS Excel format.



		47		-			ţ.	x A	COF	H_00:	171_0	0174	_AAD														
A	В	CD	) E	F	G	н	IJ	I K	L	M	N	0	Р	Q	R	S	T	U	V	W	Х	Y	Z	AA	AB	AC	AD
1	0.18	0.16 0.14		0.10	0.09	0.08	0.07	0.05	0.04	0.03	0.02	0.01	0.00	Graph Class	Node Type		frame LG	fr. position		orig. order		assigned LG	MARKER		A score	B score	total
26 1	1 1	1	1 1	11	27	32	37 6	5 9	3 14	3 206	6 272	387	845 LINKED	GROUP_00845	DILUTEDNODE	•••	-	-	•••	4138	•••	LG_1	BGIF_00075_00078_AAD	•	67	54	123
27 1	1 1	1	1 1	11	27	32	37 6	5 9	3 14	3 206	6 272	387	845 LINKED	GROUP_00845	SATURATED_NODE	•••	LG1	12	2 •••	1851	•••	LG_1	ANII_00243_00246_AAD	•	68	54	123
28	1 1	1	1 1	11	27	32	37 6	5 9	3 14	3 206	6 272	387	1342 SINGLE	NODE_01342	SATURATED_NODE	••••	-	-	••••	3379	•••	LG_1	AXYW_00167_00188_AAD	•	68	54	123
29 1	1 1	1											1368 SINGLE		SATURATED_NODE		-	-	••••	3475		LG_1	AYXK_00389_00392_AAD	•	69	49	123
30 1	1 1	1												NODE_01357	SATURATED_NODE		-	-	••••	3443		LG_1	AYQZ_00151_00180_AAD		70	53	123
								_	_						SATURATED_NODE	_	LG1	20	, ••••	5222		LG_1	BOFS_00123_00126_AAD		95	27	123
	1 1	17 1	_			_	_	_	_	_	_	_			SATURATED_NODE	••••	-	-	••••	5223	_	LG 1	BOFS 00127 00142 AAD		96	27	123
33 2	2 2	2		2 2					_		2 2		2 LINKED		DILUTEDNODE		-	-			••••	LG_2	AABN_00123_00130_AAD		63	58	123
	222		22 22			2			2		22 22	2		GROUP_00002 GROUP_00002	DILUTEDNODE DILUTEDNODE	•••	-	-	•••	1131 1251		LG_2 LG_2	AIDH_00393_00400_AAD AJAD_00051_00054_AAD		60 62	60 59	123 123
36 3	2 2		2 2			2			_		2 2	2		GROUP 00002	DILUTED NODE	•••	-		•••	1344		LG 2	AJSB_00231_00238_AAD	•	59	60	123
37 3	2 2	2	2 2	2	2	2	2	2	2	2 2	2 2	2		GROUP 00002	DILUTED NODE			-	•••	1344		LG 2	AJZL 00049 00052 AAD	•	60	59	123
38 3	2 2	2	2 2	2	2	2	2	2	2	2 2	2 2	2		GROUP 00002	DILUTED NODE	•••		-	•••	1387		LG 2	AKGZ 00693 00696 AAD		59	60	123
	2 2		2 2	2	2	2	2	2	2	2 2	2 2	2		GROUP 00002	DILUTED NODE	•••		-	•••		•••	LG 2	ABPS 00409 00412 AAD		61	59	123
40 3	2 2		2 2	2	2	2	2	2	2	2 2	2 2	2		GROUP 00002	DILUTED NODE	•••	-	-	•••	1658		LG 2	AMAP 00199 00206 AAD		59	59	123
	2 2	2	2 2	2	2	2	2	2	2	2 2	2 2	2		GROUP_00002	DILUTED NODE	•••	-	-	•••	1659		LG 2	AMAP 00309 00312 AAD		58	59	123
42	2 2	2	2 2	2	2	2	2	2	2	2 2	2 2	2		GROUP_00002	DILUTED NODE	•••	-	-	•••		•••	LG 2	ACDJ 00645 00650 AAD		61	60	123
	22	2	2 2	2	2	2	2	2	2	2 2	22	2		GROUP_00002	DILUTEDNODE	•••	-	-	•••	205	•••	LG 2	ACGH_00049_00052_AAD	•	59	61	123
44 3	22	2	2 2	2	2	2	2	2	2	2 2	2 2	2	2 LINKED	GROUP_00002	DILUTEDNODE	•••	-	-	••••	211	•••	LG_2	ACHU_00187_00190_AAD	•	62	59	123
45 3	22	2	2 2	2	2	2	2	2	2	2 2	2 2	2	2 LINKED	GROUP_00002	DILUTEDNODE	•••	-	-	•••	239	•••	LG_2	ACOH_00149_00156_AAD	•	63	59	123
46 3	2 2	2	2 2	2	2	2	2	2	2	2 2	2 2	2	2 LINKED	GROUP_00002	DILUTEDNODE	•••	-	-	•••	240	•••	LG_2	ACOH_00165_00168_AAD	•	62	59	123
47 2	22	2	2 2	2	2	2	2	2	2	2 2	2 2	2	2 LINKED	GROUP_00002	DILUTEDNODE	•••	-	-	•••	241	•••	LG_2	ACOH_00171_00174_AAD	•	57	59	123
48 2	22	2	2 2	2	2	2	2	2	2	2 2	2 2	2	2 LINKED	GROUP_00002	DILUTEDNODE	••••	-	-	•••	2422	•••	LG_2	AQWO_00127_00132_AA	•	61	61	123
49	22	2	2 2	2 2	2	2	2	2	2	2 2	2 2	2	2 LINKED	GROUP_00002	DILUTEDNODE	••••	-	-	••••	2423	•••	LG_2	AQWO_00133_00148_AA0	•	62	60	123
50	22	2	2 2	2 2	2	2	2	2	2	2 2	2 2	2	2 LINKED	GROUP_00002	DILUTEDNODE	••••	-	-	••••	2647	•••	LG_2	ASCK_00193_00196_AAD	•	62	59	123
51	2 2	2	2 2	2 2	2	2	2	2	2	2 2	2 2	2	2 LINKED	GROUP_00002	DILUTEDNODE	••••	-	-	••••	2648	•••	LG_2	ASCK_00429_00432_AAD	•	63	58	123
52 2	22	2	2 2	2 2	2	2	2	2	2	2 2	2 2	2	2 LINKED	GROUP_00002	DILUTEDNODE	••••	-	-	••••	2649	•••	LG_2	ASCK_00435_00438_AAD	•	61	59	123
	22	2	2 2	2	2	2	2	2	2	2 2	22	2		GROUP_00002	DILUTEDNODE	••••	-	-	••••	2697		LG_2	ASJA_00399_00424_AAD	•	63	59	123
	22	2	2 2	2	2	2	2	2	2	2 2	22	2	2 LINKED	GROUP_00002	DILUTEDNODE	••••		-	••••	2698		LG_2	ASJA_00425_00428_AAD		64	58	123
	22		2 2	2 2	2	2	2	2	2	2 2	22	2		GROUP_00002	DILUTEDNODE	••••		-	••••		••••	LG_2	ACYH_00535_00544_AAD	•	61	58	123
	2 2	-	2 2	2	2	2	2	2	2	2 2	2 2	2		GROUP_00002	DILUTEDNODE	••••		-	••••	2812		LG_2	ATAI_00993_00996_AAD	•	63	60	123
	2 2	2	2 2	2	2	2	2	2	2	2 2	2 2	2		GROUP_00002	DILUTEDNODE	••••		-	••••	2813		LG_2	ATAI_01007_01012_AAD	•	64	58	123
	2 2	2	2 2	2	2	2	2	2	2	2 2	2 2	2	2 LINKED		DILUTEDNODE	••••		-	••••	2939		LG_2	ATUV_00125_00134_AAD		61	57	123
	2 2	2	2 2	2	2	2	2	2	2	2 2	2 2	2	2 LINKED		DILUTEDNODE	••••		-		2940		LG_2	ATUV_00431_00434_AAD		57	61	123
60	2 2	2	2 2	2	2	2	2	2	2	2 2	2 2	2	2 LINKED		DILUTEDNODE	••••	-	-		2958		LG_2	ATYG_00179_00184_AAD		62	60	123
	2 2	2	2 2	. 9	9	9	9 1	0 1	1 10	9 154	4 195	264	365 LINKED		SATURATED_NODE		-	-				LG_2	AFML_00751_00754_AAD		50	71	123
62 2 63 2	2 2	2	2 2	9	9	9	9 1 9 1						1490 SINGLE 1204 SINGLE		SATURATED_NODE SATURATED_NODE		-	-	•••	3824 2942		LG_2	BDDA_00077_00080_AAD	•	50	72 60	123 123
		14 1			-	-			_		0 446 8 315		692 SINGLE		SATURATED_NODE		-	-	•••	1479		LG_2	ATVI_00321_00324_AAD	•	34	00	123
65 3			3 3			3	3	-			3 3	3	3 SINGLE		SATURATED NODE				•••		•••	LG 3	AABS 00075 00078 AAD		66	56	123
	3 3			3				_	_		3 3				DILUTEDNODE	•••	-	-	•••	2125		LG_3	AOVW_00045_00066_AAD		65	58	123
	3 3				3	3	3		3		3 3				DILUTEDNODE			-	•••	2125		LG_3	AOVW_00683_00688_AAD		64	58	123
	3 3				3	3	3				3 3				DILUTEDNODE	•••	-	-		5205		LG_3	BODC_00337_00342_AAD		66	57	123
	3 3	-		3							3 3				SATURATED_NODE			-	•••	5205		LG_3	BODC_00347_00354_AAD		63	57	123
	3 3	-									3 3				SATURATED NODE			-	•••	2968		LG_3	AUBH_00453_00464_AAD		61	61	123
71 3		-	3 3		-	-	3				3 3				SATURATED_NODE			-	•••	2969		LG_3	AUBH_00465_00468_AAD		61	58	123
	3 3	3									3 3				SATURATED_NODE			-	•••	3645		LG_3	BBCX_00493_00518_AAD		64	59	123
73 3		3	3 3		-	-	3				3 3				SATURATED_NODE			-	•••	3873		LG_3	BDLF_00205_00208_AAD		63	60	123
	3 3	3	3 3								3 3				SATURATED_NODE			-	•••	5086	•••	LG_3	BNEA_00235_00278_AAD		64	59	123
	3 3	3	3 3	-	-	-	3				3 3				SATURATED_NODE			-	•••	5087		LG_3	BNEA_00283_00290_AAD		62	59	123
	3 3	3	3 3	3	3								1787 COMPL	ETE_GRAPH_01787	SATURATED_NODE	•••	-	-	•••	5088	•••	LG_3	BNEA_00365_00370_AAD	•	61	59	123
77 3	2 2	2	° °	9	2	2	2	2	2	2 :	o o	2	1797 COMBI	ETE CDADH 01797	SATURATED NODE	•••			•••	5707	•••	10.2	BSCV 00447 00460 AAD	•	62	50	172

		_			AAD.x_tre	e_clust - Mic	rosoft Excel		
Home Insert	Page Layout			_		_		_	@ _ ¤ x
Cut	Calibri	* 11 * A	hhΔ	ing	o Segreg	ati	on	Data	a by Group
ste 🛷 Format Painter	B I U -		Auu	3111				Dutt	Clear Sort & Find & Clear Sort & Find & Find & Find & Clear & Find & Fin
Clipboard 😼		Font	*		A A				Editing
	(• × ✓   <b>f</b> *		,Sheet1!\$A\$2:\$DT	\$5999,Sheet					×
R S	Т	UV	w x	Y	Z AA	AB	AC	AD AE	AF AG AH AI AJ AK
Node Type	frame LG	ft. position	orig. order	assigned LG	MARKER	A score	B score	total	DriginalSPPs
SATURATE ***		***	0 ***	LG_1	AAAH_00127_00130_AAD *	66	56	123 *	0 \$5999,Sheet1!A\$1,FALSE) A B B A B
SATURATE ***		***	9 ***	LG_1	AADK_00505_00522_AAD *	69		123 *	1 AADK_00505_00522_AAD A B B A B 2 AQBK_00375_00396_AAD A B B A B
SATURATE ***		***	2322 ***	LG_1	AQBK_00375_00396_AAD *	70		123 *	2 AQBK_00375_00396_AAD A B B A B
SATURATE ***		***	2639 *** 2285 ***	LG_1	ASBS_00243_00246_AAD *	71		123 *	3 ASB5_00243_00246_AAD A B B A B 4 APVE 01007_01020_AAD A B B A B
SATURATE ***		***	2285 ***	LG_1 LG_1	APVE_01007_01020_AAD * AQXF_00623_00628_AAD *	70 68		123 * 123 *	4 APVE_01007_01020_AAD A B B A B 5 AQXF_00623_00628_AAD A B B A B 6 AUBX_00195_00212_AAD A B B A B
SATURATE ***		***	2974 ***	LG_1	AUBX_00195_00212_AAD *	69		123 *	6 AUBX_00195_00212_AAD A B B A B
SATURATE ***		***	3045 ***	LG_1	AUPW 00825 00828 AAD *	65		123 *	7 AUPW_00825_00828_AAD A B A B
D SATURATE ***		***	3046 ***	LG_1	AUPW 00829 00842 AAD *	67	56	100 *	
1 SATURATE ***		***	3035 ***	LG_1	AUOA_006 Function Arguments				
2 SATURATE ***		***	316 ***	LG_1	ADEP_0007 VLOOKUP				AAD A A B A B
SATURATE ***		***	4721 ***	LG_1	BKAA_0006 Lookup_val	ue <b>672</b>		📧 = "AA#	AAD A A B A B AAD A A B A B AAD A A B A B AAD A A B A B
4 SATURATE ***		***	3233 ***	LG_1	AVVE 0092		A\$2:\$DT\$5999		
5 SATURATE ***		***	2979 ***	LG_1	AUEC_0031			= (AA	AAD A B B A B
5 SATURATE ***		***	576 ***	LG_1	AEQ5_0033		\$1		AAD A B B A B
7 SATURATE ***		***	577 ***	LG_1	AEQS_0055 Range_look	up FALSE		🔣 = FALS	AAD A B B A B
SATURATE ***		***	2980 ***	LG_1	AUEC_0093				AAD A B A B
SATURATE***		***	3196 ***	LG_1	By default, the table must be			returns a value in the s	VAH_00127_00130_AAD";         AAD         A         B         A         B           name row from a column you specify.         AAD         A         B         B         A         B           AAD         A         B         B         A         B         B         A         B
SATURATE ***		***	3197 *** 3198 ***	LG_1	AVPX_0004 by default, the table must be	kup value is	s the value to be	found in the first colum	In of the table, and can be a value, a
SATURATE ***		***	4125 ***	LG_1 LG_1	AVPX_0005 Loo BGFQ_0020		eference, or a te		in of the table, and can be a value, a AAD A B B A B AAD A B B A B
3 SATURATE ***		***	5969 ***	LG_1	BTVA_0035				
4 SATURATE***		***	5391 ***	LG 1	BPNP 0011 Formula result = AAAH_001	27 00130 AAD	D		
SATURATE***		***	5961 ***	LG_1	BTUN 0021				
5 SATURATE ***		***	5962 ***	LG_1	BTUN_0022				
7 SATURATE ***		***	5968 ***	LG_1	BTVA_00205_00224_AAD *	74	49	123 *	25 BTVA_00205_00224_AAD A B B A B
B DILUTED_***		***	1439 ***	LG_1	AKLF_00589_00600_AAD *	66	55	123 *	AND THE REAL PROPERTY AND
DILUTED_***		***	1440 ***	LG_1	AKLF_00613_00618_AAD *	65	58	123 *	Rome Dente Apgragent Families Det Rome View Annual A data - 11 - A dat
DILUTED_***		***	1441 ***	LG_1	AKLF_00945_00948_AAD *	67		123 *	
DILUTED_***		***	1442 ***	LG_1	AKLF_01345_01352_AAD *	66		123 *	
2 DILUTED_***		***	1443 ***	LG_1	AKLF_01353_01360_AAD *	63		123 *	A         B         C         D         E         F         G         H           1         3         2         3         4         5         6         7         3         15         16         17         18         19           2         AAAH (0227) (0230) AAD         A         B         B         A         B         B         A         B         A         A         B         A         B         B         B         B         B         A         A         A         A         A         B         A         B         B         B         B         B         B         A         A         A         A         A         B         B         A         B         B         B         B         B         A         A         A         A         A         B         B         B         B         B         A         A         B         A         B         B         B         B         B         A         A         A         A         A         B         A         B         B         B         B         B         B         B         B         B         B         B<
DILUTED ***		***	1522 *** 1523 ***	LG_1	AKXB_00069_00076_AAD *	66		123 * 123 *	3         AABN_00223_00300_AAD         A         A         A         A         A         A         A         B         A         A         A         B         B         B         B         B         B         B         B         A         A         A         B         B         B         B         B         B         A         A           4         AABS_00075_00076_AAD         A         B         A         A         B         B         B         A         B         A         B         A         A         A         B         B         B         A         A         A         A         B         B         B         A
DILUTED_ ***		***	1523 ***	LG_1 LG_1	AKXB_00337_00360_AAD * AKXB_00431_00442_AAD *	66 66		123 *	S         AMP (2014) (2014) (AAD         A         A         A         A         A         A         A         A         A         A         A         B         B         A         -         A         B         B         A         -         A         B         B         A         -         A         B         B         A         -         A         B         B         A         -         A         B         B         A         A         B         B         A         A         B         B         A         A         B         B         A         B         B         A         B         B         A         B         B         A         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B
DILUTED ***		***	1525 ***	LG_1 LG_1	AKXB_00533_00558_AAD *	66		123 *	B         AAD         B         B         B         B         B         B         B         A         B         B         A         B         B         B         B         B         B         B         A         B         B         B         B         B         A         B         B         B         B         A         B         B         A         B         B         B         B         A         B         B         B         B         B         B         A         B         B         B         A         B         B         B         A         B         B         B         A         B         B         B         A         B         B         B         B         A         B         B         B         A         B         B         B         B         A         B         B         B         B         A         B         B         B         B         A         B         B         B         A         B         B         B         A         B         B         B         A         B         B         B         A         B         B         B         B
DILUTED_***		***	1526 ***	LG_1	AKXB_00555_00558_AAD *	64		123 *	12         AAAD         0515         OHIG         AAA         B         B         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         <
BILUTED ***		***	2721 ***	LG 1	ASOA 00389 00410 AAD *	71		123 *	ID         Add (get)         A         A         B         B         A         B
DILUTED_***		***	2897 ***	LG_1	ATPE_00581_00584_AAD *	69		123 *	15 AAHT 01135 00133 AAD A A B A A B B B B B A A B A B A B A
DILUTED_***		***	3943 ***	LG_1	BEKT_00069_00074_AAD *	65		123 *	11         AAB_00327_00032_AAD         A         B         A         B         A         B         A         B         A         B         B         B         A         B         A         B         B         B         A         A         A         B         B         B         A         A         A         B         B         B         A         A         A         A         B         B         B         A         A         A         A         B         B         B         A         A         A         A         B         B         A         A         B         B         A         A         B         B         A         B         B         B         A         A         A         A         B         B         A         A         B         B         B         A         B
DILUTED_***		***	3944 ***	LG_1	BEKT_00075_00088_AAD *	65	57	123 *	20 AAK 002735 00228 AAD 8 8 8 A A A A A A A B 8 A A A A B B A A A A
DILUTED_***		***	3945 ***	LG_1	BEKT_00093_00116_AAD *	67	56	123 *	121         AMS 50125 0022 AAD         A         B         B         A         A         A         B         A         A         B         A         B         A         A         B         A         A         B         A         B         A         B         A         B         A         A         B         A         A         B         A         B         B         A         B         A         B         A         A         B         A         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B
DILUTED_ ***		***	4679 ***	LG_1	BJTU_00233_00238_AAD *	66	51	123 *	25 AAUX,00051,00054,AAD - B B A A A A B A B A B A B A B A B A B
AAD.x_tree_c	lust / Sheet1	. / 😒 /			· · · · ·			I 4	27 MAX_00051 0005 AAD 8 A 8 A 8 A 8 A 8 A A A A A A A A A
t 3 2 2				1		13			Image: state of the s
									Image: Note of the state of the st
									M AANNY (2005) 2002 JAAD         B         B         B         A         A         B         B         A         A         B         B         A         B         B         A         B         B         A         A         B         B         A         A         B         B         A         B         B         A         B         B         A         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         A         B         B         B         B         A         B         B         A         B         B         A         B         B         B         A         B         B         B         A         B         B         B         A         B         B         B         B

## Save for each linkage group

🗶 🔙 🔊 • (° - 💼 ź.	Z    -												ChMTR	X test d	ata - Micro	soft Excel														x
	ert Page	lavout	Form	ular	Data	Daviau	View	JMP	Acrobat					orresta	oto micro	Joirt Excer													~ @	- # X
	ert Page	e Layout	Form	iulas	Data	Review	view	JMP	Acrobat																	-	_			
Cut	Calibri		11 .	· A A	= =	: 🔚 👌	\$7	📑 Wrap Tex	đ	General	*	1 8	≤ę		Normal	E	lad		Good		Neutral	Calcu	lation	^		× 🔲	Σ AutoSum	ž Ž	ñ	
Paste	BZI	n - L 🗆	- 8	- A -	= =		= =	Han Mercie &	Center v	\$ - %	• • 0 .00		ditional Fo	ormat	Check Cel	F	xplanat	tory	Input		Linked Cell	Note				elete Format	🐺 Fill 🔻	Sort &		
<ul> <li>V Format Painter</li> </ul>												Form	atting + as 1	Fable 🔻			- and a							-	*	• •	🖉 Clear 🔻	Filter 🔻	Select *	
Clipboard 🖬		Font		6	ž.		Alignme	nt	Gi.	Numb	er G	i						Styl	es						(	ells	E	diting		
AW47	• (*	<i>f</i> <sub>*</sub> =	VLOOK	UP(\$Z47	,Sheet1!	\$A\$2:\$D	)T\$5999	,Sheet1!Q\$	1,FALSE)																					¥
AG		AH	AI	AJ	AK	AL	AM	AN	AO	AP AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB BC	BD	BE	BF	BG	BH	BI BJ	BK	BL	BM 🔺
OriginalSPP	s																													
	<u> </u>																													
1		1	2	3	4	5	6	7	8	9	10 11	1	12 13	1	4 15	16	17	18	19	20	21	22 2	3 24	4	25	6 27	28	29 3	0 3	1 -
35 AJAD_00051_00054_AAD		А	A	<u>ر ۱</u>	A A	. <i>1</i>	A	A A	В	В	В	A	В	в	В	В	в	A	A A		B B	В	В	A	В	- /	В	A	А	B 🔺
36 AJSB_00231_00238_AAD					а а	. 4	A -		В	В	В		В	В	В	B	в	В	A A		B B	В	В	A	В	- /	В	A		В
37 AJZL_00049_00052_AAD 38 AKGZ_00693_00696_AAD		A	A				Δ		B	8	B	A	в	B	8	B	B	B	A A		о В в в	B	в	A A	B		B	A	A	B
39 ABPS_00409_00412_AAE		A			A A	í	A	A A	B	В	в	A	в	в	в	-	8	в	A A		B B	В	в	A	В	A A	В	Â	A	в
40 AMAP_00199_00206_AA	D A				A A		A		в	В	В		в		в	в	в	В	A A		B B		в	A	в	- 4	В	A	А	в
41 AMAP_00309_00312_AA			-		A A	1	A		В	В	В		В	В	В	B	В	A	A A		B B	В	В	A	В	- /	В	A	-	В
42 ACDJ_00645_00650_AAD 43 ACGH_00049_00052_AA		A	A			i i i	A A	A A	B	8	B	A	B	B	8	B	8 R	B	A A		6 6 8 8	B	8	A 	B	- B	8	A	A	B
44 ACHU_00187_00190_AA		A	- A		A A			A A	В	8	в	A	в	в	в	в	8	Α	A A		B B	в	в	A	в	A A	в	A	A	в
45 ACOH_00149_00156_AA	D A	A	A	, j	A A	1	A		В	В	в		в	в	в	в	в				в в	в	в	A	в	A A	В	A		в
46 ACOH_00165_00168_AA		A	A		A A	4	A		В	В	В		В	В	В	В	8				8 B	В	В	A	В	A A	В	A	A	В
47 ACOH_00171_00174_AA		A	-		а а	. 4	A A	A A	B	В	B	A	B	B	В	B	B 9	A	A A		B B	B	В	A	B	A A	В	A	-	B B
49 AQWO_00127_00132_AA		A	-		ч <i>н</i> 4 А		а А		B	B	B	A A	B	B	B	B	D R	B			о о в в	B	B	A	B		B	A	A	B
50 ASCK_00193_00196_AAD					A A		A		в	в	в		в		в	в	в	-	A A		8 B		в	A	в	A A	в	A		в
51 ASCK_00429_00432_AAD		A	A	<u> </u>	а а	. 4	A		В	В	В		В		В	В	в				в в		в	A	в	A A	В	A		в
52 ASCK_00435_00438_AAD		-	-		а а		A •	A A	В	В	В	A	В	В	В	В	В	A	A A		8 B	В	В	A	В	A A	В	A	A	В
53 ASJA_00399_00424_AAD 54 ASJA_00425_00428_AAD		A 	-		α, Α Δ Δ		A A		B	B	B	A A	B	B	B	B B	B B	А Δ			666 88	B	B	A A	B		B	A A	A	В
55 ACYH_00535_00544_AAD			A	i j	A A		A		В	В	в		В	в	B	-	8	В	A A		вв	В	в	A	В	A A	в	A		в
56 ATAI_00993_00996_AAD	A		A	د <u>ب</u>	A A	. 4	A		в	В	в		в		в	В	в				в в		в	A	в	B /	В	A		в
57 ATAI_01007_01012_AAD			A		A A	. 4	A		В	В	В		В	В	В	В	В	A	A A		в в	В	В	A	В	A A	В	A		В
58 ATUV_00125_00134_AA0 59 ATUV_00431_00434_AA0		A	A		Α Α Δ Δ		A A		B	B	- R	A A	B	B	в   в	- 8	B R	B			B B R R	B	B	A A	B		B	A	A	В
60 ATYG_00179_00184_AAD			A		ч , , , ч , , , , , , , , , , , , , , ,		A	A A	В	В	в	A	в	в	В	в	B	в	A A		в в	В	в	Ā	В	A A	в	-	Â	в
61 AFML_00751_00754_AA		A	В	3 /	A A	. 6	в	A B	A	В	в	В	в	A	в	-	A	-	A B		A B		в	в	в	B /	в	в	А	A
62 BDDA_00077_00080_AA		A	В	3 /	A A	. 6	в	A B	A	в	в	В	В		В	A	в	A	A B		A B	В	в	в	В	в	В	в	A	в
63 ATVI_00321_00324_AAD		B	В	· ·	A B		B	A A	A	B	B	A	B	A	A	A	B 	B	A A		- B	B	в	В	В	A A		B	В	В
64 AKQU_00747_00750_AA		B	В		A A		A	A B	B	B	A	B	в	A	в	A	в	A _	A B		A B	A	A	A	B	в		A	в	в
66 AOVW_00045_00066_AA		В	В		A A		A	A B	в	В	A	в	в		в	A	8	в	A B		A B	A			в	в			в	в
67 AOVW_00683_00688_AA		В	В		A A		A	A B		В	A	В			В	A	в	В	A B		A B	A			В	В			в	В
68 BODC_00337_00342_AA		В	8		A A		A A	A B	В	В	A	8	B		В	A	8	A	A B		A B	A			В	B			B	В
69 BODC_00347_00354_AA		B	B	· ·			A	A B	B	B	A	B	в	A	в	A	8	в	A B		A B	A	A	A	в	в		A	в	в
71 AUBH_00465_00468_AA		В	В	· ·	A A	, j	Ą	A B	В	В	A	в	в		в	A	в	в	A B		A B	A			в	в			в	в
72 BBCX_00493_00518_AAE	A C	В	В	, <b>,</b>	A A	ļ	A	A B		В	A	В			В	А	8	в	A B		A B	А			В	в			в	в
73 BDLF_00205_00208_AAE		В	B	<u> </u>	A A	1	A	A B	В	В	A	В	В		В	A	8		A B		A B	A			В	В /			В	В
74 BNEA_00235_00278_AAD 75 BNEA_00283_00290_AAD		В	B				Δ	A B	8	в	A 	B	B	A A	8 8	A	B	A A	A B		A B	A	A A	Α	B B	B		A	B	B
76 BNEA_00365_00370_AA		в	B		A A	í a l	A	A B	В	В	A	в	в	A	в	A	8	A	A B		A B	A	A	A	В	в	Â	A	в	в
77 BSGV_00447_00460_AA	D A	В	В	s <mark>/</mark>	A A		A	A B		В	A	В			В	A	8	A	A B		A B	A			в	в			в	В
78 BSFO_00315_00348_AAD		В	В	· ·	A A	4	A	A B	В	В	A	В	В		В	A	В	В	A B		A B	A	A	A	В	B /			В	В
79 AAIX_00251_00260_AAD		B	B		A A		A A	A B	A	B	A	B	В	A	В	A	8	B	A B		A B	A	B	A	B	A A	A	A	В	8
80 AAIX_00261_00264_AAD 81 ABLK_00311_00314_AAD		B	B				A	A B	A	B	A	B	в	A	в	A	8	-	A B		A B	A	B	A	в	A A		A	B	в
82 ADFG_00189_00192_AA		В	В		A A	, j	Ą	A B	A	В	A	в	в		в	A	в	A	A B		A B	A	в	A	в	A A			в	в
83 ANZV_00481_00500_AA	D A	В	В		A A	1	A	A B	A	В	A	В			В	А	В	В	A B		A B	А	А		В	В /			в	В
84 BCJO_00093_00096_AAD		В	B		а А		A	A B	A	В	A	В	В		В	-	8		A B		A B	A			В	A A			В	В
85 BCJO_00103_00106_AAD		В	В		A A		A •	A B	A	В	A	8	8	A	6	A	0	A	A B		A B	A	A	A	В	A	A	A	8	B 🗸
AAD.x_tree_	_clust / Sh	ieet1 /1	2/															j 4											_	
Ready												_															🖽 🗆 🛄 🔞	% —	-0	

## **Further processing the result files**

From this group info file we will extract marker IDs belonging to the linkage group 8 <u>LG8.tab</u> and create new locus file <u>LG8.loc</u> which contains marker scores for linkage group 4 only. It has raw scores for 247 markers. We <u>vlookup</u> function in Excel.

=VLOOKUP(\$Z5223,Sheet1!\$A\$2:\$DT\$5999,Sheet1!A\$1,FALSE)

#### **OPTIONAL ANALYSIS**

Now we will run Python\_MadMapper\_V248\_RECBIT\_012.py again on LG8.loc file with "**TRIO**" option:

#### \$python Python\_MadMapper\_V248\_RECBIT\_012.py LG8.loc LG8.out 0.2 100 25 X 0.33 50 TRIO 3

83 files will be generated. We are interested in the Marker Summary file: <u>LG8.out\_marker\_sum</u> which contains information about <u>loss of data</u> and <u>allele distortion</u> for each marker, as well as useful information derived from "TRIO" analysis. From this file we will extract marker IDs which have "GOOD\_\_MARKER" label (grade) only. It means we will use high-quality markers for further analysis. These markers have low fraction of data loss and meaningful ratio of "A"/"B" scores (allele distortion). There are 171 'good' markers in the dataset.

So, from here we are ready to jump to the **Part 2** to infer linear order of markers on linkage group 8.



**Pairwise matrices-** We will run the script with the following options/arguments <u>for each linkage group</u>

Prompt> python Python\_MadMapper\_V248\_RECBIT\_012.py LG\_8.loc LG\_8out 0.2 100 25 X 0.33 50 NOTRIO 3

See CheckMatrix analyses

# Prompt> python Python\_MadMapper\_V248\_RECBIT\_012.py

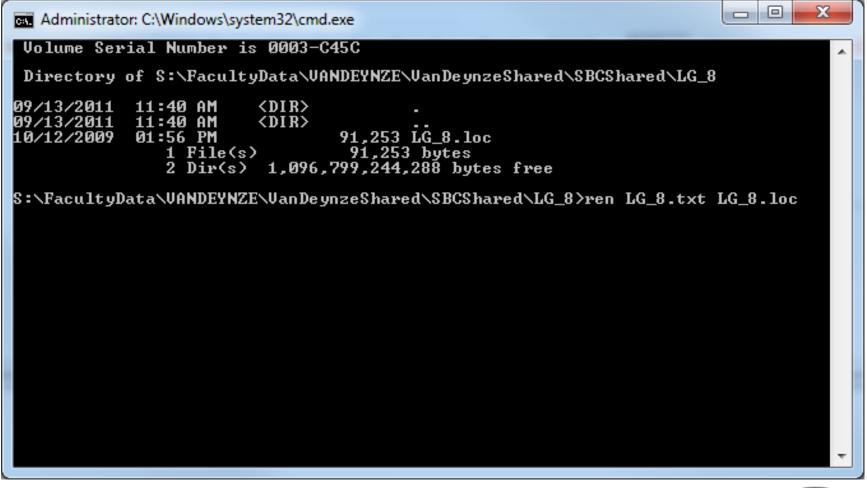
PROGRAM USAGE: MAD MAPPER TAKES 10 ARGUMENTS/OPTIONS IN THE FOLLOWING ORDER: (1)input\_file[LOC\_DATA/MARKER SCORES] (2)output\_file[NAME] (3)rec\_cut[0.2] (4)bit\_cut[100] (5)data\_cut[25] (6)group\_file[OPTIONAL] (7)allele\_dist[0.33] (8)missing\_data[50] (9)trio\_analysis[TRIO/NOTRIO] (10)double\_cross[3]



## Linkage group 8

Home Insert	Page Layout	t Formi	ulas Da	ata Revi	ew View	/ Acroba	it			LG_8 - Mici	O DITLACE											0 -	
K Cut	alibri	- 11 -		= = -	- Der	📑 Wrap Te	vt	General	•			Normal		Bad	Goo	d		- 式		Σ AutoSum		A	
Сору																	_		السي	🐺 Fill 🕋			
e 🛷 Format Painter 🛛 🖪	3 I U -		» - <u>A</u> -			Merge 8	Center *	\$ - %	• .00 →.0		nal Format 1g ∗ as Table	Neutral		Calculation	Che	ck Cell	Ins Ins	ert Delete	Format	🖉 Clear 👻	Sort & Filter ▼		
Clipboard 🕞	1	Font	Fa.		Alignm	ent	Fa	Numb	oer 🕼		-		Style	s				Cells			Editing		
A1 - 🕥	fx	name =	Flitered	SPPs																			
А		В	С	D	E	F	G	н	1	J	К	L	М	N	0	Р	Q	R	S	Т	U	V	
ame = Flitered_SPPs																							
opt = RI7																							
loc = 150																							
ind = 123																							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	5 1	.7	18 1	19	20	21
BIG_00379_00382_AAD		4	Α	В	Α	-	Α	В	Α	Α	Α	B A		B	Α	В	Α	Α	Α	В	В	Α	B
AFT_00447_00450_AAD				В	Α	Α	Α		Α	Α		B A			A	В	Α	Α	Α	В	В	Α	B
JMB_00771_00776_AAD				В	Α	Α	Α		Α	Α		B A			Α	В	Α	Α	Α	В	В	Α	B
MDI_00161_00164_AAE				B	A	-	A		Α	Α	A	B A			A .	B	A	A	A	B	B	A	B
MDI_00301_00340_AAE				B	A	A	A		A	A		B A			A	B	A	A	A	B	B	Α	B
PCM_00975_00978_AA				B	A	A	A A		A	A		B A			A	B	A	A	A	B	B	-	E
QBU_00427_00430_AAI				B	A	A	A		A A	A A		B A B A			A A			A	A	B	B	A	E
QBU_00553_00568_AAI				в	A	-	A		A	A		B A			А А	в	A	A	A	B	B	A	E
HCE_00333_00336_AAD				B	A	A	A		A	A		B A			A	B	A	A	A	B	B	A	B
HCE_00359_00362_AAD				B	A	-	A		A	A		B A			A	B	A	A	A	B	B	A	B
HCE 00597 00600 AAD				В	A	A	A		A	A		B A			A	В	A	A	Ā	В	в	Ā	F
HCE 00601 00614 AAD				В	A	A	A		A	A		B A			A	В	A	A	A	B	B	A	E
HZD_00177_00180_AAD				В	A	A	A		A	A		B A			A	В	A	A	A	в	B	A	B
LMD 00389 00392 AAD				в	A	-	A		A	A	Α	B A			Α	в	A	А	Α	в	в	А	В
MCU_00073_00078_AA				в	Α	Α	Α		Α	Α	Α	B A			Α	в	Α	Α	Α	в	в	Α	E
QAS_00047_00050_AAE	D E	3	A	в	Α	A	A	В	A	A	А	B A		В	А	в	A	А	A	в	в	A	B
SBC_00389_00406_AAD	) E	3	A	в	Α	Α	Α	В	Α	Α	Α	B A		В	А	в	Α	Α	Α	в	в	Α	B
QMI_00249_00252_AAE	D -		Α	В	Α	-	Α	В	Α	Α	Α	B A		В	А	В	Α	А	Α	в	в	Α	B
HXI_00395_00418_AAD	) 4	4	Α	Α	Α	Α	Α	В	A	A	Α	B A		В	А	В	Α	Α	Α	В	В	Α	B
OTA_00379_00388_AA	D A	4	Α	Α	Α	Α	Α	В	Α	Α	Α	B A		В	А	В	Α	Α	Α	В	в	Α	B
OTA_00389_00400_AA	D /	4	Α	Α	Α	Α	Α	В	Α	Α	Α	B A		B	А	В	Α	Α	Α	В	В	Α	B
MEH_00353_00356_AAI				Α	Α	-	Α		Α	Α	Α	B A			А	В	Α	Α	Α	В	В	Α	B
HYQ_00053_00056_AA				В	Α	Α	Α		Α	Α	Α	B A		-	В	В	В	В	Α	Α	В	Α	B
JTV_00219_00222_AAD				В	Α	-	Α		Α	A		B A			В	В	В	В	Α	A	В	Α	B
KEP_00241_00244_AAD				B	A	-	A		A	A		B A			B	B	B	B	A	A	B	A	B
KLS_00665_00670_AAD				B	A	A	A		A	A		B A			B	B	B	B	A	A	B	A	B
LTN_00331_00334_AAD				B	A	A -	A A		A A	A A		B A			B	B	B	B	A	A	B	A	B
LUB_00133_00136_AAD				B	A		A		A	A					B B	B	B	B	A	A	B	A	B
CBB 00117 00120 AAE				в	A	A A	A		A	A		B A B A			в В	в	-	В	A	A	B	A	B
CBB_00117_00120_AAL				В	A	A	A		A	A		B A			B	В	-	B	A	A	B	A	B
NFK_00645_00648_AAD				B	A	A	A		A	A		B A		-	B	B	B	B	A	A	B	A	B
SGS 00527 00530 AAD				B	A	-	A		A	A		B A			B	B	B	B	A	A	B	A	E
SVS_00215_00222_AAD				В	A	A	A		A	A		B A			В	В	-	В	A	A	В	Ā	E
TOT 00141 00148 AAE				В	A		A		A	A		B A			В	В	в	В	A	A	В	A	E
TOT_00981_01000_AAD				В	A	A	A		A	A		B A			В	В	В	В	A	A	B	A	E
VOE_00201_00208_AAD				В	A	A	A		A	A	A	B A			В	В	в	в	A	A	В	A	E
VOE_00209_00212_AAD				В	A	-	A		A	A		B A			В	В	В	В	A	A	В	A	E
DFV_02233_02236_AAE	D A	4	A	в	Α	-	A	в	A	A	Α	B A		В	в	в	-	в	Α	A	в	Α	E
		4	A	в	A	Α	A	В	A	A	A	B A		В	В	в	в	в	A	A	в	А	E
YNL_00227_00230_AAD																							

## Rename txt file to loc file in command line of DOS





### Inference of linear order of markers on linkage groups using WinRecord

#### <u>STEP 2</u>

We use Record software to order the markers

http://www.plantbreeding.wur.nl/UK/software record.html

<b>W</b>		NUR			Ь	og in 💻 🕻	*		Search	Links
	AGENINGE For qua	lity of life					This Site	 Search	Advance	d Search
Education	Research	Publications	News & Calendar	About Wageningen UR	Work at	Phone book	Contact			



Plant Breeding
Education
Research
Publications
News & Calendar
About Plant Breeding
Work at
Phone book
Links
Contact

The program RECORD (REcombination Counting and ORDering) can be used for the ordering of loci on genetic linkage maps. The method minimises the total number of recombination events. The search algorithm is a heuristic procedure. combining elements of branch-and-bound with local reshuffling

Since the criterion does not require intensive calculations, the algorithm rapidly produces an optimal ordering as well as a series of near-optimal ones. The latter provides insight into the local certainty of ordering along the map.

RECORD can deal with the following types of mapping populations: BC1, F2, F3, RILs (in fact any generation obtained by repeated selfing of a hybrid between homozygous parents). Mapping populations from non-inbreds should be split into BC1 or HAP data that represent the maternal and paternal gametes, according to the two-way pseudo-testcross method. Please provide the data in a .LOC file in JoinMap format.

Quotes:

- Having established the superior performance of RECORD, we used a combination of RECORD and a purpose-built perl script to construct pilot maps of individual populations. Wenzl et al. (2006) A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics 7:206 doi:10.1186/1471-2164-7-206
- The number of crossovers forms the basis for RECORD since the algorithms minimise the total number of recombination events (Van Os et al. 2005). Total map length has been used to evaluate different marker orders, and generally the shorter the map (produced from the datasets under investigation) the more accurate the marker order.
- Comparison of SALOD with number of crossovers between true, MMQTX, and RECORD maps clearly indicated that the minimum number of crossovers is a far superior determinant of marker order accuracy compared with SALOD Collard et al (2009) How accurate are marker orders in crop linkage maps. Crop & Pasture Science 60(4)362-372. doi:10.1071/CP08099
- The order of markers as given by RECORD is better than the order of markers as given by traditional linkage mapping software programs ... Varshney et al (2007) A high density barley microsatellite consensus map with 775 SSR loci. Theor Appl Genet 114(6):1091-1103. doi:10.1007/s00122-007-0503-7
- Concepts from RECORD and SMOOTH have inspired the development of MSTMAP. Tests show that MSTMAP consistently outperforms the best available methods in the literature. This is because of the integration of a mapping algorithm with detection and removal of noise from the dataset. However, when the error detection feature is turned off and the dataset is noisy, the authors conclude:
  - "CARTHAGENE appears to be better than RECORD when the data are clean (g =0). When the data are noisy, RECORD constructs more accurate maps than CARTHAGENE. MSTMAP and RECORD are both very efficient in terms of running time, and they are much faster than CARTHAGENE. Wu et al (2008) Efficient and Accurate Construction of Genetic Linkage Maps from the Minimum Spanning Tree of a Graph. PLoS Genet 4(10):e1000212. doi:10.1371/journal.pgen.1000212
- Cheema and Dicks (2009) Approaches and software tools for genetic linkage map estimation in plants. Briefings in Bioinformatics 2009 10(6):595-608; doi:10.1093/bib/bbp045

For questions, contact Herman van Eck (herman vaneck AT wur.nl).

Download RECORD

Download RECORD For Windows

last update: RF22/11/2009



#### Welcome to RECORD

#### Tue, 10 Apr 2012, 12:21

Freeware distributed by the Laboratory of PLant Breeding of Wageningen University

Copyright Piet Stam 2006

What does RECORD do?	Fast mapping of markers on a single linkage group
What population types can RECORD deal with?	BC1, F2 and RILs of any generation (selfed)
Which object function is used by RECORD?	RECORD minimises the total number of recombination events in the data

RECORD can be used as an additional tool to the JoinMap package, distributed by Kyazma company, Wageningen (www. kyazama.nl)

#### Input:

Raw genotypic data in JoinMap format

#### Output:

1. Best order of markers and their positions on the map. (Duplicate markers are removed before calculation starts.)

2. Semi-graphical output that displayss a set of 'plausible' alternative maps

Output is stored in files that reside in a 'project directory', to be created by the user at run time

Reference: In publications that are (partially) based on this software refer to:

Os, van, H, P Stam, R G F Visser and H J van Eck (2005) : RECORD, a novel method for ordering loci on a genetic linkage map Theoretical and Applied Genetics 112:30-40

Click to continue



# RECORD 2.0												x
	ons EXIT A	About	-	-	-	-	-	-	-	-	-	
New Project												
🔁 Load Data												1
Exit												
		JMORD Lie	censed to Lab	oratory of P	lant Breedin	ig by Piet S	tam					

# RECORD 2.0												- • ×
<u>File Calculate Opti</u>	ons <u>E</u> XIT <u>A</u> bout											
	Loc File											
	name = Flitered_SPPs										*	
LOAD DATA	popt = RI7											
	nloc = 150											
	nind = 123										=	
	; 1 2 3	4	5	6	7	8	9	10	11	12	:	
	ABIG_00379_00382_AAD A	A	в	A	-	A	в	A	A	A	1	
	AAFT_00447_00450_AAD -	A	в	A	A	A	в	A	A	A	1	
	AJMB_00771_00776_AAD A	A	В	A	A	A	В	A	A	A	1	
	AMDI_00161_00164_AAD -	A	B	A	-	A	B	A	A	A A	1	
	AMDI_00301_00340_AAD A	A A	В	A A	A A	A A	В	A A	A A	A		
	APCM_00975_00978_AAD A AQBU_00427_00430_AAD A	A	В	A	A	A	В	A	A	A	1	
	AQBU_00553_00568_AAD -	A	в	A	A	A	в	A	A	A	1	
	BHCE_00201_00204_AAD -	A	в	A	-	A	в	A	A	A	i	
	BHCE_00333_00336_AAD -	A	в	A	А	A	в	A	A	A	i	
	BHCE_00359_00362_AAD -	A	в	A	-	A	в	A	A	A	i	
	BHCE_00597_00600_AAD -	A	в	A	А	А	в	A	А	А	1	
	BHCE_00601_00614_AAD -	A	в	A	A	A	в	А	А	A	1	
	BHZD_00177_00180_AAD -	A	в	A	A	A	в	A	А	A	1	
	BLMD_00389_00392_AAD A	A	в	A	-	A	в	A	А	A	1	
	BMCU_00073_00078_AAD A	A	в	A	A	A	в	A	A	A	1	
	BQAS_00047_00050_AAD B	A	в	A	A	A	в	A	A	A	1	
	BSBC_00389_00406_AAD B	A	в	A	A	A	в	A	A	A	1	
	BQMI_00249_00252_AAD -	A	в	A	-	A	в	A	A	A	1	
	AHXI_00395_00418_AAD A	A	A	A	A	A	в	A	A	A	1	
	AOTA_00379_00388_AAD A	A	A	A	A	A	в	A	A	A	1	
	AOTA_00389_00400_AAD A	A	A	A	A	A	В	A	A	A	1	
	BMEH_00353_00356_AAD A	A	A	A	-	A	В	A	A	A	1	
	AHYQ_00053_00056_AAD -	A -	В	A -	A	A -	В	A -	A -	A -	1 +	
	<										- F	

JMORD Licensed to Laboratory of Plant Breeding by Piet Stam



															- 0 <b>X</b>
<u>File</u> <u>Calculate</u>	Options	<u>e</u> xit	About												
D h 🛠 🔳															
		oc File													1
Options		popt nloc	e = Flitered ; = RI7 ; = 150 d = 123	d_SPPs											
		;		2 3	4	5	6	7	8	9	10	11	12	-	
			00379_003		A	в	A	-	A	В	A	A	A		
			00447_004		A A	B	A A	A A	A A	B	A A	A A	A A	1	
			3_00771_007 [_00161_001		A	В	A	-	A	В	A	A	A	1	
			00301_003		A	в	A	A	A	в	A	A	A		
			00975_009		A	в	A	A	A	в	A	A	A	i	
			00427_004		A	в	A	A	A	в	A	A	A	i	
			00553_005		A	в	A	A	A	в	A	A	A	1	
			00201_002		A	в	А	-	A	в	A	A	А	1	
			00333_003		А	в	А	А	A	в	А	A	А	1	
			00359_003		A	в	A	-	A	в	A	A	А	1	
			00597_006		A	в	A	A	A	в	A	А	А	1	
			00601_006		A	в	А	А	A	в	A	A	А	1	
			00177_001		A	в	А	А	A	в	A	A	А	1	
			00389_003		A	в	A	-	A	в	A	A	A	1	
		BMCU	J_00073_000'	78 AAD A	A	в	A	A	A	в	A	A	A	1	
		BQAS	00047_000	50_AAD B	A	в	A	A	A	в	A	A	A	1	
			00389_004		A	в	A	A	A	в	A	A	A	1	
			_00249_002		A	в	A	-	A	в	A	A	A	1	
		AHXI	_00395_004	18_AAD A	A	A	A	A	A	в	A	A	A	1	
			00379_003		A	A	A	A	A	в	A	A	A	1	
		AOTA	_00389_004	00_AAD A	A	A	A	A	A	в	A	A	A	1	
		BMEH	I_00353_003	56_AAD A	A	A	A	-	A	в	A	A	A	1	
		AHYQ	2_00053_000	56_AAD -	A	в	A	A	A	в	A	A	A	1 👳	
		•				-	-			-	-			- F	

JMORD Licensed to Laboratory of Plant Breeding by Piet Stam



## **Options in WinRecord**

	eeking plausible alternatives (fraction of total) 0.10	(A number between 0 and 1) (A number >1)
-Mapping function	Critical Gap Size	
O Haldane	Gap size (cM) 25	
Kosambi	(Gaps larger than this size result in multiple linka	ge groups)
	οκ	
		Sol

# RECORD 2.0												- • ×
File Calculate Options EXIT About												
Calculato name = Flite	ered_SPPs											
Calculate popt = RI7												
nloc = 150 nind = 123											-	
; 1	2 3	4	5	6	7	8	9	10	11	12	. =	
ABIG_00379_0		A	в	A	_	A	в	A	A	A	i	
AAFT_00447_0		A	в	A	А	A	в	A	A	A	1	
AJMB_00771_0		A	в	A	A	A	в	A	A	А	1	
AMDI_00161_0		A	в	A	-	A	в	A	A	A	1	
AMDI_00301_0	00340_AAD A	A	в	A	A	A	в	A	A	A	1	
APCM_00975_0		A	в	A	A	A	в	A	A	A	1	
AQBU_00427_0		A	в	A	A	A	в	A	A	A	1	
AQBU_00553_0		A	в	A	A	A	в	A	A	A	1	
BHCE_00201_0		A	в	A	-	A	в	A	A	A	1	
BHCE_00333_0		A	в	A	A	A	в	A	A	A	1	
BHCE_00359_0		A	в	A	-	A	В	A	A	A	1	
BHCE_00597_0		A	B	A A	A A	A A	B	A A	A A	A A	1	
BHCE_00601_( BHZD_00177_(		A A	В	A	A	A	В	A	A	A	1	
BLMD_00389_0		A	В	A	-	A	В	A	A	A	1	
BMCU_00073_0		A	в	A	А	A	в	A	A	A	i	
BQAS_00047_0		A	в	A	A	A	в	A	A	A	i	
BSBC_00389_0		A	в	A	A	A	в	A	A	A	1	
BQMI_00249_0		А	в	A	-	A	в	A	А	А	1	
AHXI_00395_0		A	A	A	A	A	в	A	А	A	1	
AOTA_00379_0		A	A	A	A	A	в	A	A	A	1	
AOTA_00389_0	00400_AAD A	A	A	A	A	A	в	A	A	A	1	
BMEH_00353_0	00356_AAD A	A	A	A	-	A	в	A	A	A	1	
AHYQ_00053_0	00056_AAD -	A	в	A	A	A	в	A	A	A	1 👳	
٠ III		-	-				_		-	-	•	
,												

JMORD Licensed to Laboratory of Plant Breeding by Piet Stam



# RECORD 2.0	
<u>File Calculate Options EXIT About</u>	
Loc File Log File	
log file for JMORD	<u> </u>
Tue, 10 Apr 2012, 12:23	
raw data read from \\plantsciences\dfs\FacultyData\VANDEYNZE\VanDeynze\Education\Breedin	a
	-
Parameter settings for this run:	
Gap Size : 25.0 cM	
Mapping function : Kosambi	
Fraction extra recombinations allowed: 0.1000	
Nr. of steps neeeded : 5	
Bins: (markers on a single line have identical segregation and belong to a bin)	
ABIG_00379_00382_AAD	
AAFT_00447_00450_AAD	
AJMB_00771_00776_AAD	
AMDI_00161_00164_AAD	
AMDI_00301_00340_AAD	
APCM_00975_00978_AAD	
AQBU_00427_00430_AAD	
AQBU_00553_00568_AAD BHCE 00201 00204 AAD	
BHCE_00333_00336_AAD	
BHCE_00359_00362 AAD	
BHCE_00597_00600_AAD	
BHCE_00601_00614_AAD	-
	•

# RECORD 2.0		
File Calculate Options	<u>EXIT A</u> b	out
	File Log F	
LOCT	lie cogi	
		523_00540_AAD
1	APND_00:	323_00334_AAD
	marker 1	names after removal of duplicates:
	1 2	ABIG_00379_00382_AAD
	2 2	AAFT_00447_00450_AAD
		AJMB_00771_00776_AAD
		AMDI_00161_00164_AAD
		AMDI_00301_00340_AAD
		APCM_00975_00978_AAD
		AQBU_00427_00430_AAD
		AQBU_00553_00568_AAD
		BHCE_00201_00204_AAD
		BHCE_00333_00336_AAD
		BHCE_00359_00362_AAD
		BHCE_00597_00600_AAD



# RECORD 2.0	
File Calculate Option	ns <u>E</u> XIT <u>A</u> bout
	Loc File Log File
	<pre>rep nr. = 1 TestCoreValue = 176.9798</pre>
	rep nr. = 2 TestCoreValue = 176.9792
	rep nr. = 3 TestCoreValue = 176.9798
	rep nr. = 4 TestCoreValue = 162.5153
	rep nr. = 5 TestCoreValue = 176.9792
	rep nr. = 6 TestCoreValue = 168.7238
	rep nr. = 7 TestCoreValue = 170.7707
	rep nr. = 8 TestCoreValue = 170.8015 best sequence after 8 replicates:
	0 APND 00323 00334 AAD
	•         •
_	
READY	JMORD Licensed to Laboratory of Plant Breeding by Piet Stam
READY	JMORD Licensed to Laboratory of Plant Breeding by Piet Stam

Sol CAP

jmord - Notepad		×
<u>F</u> ile <u>E</u> dit F <u>o</u> rmat <u>V</u> iew <u>H</u> elp		
best sequence after 8 replicates:		
0 APND_00323_00334_AAD		
1 ACCM_00365_00374_AAD 2 BGNG_00275_00282_AAD		
3 AEBW_00523_00540_AAD 4 ACCM_00375_00382_AAD		
5 BGNG_00261_00274_AAD 6 BQHV_00391_00414_AAD	Order	
7 ABW0_00273_00276_AAD		
8 BQHV_00415_00420_AAD 9 ABW0_00067_00070_AAD		_
10 AIPG_00121_00124_AAD 11 BDBZ_00121_00142_AAD		1
12 AONS_00509_00512_AAD 13 AFWK_00477_00480_AAD		
13 APWR_00477_00480_AAD 14 AQRR_00085_00098_AAD 15 AQRR_00377_00384_AAD		
141 AOTA_00389_00400_AAD 142 BMCU_00073_00078_AAD 143 APCM_00975_00978_AAD 144 AMDI_00161_00164_AAD 145 BQMI_00249_00252_AAD 146 BHCE_00201_00204_AAD 147 BHZD_00177_00180_AAD		
148 BSBC_00389_00406_AAD 149 BIFE_00379_00382_AAD		=
count = 162.5153		-
map in JoinMap format (up & down)	Distance	
group UP APND_00323_00334_AAD 0.000		
ACCM_00365_00374_AAD 0.417		
AEBW_00523_00540_AAD 0.417		
ACCM_00375_00382_AAD 0.417 BGNG_00261_00274_AAD 0.418		10
BQHV_00391_00414_AAD 3.461 ABW0_00273_00276_AAD 3.461		
0011/ 00415 00420 ++D 2 464		

## DOS version the same but faster (No distances though)

S:\FacultyData\VANDEYNZE\VanDeynzeShared\SBCSh	ared\LG_8\ord.exe	- 0 X
		A
***** ORD: Quick Ordering	<del></del>	
**** by counting rec. events	****	
****	****	
**** copy right P. Stam 1999	****	
**** Laboratory of Plant Breeding	XXXX	
**** Dept. Plant Sciences	****	
**** Wageningen Agricultural University	××××	
**** PO Box 386	XXXX	
**** 6700 AJ Wageningen	XXXX	
**** The Netherlands	****	
	*****	
Tue, 13 Sep 2011, 12:58	1 (default out in lea) IC 9 la	
raw data (.loc) file? [path] name [.ext output (.log) file? [path] name [.ext		C
output file LG_8.log already exists; o		у
ripple interval? ( $0 = no ripples$ ) 5	voiwiitee, (yyes / (nye / (q/ait	3
EQU-threshold? (for (almost) equivalent	orders) 5	
Preparing Bins		
Estimating recombination frequencies		
rep 1 of 10		
rep 2 of 10		
rep 3 of 10		
 rep 4 of 10		
· · · · · · · · · · · · · · · · · · ·		

#### Checkmatrix: Requirements mad

CheckMatrix py\_matrix\_2D\_V248\_RECBIT.py takes as input three files:

#### 1. Pairwise Distance Matrix File madmapper test small.out.pairs all pairwise distance matrix file can be generated by Python MadMapper from Locus File

GM01	GM07	0.36
GM01	GM08	0.40
GM01	GM09	0.48
GM01	GM10	0.52
GM01	GM11	0.60
GM01	GM12	0.68
GM02	GM01	0.04
GM02	GM02	0.00
GM02	GM03	0.08
GM02	GM04	0.16
GM02	GM05	0.20
GM02	GM06	0.24

Convert 🕀

× Find:

Q Complete gran

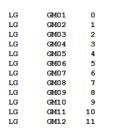
## Pairwise Distance matrices from MadMapper

- Alexander Kozik, UC Davis Genome Center
- http://cgpdb.ucdavis.edu/XLinkage/MadMapper/

ucdavis.edu

#### 2. Genetic Map File madmapper test small.map.right

(on this example last column reflects the order markers)



#### 3. Locus File (Raw Marker Scores) madmapper test small.loc

	1									10	ו									20					25	
																									1	
GM01	А	А	А	А	А	А	А	А	А	А	А	А	А	A	А	А	в	в	в	в	в	в	в	в	в	
GM02	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	в	в	в	в	в	в	в	в	в	в	
GM03	А	А	А	А	Α	А	А	А	А	А	А	А	Α	в	в	в	в	в	в	в	в	в	в	в	в	
GM04	А	А	А	А	А	А	А	А	А	А	А	в	в	в	в	в	в	в	в	в	в	в	в	в	в	
GM05	А	А	A	А	A	А	A	А	A	А	в	в	в	в	в	в	в	в	в	в	в	в	в	в	в	
GM06	Α	А	А	А	Α	А	А	А	А	в	в	в	в	в	в	в	в	в	в	в	в	в	в	в	в	
GM07	А	А	А	А	А	А	А	А	А	в	в	в	в	в	в	в	в	в	в	в	в	в	в	А	Α	
GM08	А	А	A	А	A	А	A	А	A	в	в	в	в	в	в	в	в	в	в	в	в	в	A	A	Α	
GM09	Α	А	А	А	Α	А	А	А	А	в	в	в	в	в	в	в	в	в	в	в	А	А	А	Α	Α	
GM10	в	А	А	А	A	А	А	А	А	А	в	в	в	в	в	в	в	в	в	А	Α	А	А	A	Α	
GM11	в	в	Α	А	Α	Α	А	А	А	А	в	в	в	в	в	в	в	в	А	А	Α	А	А	Α	Α	
GM12	в	в	в	А	A	A	A	А	A	А	в	в	в	в	в	в	в	А	А	A	A	А	A	A	Α	

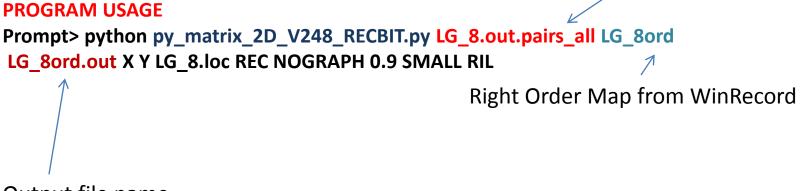
### From Record (reformatted)

Loc File (input for Record)

## Visualization and validation of genetic maps using 2dimensional heat plots using Check Matrix program

http://cgpdb.ucdavis.edu/XLinkage/Genetic\_Map\_Matrix\_Plot\_Art.html

Run from MadMapper on LG\_8 to get this file



Output file name

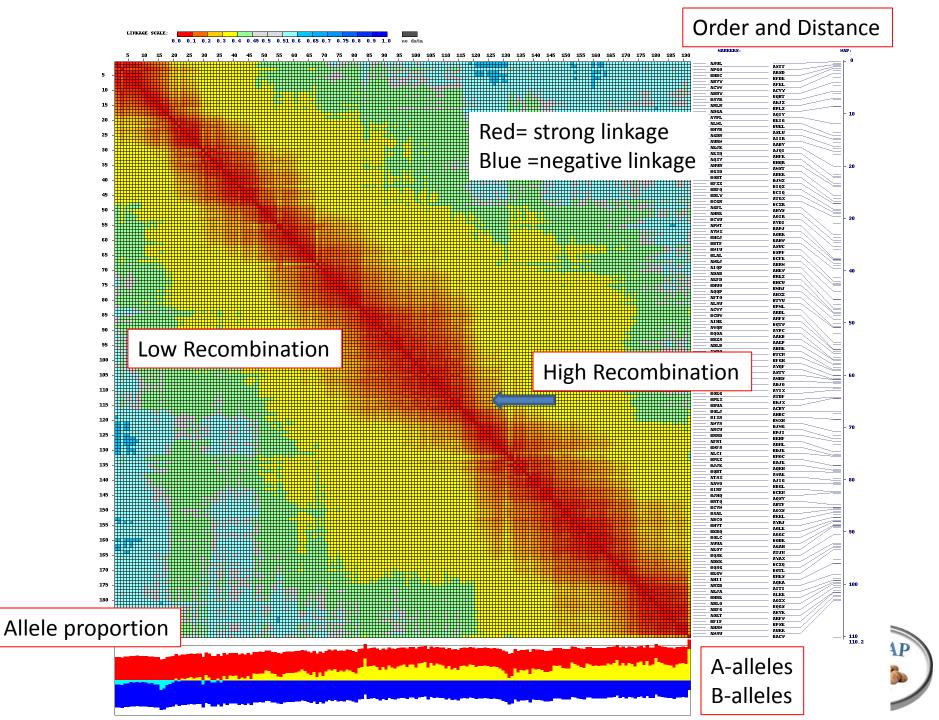
py matrix 2D V248 RECBIT.py takes as input matrix file, map file and three optional files: framework markers list, list of IDs to highlight in red and \*.loc file with recombination data (raw marker scores).

[matrix\_file] [map\_file] [output\_file] [frame\_marker\_list] [red\_list] [loc\_file] Sol CAP [REC/BIT/LOD] [GRAPH\_OPT] [LINK\_CUT] [LARGE/SMALL] [Popn\_TYPE]

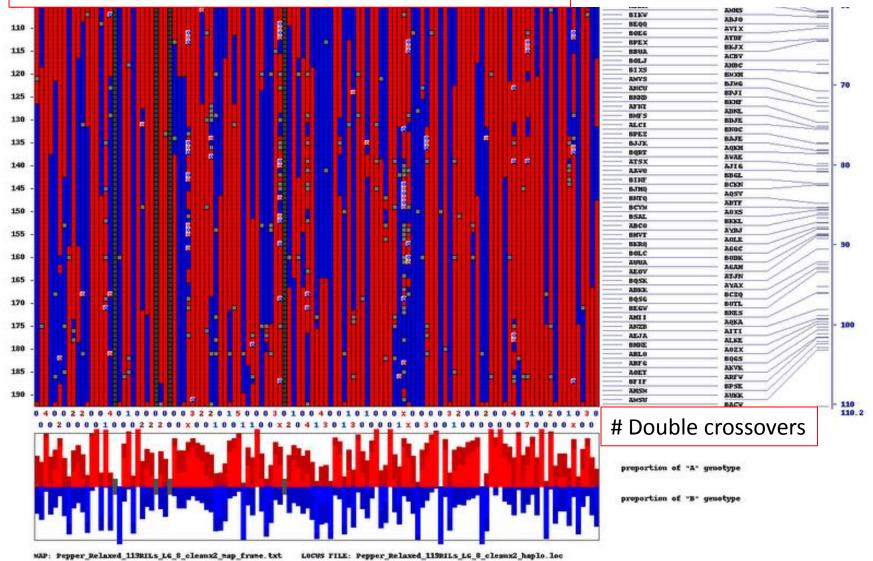
## Transfer <u>PNG</u> files from Unix to windows or Mac OS system.

• WinSCP connects the unix server to windows environment. http://winscp.net/eng/download.php

		<mark>] home                                   </mark>
Image: Second secon		<mark>] home                                   </mark>
My documents     Image: Second		<mark>] home                                   </mark>
Users\ashrafi\Documents		/home
	_	
Ali Pey		
Ali Pey		Name Ext
		불
App Powel Tomato Project	Ξ	🔰 ashrafi
Ann owei_romato_roject		🕌 sreyesch
Ava		iii msolanor
Books		🕛 zwłuxx
Borevitz_files		🕌 syarnes
Bowtie Commands		🔑 yfy
Breeding With Molecular Markers		\mu jfass
Cotton		🔑 thill
Courses		퉬 akozik
Data		📙 basickler
Downloaded Software		🕌 dlin
Downloads		
Emploee		
EndNote_Library		
Farzaneh		
Fax		
GC		
How to merge many excel sheets		
IntronFinder		
JobApplications		
JoinMap4		
KBiosciences		
Kevin		
LabVIEW Data		
LB_Heather Slide	-	
		۲ III کې ا
of 403 MiB in 0 of 179		0 B of 0 B in 0 of 11
🖻 405 MiB in 0 01 175 P F2 Rename 📝 F4 Edit 📸 F5 Copy 📸 F6 Move 💣 F7 Create Directory 🗙 I		
rz Rename 🔄 ra Edit 🖄 PS Copy 🖄 Po Move 🎯 P/ Create Directory 🗙	F8 L	
		🔒 SFTP-3 🗐 0:00:19



## Double crosser- White, Grey Missing data





Wu et al 2009 COSII F2 map "Pseudolinkage of loci near breakpoints in reciprocal translocation."

Markers linked to P1-wild have weak linkage to P8wild

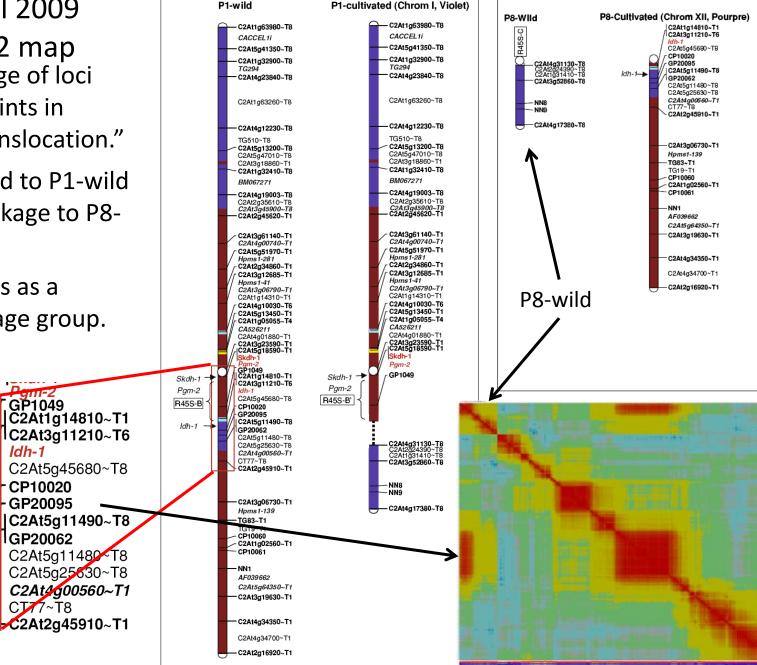
P8-wild stands as a separate linkage group.

Skdh-1

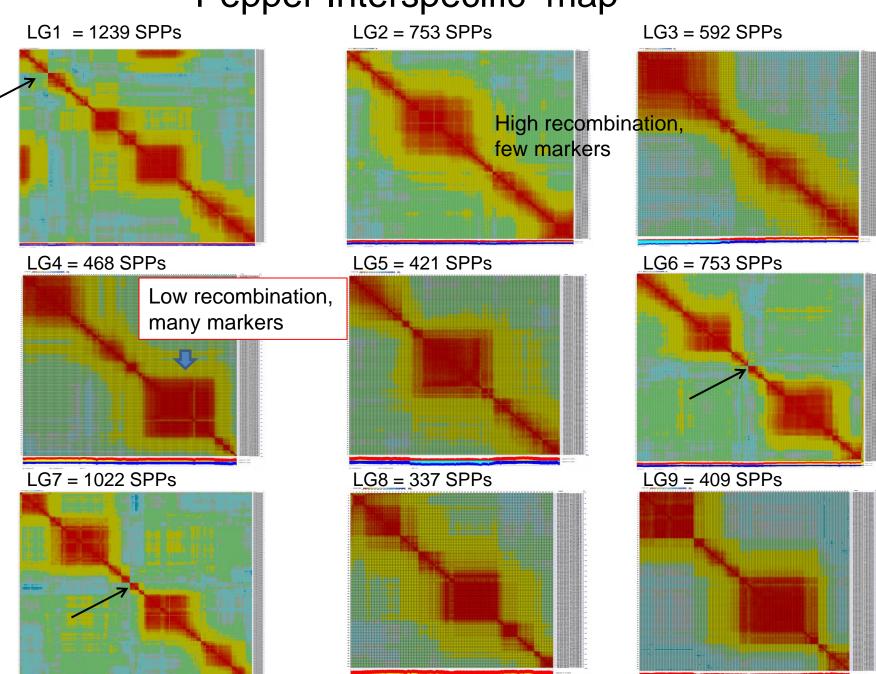
Pam-2

R45S-B

Idh-1



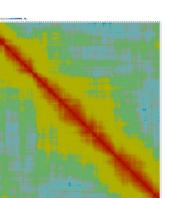
## Pepper Interspecific map

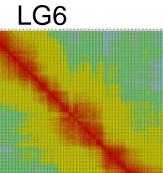


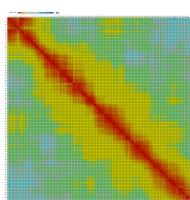


LG5

## LG2



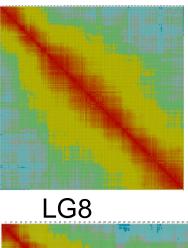


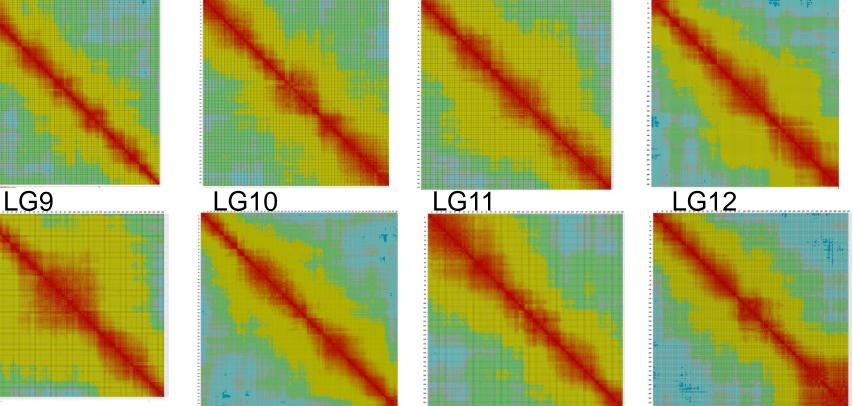


LG3

LG7







C. Frutescens x C. annuum. 12 Linkage Groups / 12 Chromosomes, 2,886 bins

# Summary

- Group with MadMapper
- Define each linkage groups-Vlookup
- Create paired matrices file
- Determine order with RECORD
- Visualize with Checkmatrix



# Thank you



Please fill out the survey evaluation! (You will be contacted via email)

Today's presentation available at: http://www.extension.org/pages/63330

> Sign up for PBG News http://pbgworks.org

Watch past webinars and sign up for future webinars!

http://www.extension.org/pages/60426

