

LD Adjacent Pairs Analysis

Author: Greta Linse Peterson, Golden Helix, Inc.

Overview

LD Analysis is performed on all adjacent pairs within a chromosome (if a marker map is applied) or within a haplotype block. The results contain values for both the EM and CHM methods and both R^2 and D' values.

This script can be run on a spreadsheet that includes both genotype and phenotype data, with or without a marker map. It can also be run on a subset spreadsheet created from an LD plot by selecting **Subset Markers**. You can also run this script on markers in haplotype blocks as specified in a Haplotype Block Spreadsheet.

Recommended Directory Location

Save the script to the following directory:

*..\Application Data\Golden Helix SVS\UserScripts\Spreadsheet\Analysis\

Note: The **Application Data** folder is a hidden folder on Windows operating systems and its location varies between XP and Vista. The easiest way to locate this directory on your computer is to open SVS and select the **Tools >Open Folder > UserScripts Folder** menu option. If saved to the proper folder, this script will be accessible from the spreadsheet **Analysis** menu.

Using the Script

- 1) Method One:
 - a) Open a spreadsheet containing genotype data (possibly including phenotype data and may or may not be marker mapped).

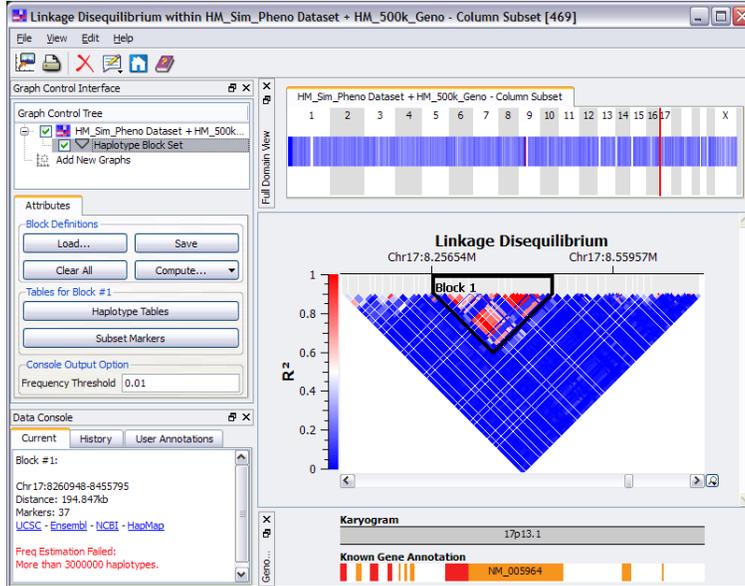
Map	Sample ID	Height (In)	Age	SNP_A-1909444	SNP_A-2237149	SNP_A-4303947	SNP_A-1886933	SNP_A-2236359
Chromosome				1	1	1	1	1
Position				742429	767376	769185	775852	782343
Cytoband				p36.33	p36.33	p36.33	p36.33	p36.33
1	NA12813	60.27476	55.48754	B_B	B_B	A_A	B_B	B_B
2	NA12740	67.06439	54.33223	A_B	B_B	A_A	B_B	B_B
3	NA19193	65.06428	29.11164	A_A	B_B	B_B	A_A	B_B
4	NA18501	66.4639	46.90015	A_B	B_B	A_B	A_B	A_B
5	NA12874	55.6836	34.07425	B_B	B_B	A_A	B_B	B_B
6	NA19154	70.79579	31.26026	A_A	B_B	A_B	A_A	A_B
7	NA19129	71.88845	32.83759	A_A	B_B	A_B	A_A	A_B
8	NA12761	74.85801	50.45818	B_B	B_B	A_A	B_B	B_B
9	NA18537	68.62109	42.67434	A_B	B_B	A_B	A_B	B_B
10	NA19201	60.72689	52.96835	A_A	B_B	B_B	A_A	B_B
11	NA12760	63.3528	23.07136	B_B	B_B	A_A	B_B	B_B
12	NA18998	65.56057	23.33213	B_B	B_B	A_A	B_B	B_B
13	NA19132	71.94062	29.21608	A_A	B_B	B_B	A_A	B_B
14	NA06985	67.98739	27.66916	B_B	?	A_A	B_B	B_B

- b) Select **Analysis > LD Adjacent pairs Analysis**.

c) Select **No** when asked if using a haplotype block spreadsheet.

2) Method Two:

a) Open a Plot viewer containing an LD plot. Select a block of markers, and under the Haplotype Block Set attributes, select Subset Markers.



b) This creates a subset spreadsheet with these markers.

Map	Sample ID	G 1	G 2	G 3	G 4	G 5	G 6	G
Chromosome		17	17	17	17	17	17	17
Position		8260948	8264129	8270690	8273323	8285267	8291756	8316
Cytoband		p13.1	p13.1	p13.1	p13.1	p13.1	p13.1	p13
1	NA12813	B_B	B_B	A_A	A_B	B_B	B_B	B_B
2	NA12740	?_?	B_B	A_A	A_B	B_B	B_B	B_B
3	NA19193	B_B	B_B	B_B	A_A	B_B	B_B	B_B
4	NA18501	A_B	A_B	B_B	A_A	B_B	B_B	B_B
5	NA12874	B_B	B_B	A_A	A_B	B_B	B_B	B_B
6	NA19154	B_B	A_B	B_B	A_A	B_B	B_B	B_B
7	NA19129	B_B	A_B	B_B	A_A	A_B	A_B	A_B
8	NA12761	A_B	A_B	B_B	A_B	B_B	B_B	B_B
9	NA18537	B_B	B_B	A_A	A_B	B_B	B_B	B_B
10	NA19201	A_B	A_B	A_B	A_B	B_B	B_B	B_B
11	NA12760	B_B	B_B	B_B	A_A	B_B	B_B	B_B

c) Select **Analysis > LD Adjacent pairs Analysis**.

d) Select **No** when asked if using a haplotype block spreadsheet.

3) Method Three:

a) Open a spreadsheet containing genotype data (possibly including phenotype data and may or may not be marker mapped).

b) Select **Analysis > LD Adjacent pairs Analysis**.

- c) Select **Yes** when asked if using a haplotype block spreadsheet.
- d) Select the haplotype block spreadsheet.

Map	Markers	Block #
1	SNP_A-1865924	1
2	SNP_A-2041478	1
3	SNP_A-4274931	2
4	SNP_A-2285039	2
5	SNP_A-1941521	3
6	SNP_A-1860613	3
7	SNP_A-1817285	4
8	SNP_A-1791886	4
9	SNP_A-1938509	5
10	SNP_A-1904645	5
11	SNP_A-1802724	5
12	SNP_A-4174969	5
13	SNP_A-2005048	5
14	SNP_A-2064632	6
15	SNP_A-2241518	6
16	SNP_A-2052894	7
17	SNP_A-1820470	7
18	SNP_A-4281588	7
19	SNP_A-1800897	8
20	SNP_A-2109637	8
21	SNP_A-2245307	9
22	SNP_A-2178502	9
23	SNP_A-1944345	10
24	SNP_A-4196448	10

The result is a new spreadsheet with the row labels and first column listing the markers that were compared (all adjacent pairs comparisons, only listing one occurrence of the pair), the distance in markers, the distance in kb (if the genotype spreadsheet was marker mapped), both R^2 and D' values for the EM method and the CHM method.

Unsort	Label	C 1	C 2	Distance in markers	Distance in kb	EM - R Squared	EM - D Prime	CHM - R Squared	CHM - D Prime
19	SNP_A-2057889	SNP_A-2241518	19	107.925	0.1328166604869349	0.313722169235354	0.3293442368417157	0.24520263885569	
20	SNP_A-2057889	SNP_A-4244414	20	107.957	0.011025078864464	0.935280100491622	0.00725897678881073	0.758906993676153	
21	SNP_A-2057889	SNP_A-2067066	21	129.944	0.0142806056286833	0.978368968002282	0.115006143924997	0.877996622435284	
22	SNP_A-2057889	SNP_A-1913732	22	129.926	0.014124211802054	0.977410407860535	0.0112769605075401	0.873355154626512	
23	SNP_A-2057889	SNP_A-2026046	23	134.418	0.014124211802054	0.977410407860535	0.0112769605075401	0.873355154626512	
24	SNP_A-2057889	SNP_A-4258149	24	143.653	0.013852467912468	0.973074457762629	0.107786986326023	0.858352470042257	
25	SNP_A-2057889	SNP_A-4234680	25	145.107	0.014097680299789	0.976616739907389	0.0111352172998112	0.868665464301792	
26	SNP_A-2057889	SNP_A-1836491	26	147.266	0.013637936703397	0.9823628051233	0.0137346014400237	0.986247405712691	
27	SNP_A-2057889	SNP_A-4282373	27	154.134	0.014124211802054	0.977410407860535	0.0112769605075401	0.873355154626512	
28	SNP_A-2057889	SNP_A-2149200	28	154.503	0.014124211802054	0.977410407860535	0.0112769605075401	0.873355154626512	
29	SNP_A-2057889	SNP_A-1838915	29	156.52	0.0514968734109784	0.474563835562549	0.0597440993796958	0.511153967523242	
30	SNP_A-2057889	SNP_A-1836510	30	163.752	0.0149201052010058	0.97941085915853	0.0133775805828589	0.927401427646584	
31	SNP_A-2057889	SNP_A-2056328	31	164.351	0.0142554642092355	0.977647914330333	0.0113750979078892	0.873311630712922	
32	SNP_A-2057889	SNP_A-2053994	32	164.46	0.0480284524934604	0.590905399665188	0.0636203268592923	0.50518191377441	
33	SNP_A-2057889	SNP_A-1820470	33	168.575	0.0514579067102755	0.614862602790051	0.0351854662926527	0.508100939910701	
34	SNP_A-2057889	SNP_A-4281588	34	176.835	0.0518240867956488	0.617830525084585	0.0352082847655593	0.508506408790012	
35	SNP_A-2057889	SNP_A-2050549	35	183.595	0.009950306630303	0.9518673892162	0.00799370240452801	0.831543597116625	
36	SNP_A-2057889	SNP_A-1800897	36	194.847	0.000114316447173746	0.0185212839436728	0.000154187036964644	0.0215100124530124	
37	SNP_A-2032943	SNP_A-2048970	1	6.561	0.179527678745207	0.999473967316339	0.17352022974388	0.982614314960071	
38	SNP_A-2032943	SNP_A-2159364	2	9.194	0.00809708932880188	0.220727609366008	0.0102023289601539	0.24885640702853	
39	SNP_A-2032943	SNP_A-1938959	3	21.138	0.04835628037842	0.99952952442232	0.077366007744058	1.10489176137792	
40	SNP_A-2032943	SNP_A-1802724	4	27.627	0.0801588187206131	0.99961996239913	0.098820722893597	1.10989853802636	
41	SNP_A-2032943	SNP_A-2272371	5	51.913	0.0638886189404851	0.9901990494482	0.0799136056481325	1.1032362446392	
42	SNP_A-2032943	SNP_A-1910767	6	57.026	0.003171485223109944	0.220254410798884	0.00135367342766707	0.195684104818262	
43	SNP_A-2032943	SNP_A-4252261	7	57.212	0.04445880156575	0.99676887245318	0.03974823982175	0.94269345761708	
44	SNP_A-2032943	SNP_A-1938509	8	61.471	0.10948203842384	0.99930623978837	0.157454853799745	1.198939457120924	
45	SNP_A-2032943	SNP_A-2289802	9	66.16	0.045357640001332	0.997418787888474	0.0418415996584366	0.957980037635631	
46	SNP_A-2032943	SNP_A-1904645	10	68.749	0.0808882394625289	0.86062421712157	0.083860306689992	0.76542289767295	
47	SNP_A-2032943	SNP_A-1802724	11	71.72	0.0303110990479407	0.44022248054232	0.109626602523642	0.483107676658847	