Uncovering cryptic asexuality in *Daphnia magna* by RAD-sequencing

Nils Svendsen, Celine M. O. Reisser, Marinela Dukić, Virginie Thuillier, Adeline Ségard, Cathy Liautard-Haag, Dominique Fasel, Evelin Hürlimann, Thomas Lenormand, Yan Galimov, and Christoph R. Haag

Supporting information

Supporting File S1

Assessing the inter-chromosomal pattern when centromere locations are unknown

The inter-chromosomal prediction for offspring heterozygosity under automixis is that all chromosomes in a given offspring should either retain 100% of parental heterozygosity or become fully homozygous at markers in the centromere regions. If centromere locations are unknown, this prediction cannot directly be assessed. However, if mapped markers are available, it is possible to test for specific "segregation patterns" by tabulating, for each marker, the individuals in which the marker becomes homozygous and in which it retains parental heterozygosity. If offspring are produced by central fusion, one would expect to find on each chromosome markers that retain parental heterozygosity in all offspring. Similarly, with pure terminal fusion, one would expect to find on each chromosome markers that become homozygous in all individuals. A sufficient number of markers is needed so that it can be assumed that each chromosome contains at least one marker that is in full linked with the centromere. If some offspring are produced by a terminal fusion and some by central fusion, one would expect to find on each chromosome markers that are heterozygous in a given set of offspring (those produced by central fusion) and homozygous in the rest (those produced by terminal fusion), with the important point being that it should be the same set of individuals that retain heterozygosity for all these markers and each chromosome should contain at least one of these markers.

We illustrate this with using a reduced set of loci with complete information (no missing genotypes) for all eight automictic offspring (N = 1693 loci). With eight offspring, there are 2^8 = 256 possible segregation patterns, each of which can be represented binary string for offspring1 to offspring8 (zero: homozygous, 1: heterozygous). For instance 00011000 is a marker, which is heterozygous in offspring4 and offspring5, and homozygous in all other offspring. We identified the segregation pattern for each of the 1693 markers and counted how many times and on how many linkage groups each specific segregation pattern occurred. Only one segregation pattern occurred on all ten linkage groups: homozygous in all individuals except individual V04_04. This pattern was shown by a total of 332 loci, with between 12 and 106 loci per linkage group. Moreover, on each linkage group, these markers were located in just one region. The ten other most common segregation patterns (Supporting Table S4) include loci that were heterozygous in all offspring (found on eight linkage groups) and loci homozygous in all offspring (found six linkage groups), but they did not occur in just

a single region in these linkage groups and probably contain some error (genotyping error alignment error, etc.). The only other pattern that was observed on more than four linkage groups is a pattern that is very similar to the presumed centromeric one (Supporting Table S4), and indeed was found in many pericentromeric regions. None of the segregation patterns among the 27 self-fertilized offspring occurred on more than three linkage groups (a total of 769 markers were investigated), except for 34 loci distributed across nine linkage groups that were heterozygous in all individuals. Within each linkage group, these loci did not occur in a single region, and show strongly differing segregation patterns compared to adjacent markers, which suggests that they may be explained by alignment errors (e.g., false mapping of paralogous loci to a single position).

Overall these results show that even without information on the centromere locations it is possible to conclusively infer the mode of reproduction, given a sufficient number of mapped markers. Conversely, the results also show that mapping of centromeres can be achieved and even if some offspring are produced by terminal fusion and others by central fusion, and that the proportion of offspring produced by terminal vs. central fusion can be directly estimated from the same data.

File S2:

Expected offspring heterozygosity under central vs. terminal fusion

Expected heterozygosity $H(d)$ at a distance d (in Morgan) from the centromere can be computed in two steps. The first step is to derive expected heterozygosity $H(x)$ for any fixed number x of crossovers between the marker and the centromere. This can be obtained by recurrence. Under terminal fusion, we have

$$H(x+1) = 1 - H(x) + H(x)/2 \qquad \text{(A1)}$$

Indeed, if the marker was homozygous (1-$H(x)$), it becomes heterozygous with an additional crossing over, and if it was already heterozygous, there is only one chance over two that it will remain heterozygous with an additional crossing over ($H(x)/2$). Hence, with $H(0) = 0$ (i.e., terminal fusion), we obtain

$$H(x) = \frac{2}{3}\left(1 - \left(-\frac{1}{2}\right)^x\right) \qquad \text{(A2)}$$

This function oscillates (0, $1, \frac{1}{2}, \frac{3}{4}, \frac{5}{8}, \frac{11}{16}, \frac{21}{32}$, ...) and stabilizes at 2/3 after many cross-overs. (Note that heterozygosity under central fusion can be obtained from the result under terminal fusion noting that $H_{cf} = 1 - H_{tf}/2$ and that $H_{cf}(0) = 1$; Engelstädter *et al.* 2011). The second step is to assume that, in absence of interference, the number of crossovers X over a distance d follows a Poisson distribution with mean $2d$ (recalling that 0.5 Morgan corresponds to one cross-over). We obtain

$$H(d) = \sum_{x=0}^{\infty} P(X = x)\frac{2}{3}\left(1 - \left(-\frac{1}{2}\right)^x\right) \qquad \text{(A3)}$$

where $P(X = x)$ is given by the Poisson distribution. We find

$$H(d) = \frac{2}{3}(1 - e^{-3d}) \qquad \text{(A4)}$$

(Engelstädter *et al.* 2011). The equivalent result under central fusion is

$$H(d) = 1 - \frac{1}{3}(1 - e^{-3d}) \quad \text{(A5)}$$

(Rizet and Engelmann 1949; Barratt *et al.* 1954). In order to compute $H(d)$ in presence of interference, we propose here to use Conway-Maxwell Poisson distribution (Sellers *et al.* 2012) that generalizes the Poisson distribution allowing for over or underdispersion (positive interference corresponding to underdispersion). This distribution adds a parameter ν to control for the level of dispersion. Its probability density function is

$$P(X = x) = \frac{\lambda^x}{Z(\lambda, \nu)(x!)^\nu} \quad \text{(A6)}$$

where $Z(\lambda, \nu)$ is a normalization equal to $\sum_x \lambda^x/(x!)^\nu$, which can be expressed using the generalized hypergeometric function

$$Z(\lambda, \nu) = \ {}_0F_{\nu-1}(\emptyset, \mathbf{1}, \lambda), \quad \text{(A7)}$$

where **1** is a vector of 1 of dimension ν-1. Using the probability density (A7) in Eq. (A6) yields an heterozygosity function $H(d)$ for various degree of interference. This is illustrated in Figure 1. Strong interference leads to a non-monotonic mapping function as more evenly spaced cross over events will cause $H(d)$ to reflect the oscillatory behavior of $H(x)$ (Eq. A2). All mapping functions have a slope of two at d=0 and tend to 2/3 for large d. Non monotonicity arises as soon as there is interference, but it becomes noticeably large for $\nu \geq 2$. This method can also be applied to obtain a standard mapping function $M(d)$ expressing the recombination fraction as a function of the genetic distance. For instance using the Mather formula (Mather 1935)

$$M(d) = \frac{1}{2}\big(1 - P(X = 0)\big) = \frac{1}{2}(1 - Z(\lambda, \nu)^{-1}) \quad \text{(A8)}$$

In both cases, the mapping requires to express $H(d)$ or $M(d)$ not in terms of λ the parameter of the COM-Poisson distribution, but in terms of d (which is half the expected number of cross over, i.e. half the mean of the COM-Poisson distribution). Here again, the mean of the COM-

Poisson can be expressed in terms of generalized hypergeometric functions, but a simpler approximation is sufficient for most purposes:

$$E(X) = 2d = (1 - e^{-2\lambda})\left(\lambda^{1/\nu} - \frac{\nu - 1}{2\nu}\right) + \lambda e^{-4\lambda} \qquad \text{(A9)}$$

Supporting Figure S2 illustrates this mapping. The case $\nu = 1$ corresponds to Haldane mapping, while $\nu = 3$ is close to the Kosambi mapping used in *Drosophila* (Chen 2013). Note that heterozygosity with interference has already been treated by Barratt *et al.* (1954) for the case of central fusion, however using a less general model (necessitating more restrictive assumptions) than the models based on the COM-Poisson distribution (see also, Nace *et al.* 1970; Zhao and Speed 1998). The latter and other count models (e.g., Zhao *et al.* 1995) are increasingly used also to model interference in classical genetic mapping (e.g., Choi *et al.* 2013).

References

Barratt, R. W., D. Newmeyer, D. D. Perkins, and L. Garnjobst, 1954 Map construction in *Neurospora crassa*. Adv. Genet. 6**:** 1-93.

Chen, Z., 2013 *Statistical methods for QTL mapping*. Chapman and Hall/CRC, New York.

Choi, K., X. Zhao, K. A. Kelly, O. Venn, J. D. Higgins *et al.*, 2013 *Arabidopsis* meiotic crossover hotspots overlap with H2A.Z nucleosomes at gene promoters. Nat. Genet. 45**:** 1327-1336.

Engelstädter, J., C. Sandrock, and C. Vorburger, 2011 Contagious parthenogenesis, automixis, and a sex determination meltdown. Evolution 65**:** 501-511.

Mather, K., 1935 Reductional and equational separation of the chromosomes in bivalents and multivalents. J. Genet. 30**:** 53-78.

Nace, G. W., C. M. Richards, and J. H. Asher, 1970 Parthenogenesis and genetic variability. I. Linkage and inbreeding estimations in the frog, *Rana pipiens*. Genetics 66**:** 349-368.

Rizet, G., and C. Engelmann, 1949 Contribution à l'étude génétique d'un Ascomycète tétrasporé: *Podospora anserina* (Ces.) Rehm. Rev. Cytol. Biol. Veg. 11**:** 201–304.

Sellers, K. F., S. Borle, and G. Shmueli, 2012 The COM-Poisson model for count data: a survey of methods and applications. Appl. Stoch. Models Bus. Ind. 28**:** 104-116.

Zhao, H., and T. P. Speed, 1998 Statistical analysis of ordered tetrads. Genetics 150**:** 459-472.

Zhao, H., T. P. T. P. Speed, and M. S. McPeek, 1995 Statistical analysis of crossover interference using the chi-square model. Genetics 139**:** 1045-1056.

File S3:

Detailed RAD-sequencing protocol and analysis of RAD-sequencing data

Prior to DNA extraction, individuals were treated for 72 hours with three antibiotics (Streptomycin, Tetracyclin, Ampicilin) at a concentration of 50 mg/L of each antibiotic and fed with microscopic glass beads (Sephadex "Small" by Sigma Aldrich: 50 µm diameter) at a concentration to 0.5g/100 mL. The aim of this treatment was to minimize contaminant DNA (i.e., bacterial DNA or algal DNA) in in the gut and on the surface of the body. Genomic DNA was extracted using the Qiagen Blood and Tissue kit following manufacturer's instructions and digested with *Pst*I (New England Biolabs). Digested DNA was barcoded with individual-specific P1 adapters and pooled to create a library containing 2100ng DNA. The pooled library was sheared on a Bioruptor using 2 times 3 cycles (1 cycle 30 seconds ON, 1 minute OFF), and fragments between 300 and 500bp were selected through agarose gel electrophoresis. DNA fragments were blunted and a P2 adapter was ligated. The library was amplified through PCR (30 seconds at 98°C, followed by 18 cycles of 10 sec. at 98°C, 30 sec. at 65°C and 30 sec. at 72°C; a final elongation step was performed at 72°C for 5 min.). A final electrophoresis was performed to select and purify fragments between 350 and 600bp. Each library were sequenced on a single lane of an Illumina HiSeq 2000, using single-end 100 cycle sequencing by the Quantitative Genomics Facility service of the Department of Biosystem Science and Engineering (D-BSSE, ETH), Basel, Switzerland.

The quality of the raw sequencing reads (library-wide and per-base) was assessed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and reads were checked for barcode integrity, absence of adapter sequences within the reads, and integrity of the *Pst*I cut site. The reads were sorted individually by barcode and filtered to remove reads with uncalled bases or an overall base-call quality score of less than 25. The last five bases of each read were trimmed due to a decrease in base-calling quality. Reads were subsequently aligned to the *Daphnia magna* genome (V2.4, 20100422; *Daphnia* Genomic Consortium, WFleaBase) using BWA v.0.7.10 (Li and Durbin 2009). Reads that did not map to the reference genome or that mapped to more than one place were discarded. The remaining reads were filtered according to mapping quality (reads that did not map end-to-end, had a mapping quality score of less than 25, or more than eight substitutions compared to the reference genome were discarded).

Assignment of reads to RAD loci (defined by unique 90 bp locations on the reference genome) and genotype calling was performed in Stacks V1.19 (Catchen *et al.* 2011), with a bounded SNP model in *pstacks* (--bound_high of 0.04, according to the base call error rate provided by the sequencing facility). We only retained loci with a maximum of two high frequency haplotypes (i.e. alleles) per locus per individual (maximally two alleles are expected in a diploid individual). Low-frequency haplotypes (i.e., representing less than 2% of the number of reads per locus in a given individual) were discarded due to the possibility of sequencing error. Routines *cstacks* and *sstacks* were operated with default settings and with the -g option to use genomic location as method to group reads. We also used the option –n with a parameter of 2 in *cstacks* (i.e., allowing a maximum of two mismatches between individuals) to reduce the risk of considering paralogous loci as alleles. For genotype calling, the distribution of the minor allele frequency indicated that a large majority of heterozygous loci had a minor allele frequency between 0.2 and 0.5 within individuals. We thus fixed the max_het_seq parameter to 0.2 in the routine *genotypes*. Consequently, genotypes with a minor allele frequency of between 0.05 (default homozygote cut-off) and 0.2 were considered ambiguous and were scored as missing data. Loci were also filtered according to sequencing depth: Loci with less than 20 reads were discarded (to reduce uncertainty in genotype calls, Han *et al.* 2014), as were reads with a more than five times higher depth than the average depth across all RAD-loci within a given individual (to reduce the risk of including repetitive elements / multi –copy genes). Finally, we used the automated correction procedure in Stacks to correct potentially miscalled offspring genotypes through a reassessment of the likelihood of genotype calls taking parental genotypes into consideration (Catchen *et al.* 2011). Only loci that were consistently called heterozygous in both replicates of the parental individual were retained.

After genotype calling, loci were mapped to the *Daphnia magna* genetic map v4.0.1 (M. Dukić et al., unpubl., deposited on Dryad). This was done by extracting, for each RAD locus, the linkage group and cM position of the nearest map-markers on the same scaffold and, if needed, by extrapolating the cM position of the RAD locus by linear extrapolation between the two nearest map-markers. Missing genotypes were inferred only if (i) two other RAD-loci were present on both sides of the missing marker on the same scaffold, (ii) these four other loci indicated that no crossing over had occurred in this region in this particular individual, and (iii) the genotypes of the other individuals for that locus were consistent with correct mapping of the locus (no more than two recombination events compared to loci on either side across the eight offspring of the AST-01-04 clone and no more than four

recombination events across the 27 offspring of the RM1-18 clone). Similarly, suspect genotypes suggesting one crossover immediately before and a second crossover immediately after the locus were removed and treated as missing if (i) the two loci immediately before and the two loci immediately after on the same scaffold suggested no crossover in that region in this particular individual (without considering loci with missing data) and (ii) if the genotypes of the other individuals for that locus were consistent with correct mapping of the locus (using the same criteria as above). We refrained from additional inference of missing genotypes or removal of suspect genotypes because the *D. magna* genetic map was based on a different population (thus some re-arrangement may be possible) and also because the scaffolding of the current assembly may contain some errors.

After all filtering and correcting, we retained 2523 loci for the analysis of the AST-01-04 family (corresponding to the number of heterozygous loci in the parent clone) and 1610 loci for analysis of the RM1-18 family. Considering suspect individual genotypes as missing (0.5% in the AST-01-04 family and 6.4% in the RM1-18 family), 12.6% and 22.9% of all individual genotypes were missing in the two families, respectively, but this could be reduced to 5.7% and 17.8% by inferring missing genotypes according to above criteria. The proportion of missing and suspect genotypes in the RM1-18 family was higher than in the AST-01-04 family, likely due to the lower average sequencing depth (43.3 reads per locus and individual in the RM1-18 library vs. 54.8 reads per locus and individual in the AST-01-04 library). Both the original and the corrected data set will be deposited on Dryad, but only the corrected one was used in the analyses.

References

Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait, 2011 Stacks: building and genotyping loci *de novo* from short-read sequences. G3 (Bethesda) 1**:** 171-182.

Han, E., J. S. Sinsheimer, and J. Novembre, 2014 Characterizing bias in population genetic inferences from low-coverage sequencing data. Mol. Biol. Evol. 31**:** 723-735.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25**:** 1754-1760.

Files S4-S6:

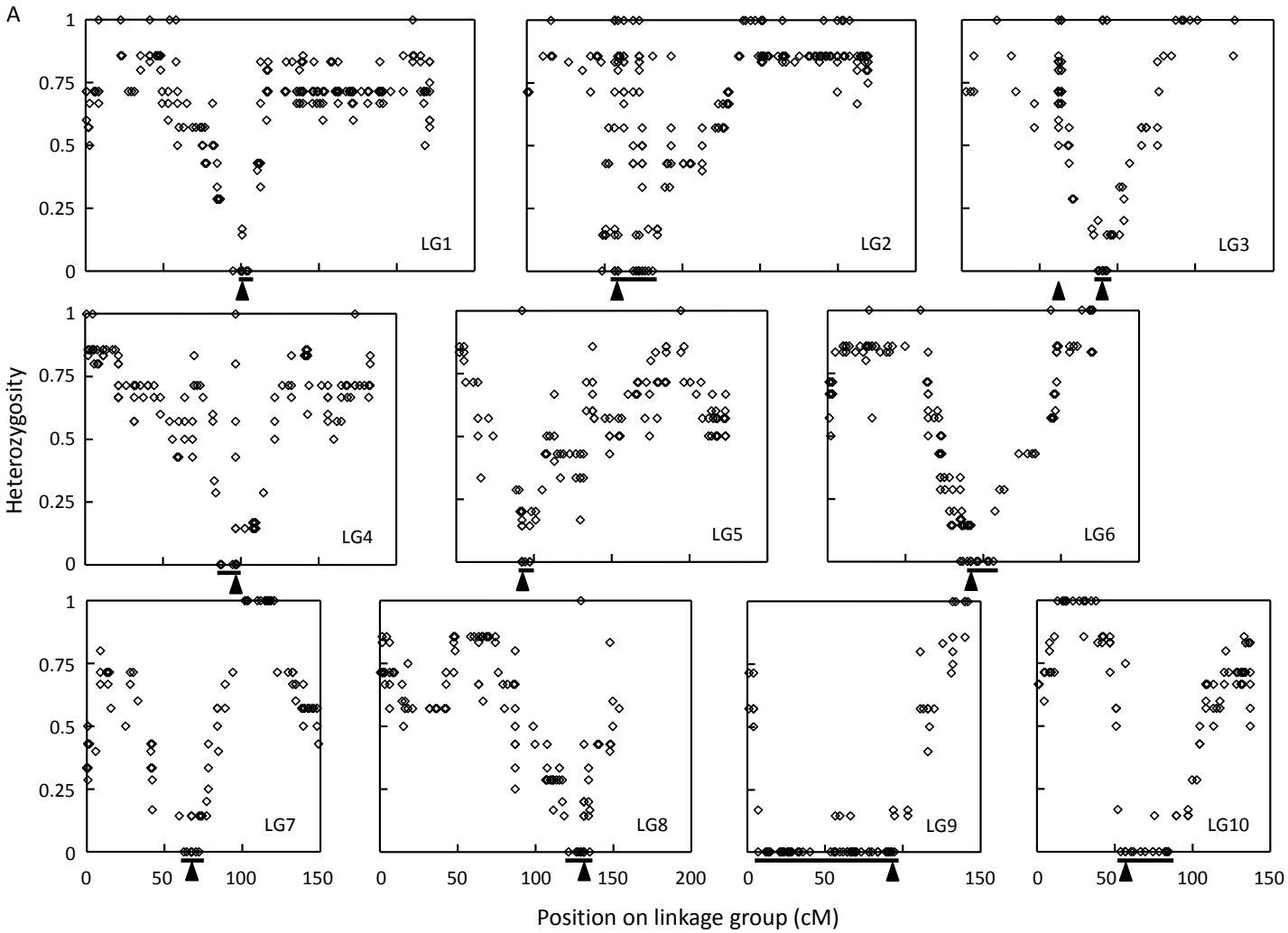
Available for download at

www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.179879/-/DC1

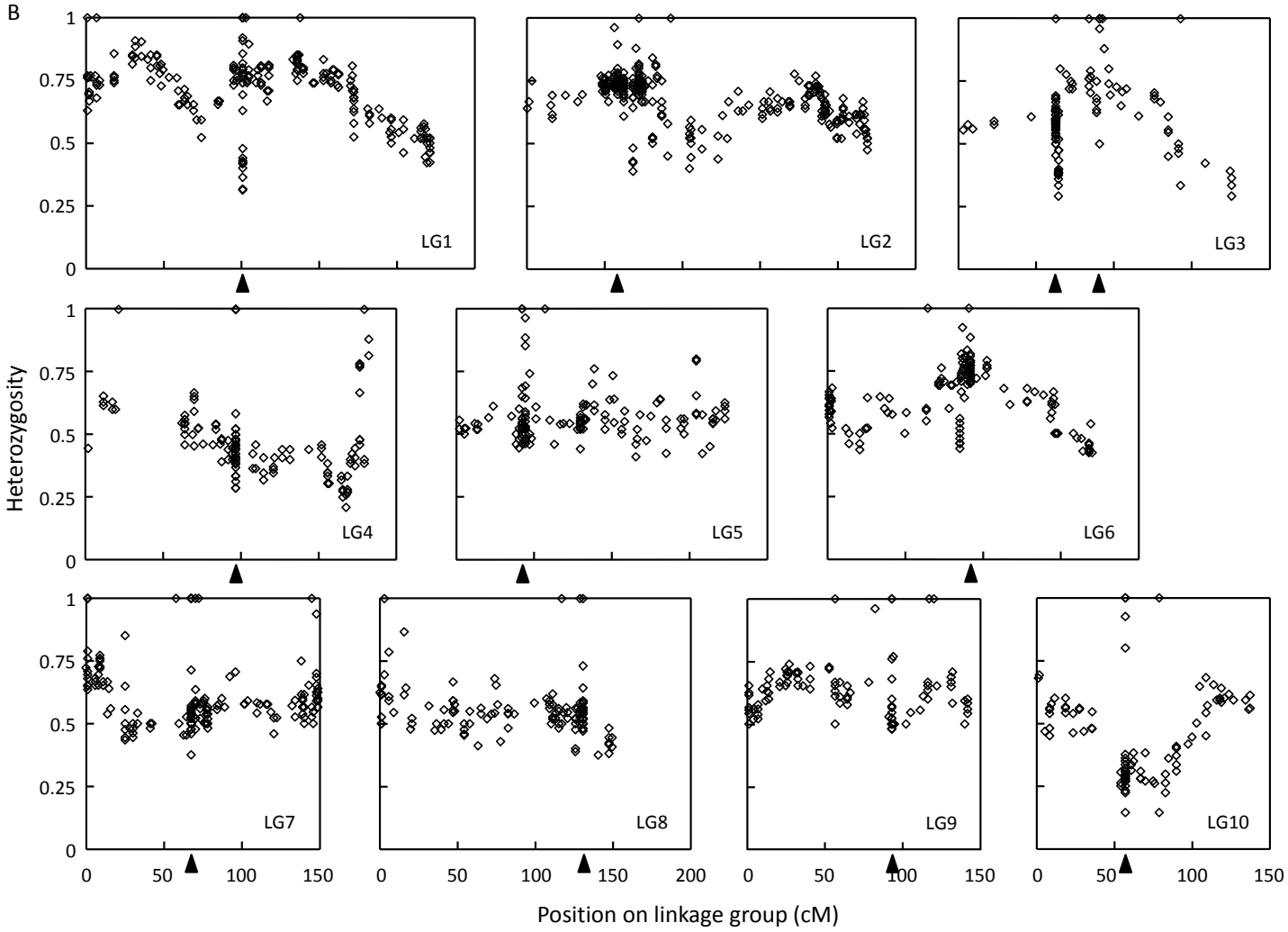
File S4: .xlsx file containing the raw and corrected SNP data set for all offspring of the AST-01-04 NMP clone.

File S5: .xlsx file containing the raw and corrected SNP data set for all offspring of the RM-1-18 MP clone.

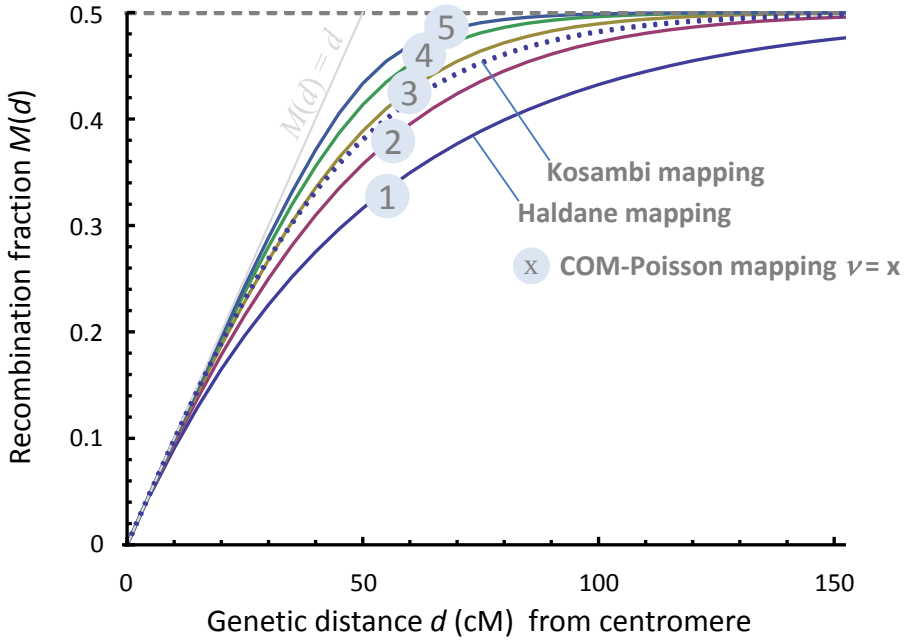
File S6: .xlsx file containing the genetic map v4.0.1.



Supporting Figure S1: (A) Heterozygosity among the seven offspring of the AST-01-04 NMP clone produced by terminal fusion, depending on marker position on the linkage groups of *D. magna*. The bar below each linkage group shows the region around the presumed centromere in which offspring V04-04 was heterozygous, and the triangle shows the presumed centromere location (two in case of LG3). (Figure continued on next page)



Supporting Figure S1: (B) Heterozygosity among the 27 offspring of the RM1-18 MP clone.



Supporting Figure S2: Mapping functions under different degrees of interference. $M(d)$ is the recombination fraction as a function of the distance d (in Morgan) from the centromere and different degrees of interference as measured by the level of underdispersion ν of the COM-Poisson distribution describing the number of crossing-over per unit of genetic map distance. The case $\nu = 1$ corresponds to zero interference (Haldane mapping), while Kosambi mapping is close to the case $\nu = 3$.

Supporting Table S1: Microsatellite genotypes of parents and offspring as well as inferred parents (for offspring only). Potential parents indicate the parental clones that were placed in the buckets from which the offspring was obtained. B008 to B135 are the nine microsatellite loci used in this study (MOLECULAR ECOLOGY RESOURCES PRIMER DEVELOPMENT CONSORTIUM *et al.* 2011), with genotypes indicated by fragment lengths. (Table continued on next pages).

A. Microsatellite genotypes of parent clones.

Clone ID	MP/NMP	Bucket(s)	B008	B030	B045	B050	B064	B074	B096	B107	B135
1MDM6	NMP	V01, V06	165/167	157/160	120/122	234/240	142/146	198/198	243/245	262/270	
AST-01-04	NMP	B20, B23, B26, V04, V10, V12	163/163	155/157	122/122	237/240	138/144	198/200	239/245	272/278	187/191
BN-08	NMP	V12	167/167	157/157	122/122	228/242		198/200	241/241	266/266	
BN-48	NMP	V05	163/179	159/159	120/122	228/245	140/140	200/200	239/241	270/272	
MOS-01-02	NMP	V09, V17	163/172	155/163	122/122	237/240	138/138	196/198	241/241	270/270	
MOS-01-04	NMP	B20, B23, B26, V12, V18	167/172	160/160	122/122	232/232	144/144	198/198	241/241	270/278	189/191
RM1-02	NMP	V21	167/167	155/160	122/122	235/235	144/146	198/200	243/243	268/280	
VV2	NMP	B20, B23, B26, V01, V02, V05, V06, V08, V09, V12	157/165	153/157	122/122	226/244	138/144	198/202	239/243	270/270	187/189
RM1-18	MP	B22, B25, B28	167/172	155/155	127/127	237/237					187/191
RM1-22	MP	B22, B25, B28	161/167	155/160	122/122	232/232					191/191
RM1-35	MP	B22, B25, B28	172/172	155/160	120/122	237/237					189/189
RM1-39	MP	B22, B25, B28	167/169	155/155	122/122	232/242					189/189

B. Microsatellite genotypes of offspring.

Hatchling ID	Treatment	Bucket ID	Potential parents	B008 25 cM	B030 33 cM	B045	B050	B064	B074	B096	B107	B135	Parent1	Parent2	Outcrossed
V02-01	NMP_single	V02	VV2	165/165	153/157	122/122	226/244	138/138	198/202	239/243	270/270		VV2		No
V02-03	NMP_single	V02	VV2	157/157	153/157	122/122	226/244	144/144	198/198	243/243	270/270		VV2		No
V04-01	NMP_single	V04	AST-01-04	163/163	155/157	122/122	237/240	138/144	198/200	239/239	272/278		AST-01-04		No
V04-02	NMP_single	V04	AST-01-04	163/163	155/157	122/122	237/240	138/138	198/200	239/245	278/278		AST-01-04		No
V04-03	NMP_single	V04	AST-01-04	163/163	157/157	122/122	237/240	138/144	198/200	245/245	272/272		AST-01-04		No
V04-05	NMP_single	V04	AST-01-04	163/163	157/157	122/122	237/240	138/144	198/200	239/239	272/278		AST-01-04		No
V04-06	NMP_single	V04	AST-01-04	163/163	155/157	122/122	237/240	138/144	198/200	245/245	272/278		AST-01-04		No
V04-11	NMP_single	V04	AST-01-04	163/163	155/157	122/122	237/237	144/144	198/200	245/245			AST-01-04		No
V04-14	NMP_single	V04	AST-01-04	163/163	157/157	122/122	237/240	138/144			272/278		AST-01-04		No
V04-18	NMP_single	V04	AST-01-04	163/163	155/157	122/122	237/240	138/138	198/200	239/239	272/272		AST-01-04		No
V04-19	NMP_single	V04	AST-01-04	163/163	157/157	122/122	237/240	144/144	198/200	245/245			AST-01-04		No
V04-22	NMP_single	V04	AST-01-04	163/163	155/157	122/122	237/240	144/144	198/200	239/239	272/272		AST-01-04		No
V04-23	NMP_single	V04	AST-01-04	163/163	155/157	122/122	237/240	138/138	198/200	239/245	272/278		AST-01-04		No
V04-24	NMP_single	V04	AST-01-04	163/163	155/157	122/122	237/240	138/144	198/200	245/245	272/278		AST-01-04		No
V04-25	NMP_single	V04	AST-01-04	163/163	155/157	122/122	237/240	138/138	198/200	239/239	272/278		AST-01-04		No
V04-27	NMP_single	V04	AST-01-04	163/163	157/157	122/122	237/240	138/144	198/200	239/239	272/278		AST-01-04		No
V08-01	NMP_single	V08	VV2	165/165	153/153	122/122	244/244	138/144	198/202	239/243	270/270		VV2		No
V08-02	NMP_single	V08	VV2	157/157	157/157	122/122	226/244	138/144	198/198	239/239	270/270		VV2		No
V08-03	NMP_single	V08	VV2	165/165	153/157	122/122	226/244	138/144	198/202	239/243	270/270		VV2		No
V10-01	NMP_single	V10	AST-01-04	163/163	155/157	122/122	240/240	138/144	198/198	239/245	272/278		AST-01-04		No
V10-02	NMP_single	V10	AST-01-04	163/163	155/157	122/122	237/237	138/144	198/200	239/245	272/278		AST-01-04		No
V10-03	NMP_single	V10	AST-01-04	163/163	155/157	122/122	237/237	138/144	198/200	239/245	272/278		AST-01-04		No
V21-03	NMP_single	V21	RM1-02	167/167	160/160	122/122	235/235	146/146	198/200	243/243	268/280		RM1-02		No
V21-04	NMP_single	V21	RM1-02	167/167	160/160	122/122	235/235	144/146	200/200	243/243	268/268		RM1-02		No

Table S1B continued.

Hatchling ID	Treatment	Bucket ID	Potential parents	B008	B030	B045	B050	B064	B074	B096	B107	B135	Parent1	Parent2	Outcrossed
V21-05	NMP_single	V21	RM1-02	167/167	160/160	122/122	235/235	144/146	200/200	243/243	268/280		RM1-02		No
V21-07	NMP_single	V21	RM1-02	167/167	155/160	122/122	235/235	144/144	198/200	243/243	268/268		RM1-02		No
V21-08	NMP_single	V21	RM1-02	167/167	160/160	122/122	235/235	144/146	198/200	243/243	280/280		RM1-02		No
V01-02	NMP_mix	V01	1MDM6, VV2	157/165	153/153	122/122	226/244	138/144	198/202	239/243	270/270		VV2		No
V01-06	NMP_mix	V01	1MDM6, VV2	165/165	157/157	122/122	226/244	138/138	198/198	239/243			VV2		No
V01-07	NMP_mix	V01	1MDM6, VV2	157/165	153/157	122/122		138/144					VV2		No
V05-01	NMP_mix	V05	BN-48, VV2	165/165	153/157	122/122	226/244	138/138	198/202	239/243	270/270		VV2		No
V05-02	NMP_mix	V05	BN-48, VV2	157/165	153/157	122/122	244/244	138/138	198/202	239/239	270/270		VV2		No
V05-04	NMP_mix	V05	BN-48, VV2	165/165	157/157	122/122	226/244	138/144	198/202		270/270		VV2		No
V05-05	NMP_mix	V05	BN-48, VV2	165/165	153/153	122/122	226/244	138/138	198/202		270/270		VV2		No
V05-07	NMP_mix	V05	BN-48, VV2	165/165	153/153	122/122	226/226	144/144	198/202	239/243	270/270		VV2		No
V05-09	NMP_mix	V05	BN-48, VV2	157/165	153/157	122/122	226/244	144/144	198/202		270/270		VV2		No
V05-10	NMP_mix	V05	BN-48, VV2	165/165	153/153	122/122	226/244	138/144	198/198	243/243	270/270		VV2		No
V06-01	NMP_mix	V06	1MDM6, VV2	165/167	157/160	122/122		142/146	198/198				1MDM6		No
V09-06	NMP_mix	V09	MOS-01-02, VV2	157/165	153/157	122/122	226/244	144/144	202/202	243/243	270/270		VV2		No
V09-09	NMP_mix	V09	MOS-01-02, VV2	157/165	153/157	122/122	244/244	138/144	198/202	239/243	270/270		VV2		No
V09-10	NMP_mix	V09	MOS-01-02, VV2	157/165	153/157	122/122	226/226	138/144	198/198	239/239	270/270		VV2		No
V12-01	NMP_mix	V12	AST-01-04, BN-08, MOS-01-04, VV2	157/165	153/157	122/122	244/244	138/138	198/202	239/243	270/270		VV2		No
V12-02	NMP_mix	V12	AST-01-04, BN-08, MOS-01-04, VV2	167/167	160/160	122/122	232/232	144/144	198/198	241/241	270/270		MOS-01-04		No
V17-01	NMP_mix	V17	MOS-01-02, MOS-01-04	167/172	160/160	122/122	232/232	144/144	198/198	241/241	270/270		MOS-01-04		No
V17-02	NMP_mix	V17	MOS-01-02, MOS-01-04	167/167	160/160	122/122	232/232	144/144	198/198	241/241	270/278		MOS-01-04		No
V17-03	NMP_mix	V17	MOS-01-02, MOS-01-04	167/172	160/160	122/122	232/232	144/144	198/198	241/241	270/278		MOS-01-04		No
V17-04	NMP_mix	V17	MOS-01-02, MOS-01-04	172/172	160/160	122/122	232/232	144/144	198/198	241/241			MOS-01-04		No

Table S1B continued.

Hatchling ID	Treatment	Bucket ID	Potential parents	B008	B030	B045	B050	B064	B074	B096	B107	B135	Parent1	Parent2	Outcrossed
B20-01	NMP_mix	B20	AST-01-04, MOS-01-04, VV2	167/172	160/160	122/122	232/232					189/191	MOS-01-04		No
B23-01	NMP_mix	B23	AST-01-04, MOS-01-04, VV2	167/167	160/160	122/122	232/232					189/189	MOS-01-04		No
B23-02	NMP_mix	B23	AST-01-04, MOS-01-04, VV2	157/165	153/157	122/122	226/244					187/189	VV2		No
B23-03	NMP_mix	B23	AST-01-04, MOS-01-04, VV2	172/172	160/160	122/122	232/232					191/191	MOS-01-04		No
B26-01	NMP_mix	B26	AST-01-04, MOS-01-04, VV2	163/163	157/157	122/122	237/240					187/191	AST-01-04		No
B22-11	MP_mix	B22	RM1-18, RM1-22, RM1-35, RM1-39	167/172	155/155	122/127	232/232					187/191	RM1-18	RM1-22	Yes
B22-12	MP_mix	B22	RM1-18, RM1-22, RM1-35, RM1-39	167/169	155/155	122/122	232/232					189/191	RM1-22	RM1-39	Yes
B22-13	MP_mix	B22	RM1-18, RM1-22, RM1-35, RM1-39	167/169	155/160	122/122	232/232					189/191	RM1-22	RM1-39	Yes
B22-14	MP_mix	B22	RM1-18, RM1-22, RM1-35, RM1-39	167/172	155/155	122/122	232/237					189/191	RM1-22	RM1-35	Yes
B22-15	MP_mix	B22	RM1-18, RM1-22, RM1-35, RM1-39	167/169		122/122						189/191	RM1-22	RM1-39	Yes
B22-16	MP_mix	B22	RM1-18, RM1-22, RM1-35, RM1-39	161/172		122/127						187/191	RM1-18	RM1-22	Yes
B22-17	MP_mix	B22	RM1-18, RM1-22, RM1-35	167/167		122/127						191/191	RM1-18	RM1-22	Yes
B22-18	MP_mix	B22	RM1-18, RM1-22, RM1-35	167/172		120/127						189/191	RM1-18	RM1-35	Yes
B25-11	MP_mix	B25	RM1-18, RM1-22, RM1-35	167/169	155/155	122/122						189/191	RM1-22	RM1-39	Yes
B25-12	MP_mix	B25	RM1-18, RM1-22, RM1-35	167/169	155/155	122/127	237/242					187/189	RM1-18	RM1-39	Yes
B25-13	MP_mix	B25	RM1-18, RM1-22, RM1-35	167/169	155/155	122/122	232/232					189/191	RM1-22	RM1-39	Yes
B25-14	MP_mix	B25	RM1-18, RM1-22, RM1-35		155/160							189/189	RM1-35	RM1-39	Yes
B25-15	MP_mix	B25	RM1-18, RM1-22, RM1-35	161/167		122/122						191/191	RM1-22		No
B25-16	MP_mix	B25	RM1-18, RM1-22, RM1-35	167/169		122/122						189/191	RM1-22	RM1-39	Yes

Table S1B continued.

Hatchling ID	Treatment	Bucket ID	Potential parents	B008	B030	B045	B050	B064	B074	B096	B107	B135	Parent1	Parent2	Outcrossed
B25-17	MP_mix	B25	RM1-18, RM1-22, RM1-35	167/169		122/122						189/191	RM1-22	RM1-39	Yes
B25-18	MP_mix	B25	RM1-18, RM1-22, RM1-35			120/122						189/191	RM1-22	RM1-35	Yes
B28-11	MP_mix	B28	RM1-18, RM1-22, RM1-35	161/172	155/155	122/127	232/237					191/191	RM1-18	RM1-22	Yes
B28-12	MP_mix	B28	RM1-18, RM1-22, RM1-35	161/172	155/160	122/127	232/237						RM1-18	RM1-22	Yes
B28-13	MP_mix	B28	RM1-18, RM1-22, RM1-35	169/172	155/155	122/122	237/242						RM1-35	RM1-39	Yes
B28-14	MP_mix	B28	RM1-18, RM1-22, RM1-35	161/167	155/155	122/127	232/237					187/191	RM1-18	RM1-22	Yes
B28-17	MP_mix	B28	RM1-18, RM1-22, RM1-35	167/172		122/127						189/191	RM1-18	RM1-35 or RM1-39	Yes
B28-18	MP_mix	B28	RM1-18, RM1-22, RM1-35	167/167		122/127						191/191	RM1-18	RM1-22	Yes
B28-19	MP_mix	B28	RM1-18, RM1-22, RM1-35	161/172		122/122						189/191	RM1-22	RM1-35	Yes
B28-20	MP_mix	B28	RM1-18, RM1-22, RM1-35	167/172		122/127						191/191	RM1-18	RM1-22	Yes

Supporting Table S2: Average heterozygosity at microsatellite loci among offspring from single clone cultures and the position in of the microsatellites with respect to the centromere

Locus	Scaffold (*D. magna* assembly 2.4)	Position on scaffold (base pairs)	Linkage group (v4.0.1)	Distance from centromere (cM)	*N*[a]	Heterozygosity (confidence limita)
B008	scaffold03124	1781052	6	25.8	5	0 (0 – 0.40)
B030	scaffold00243	463057	3	32.9	27	0.59 (0.41 – 0.76)
B050	scaffold02066	483524	3	77.6	22	0.77 (0.56 – 0.90)
B064	scaffold00443	151077	1	37.9	27	0.63 (0.44 – 0.79)
B074	contig23904	1149	unmapped	NA	27	0.78 (0.59 – 0.90)
B096	scaffold01005	1506200	9	3.6	22	0.36 (0.20 – 0.57)
B107	scaffold00763	136826	1	100.0	22	0.59 (0.30 – 0.77)

[a]*N* refers to the number of genotyped offspring from heterozygous parents

Supporting Table S3: Number of loci and average heterozygosity per linkage group in automictic and self-fertilized offspring

Linkage group	*N* loci automicitc offspring	Average heterozygosity	*N* loci selfed offspring	Average heterozygosity
LG1	386	0.64	222	0.71
LG2	355	0.69	318	0.68
LG3	282	0.58	131	0.63
LG4	263	0.56	135	0.49
LG5	209	0.51	151	0.56
LG6	250	0.54	141	0.64
LG7	200	0.45	168	0.62
LG8	192	0.48	137	0.54
LG9	212	0.25	107	0.63
LG10	174	0.54	100	0.46
Average		**0.54**		**0.60**
Sum	**2523**		**1610**	

Supporting Table S4: The ten most frequently observed segregation patterns among the eight offspring of the AST-01-04 NMP clone. All other segregation patterns occurred for fewer loci or on a smaller number of linkage groups.

Linkage group	*N* loci total[a]	*N* loci (11111111)[b]	*N* loci (00000000)[b]	*N* loci (00000001)[b]	*N* loci (01000001)[b]	*N* loci (11011100)[b]	*N* loci (10011111)[b]	*N* loci (11111101)[b]	*N* loci (11111011)[b]	*N* loci (11110110)[b]	*N* loci (11110011)[b]
LG1	275	1	10	22	1		108		28	5	4
LG2	246	17	1	24	13			81		4	6
LG3	199	21	2	21	1	109					
LG4	174	2	1	37	2				19		8
LG5	122	2		12	6				1		
LG6	173	2		33	9					11	8
LG7	119			35	5						
LG8	131	4		18						23	
LG9	150		2	106	5						
LG10	104	15	1	24	6			5			
Sum	**1693**	**64**	**17**	**332**	**48**	**109**	**108**	**86**	**48**	**43**	**26**

[a]Only without missing values for any of the eight offspring

[b]Offspring are ordered in the following way: V10_03, V04_01, V04_16, V04_18, V04_22, V04_25, V04_27, V04_04, 1 = heterozygous, 0 = homozygous.