

# Notes on the fly project

J. Ignacio Lucas Lledó

October 5, 2016

## 1 Delivery of oligonucleotides

The oligonucleotides designed on 2016-06-09 were ordered to biomers.net, and delivered. They are dry. Table 1 specifies the yield and the corresponding amount of buffer required to dilute them to 100 pmol/ $\mu$ l.

## 2 DNA extraction test

On September 13<sup>th</sup> 2009, Dr. Carazo, Mrs. Sultanova, Mr. Andic and I started the DNA extraction from 8 flies, following the GBS protocol. Table 2 summarizes the procedure. Not that we needed 4 people to run this, but they all were interested in going through the protocol, which is great. The protocol is applied to each fly individually, with the aim of measuring how much DNA we can get from one fly. Four of the flies had their oviduct previously removed by Mrs. Sultanova. We want to make sure that the ovariectomy does not significantly reduce DNA yield.

We ended up measuring DNA concentration 3 times, due to problems with the Qubit standards. The error ‘Incorrect standards’ was apparently caused by the fact that the reagent had been kept in the fridge. I thought that it would be enough to wait for it to thaw, to bring it to room temperature. Plus, it is a very small amount of it that gets dissolved in the buffer, at room temperature. For whatever reason, the kit worked once we kept the reagent at room temperature. I thank Rebeca Dominguez, from the Evolutionary Genetics lab, her help to solve the problem.

The slight difference in average DNA concentration between control and ovariectomy samples is not significant, which is good news. I expect three samples, namely 1, 3, and 6, to have enough DNA left to run the rest of the protocol with them. I could add sample 5 to see what we get from it.

Table 1: Name and available amount (yield) of oligonucleotides received from biomers.net. ‘Volume’ makes reference to the volume required to dissolve the corresponding oligonucleotides at 100 pmol/ $\mu$ l ( $\mu$ M).

Name	Yield (nmol)	Volume ( $\mu$ l)	Name	Yield (nmol)	Volume ( $\mu$ l)
Top1.1	82.6	826	Bot1.1	64.8	648
Top1.2	106.8	1068	Bot1.2	110.4	1104
Top1.3	94.9	949	Bot1.3	92.6	926
Top1.4	85.6	856	Bot1.4	75.8	758
Top2.1	78.9	789	Bot2.1	90.7	907
Top2.2	81.9	819	Bot2.2	97.4	974
Top2.3	100.5	1005	Bot2.3	84.2	842
Top2.4	88.4	884	Bot2.4	82.3	823
Top3.1	98.6	986	Bot3.1	87.1	871
Top3.2	95.2	952	Bot3.2	96.6	966
Top3.3	100.2	1002	Bot3.3	87.9	879
Top3.4	95.6	956	Bot3.4	78.7	787
A1.1	60.2	602	A2	65.9	659
A1.2	52.6	526	R1	78.7	787
A1.3	56.7	566	R2	60.0	600
A1.4	51.0	510			

Table 2: Summary of the DNA extraction test.

Sample	Treatment	Elution ( $\mu$ l)	Conc.(ng/ $\mu$ l)	Yield (ng)
1	–	100	4.240	424
2	–	100	0.978	98
3	–	100	3.600	360
4	–	100	1.790	179
5	ovariectomy	100	1.990	199
6	ovariectomy	100	3.980	398
7	ovariectomy	100	1.140	114
8	ovariectomy	100	1.530	153

### 3 Literature review

Original motivation of classic life span QTL mapping studies was the elucidation of metabolic pathways involved in aging ?. I suspect that the biomedical interest in the mechanisms of aging preceded the interest in naturally occurring variation, as well as the real effect of the QTL identified in natural conditions. Several sex-specific life span QTL exist in autosomes of *D. melanogaster* ?. These QTL were originally identified using recombinant inbred lines (RIL) descended from two strains ?. In this design, the dominance of the QTL is not known, since only homozygous genotypes are produced (?, page 432). However, the quantitative complementation test with deficiencies used to refine the QTL ? seems to assume that one of the alleles is at least partially recessive.

Crow ? mentions the genetic basis of heterosis as one of the main mid-century controversies in population genetics. He is very clear that the debate between dominance and overdominance was resolved in favor of the dominance hypothesis. That is, both the inbreeding depression and the hybrid vigor (heterosis) are mostly due to recessive deleterious alleles that become either expressed upon inbreeding or concealed upon hybridization. A review ? cites another work by James F. Crow that I cannot access (volume 9 of the Oxford Series in Evolutionary Biology), in which the author shows evidence of 30% of autosomes isolated from natural populations of *D. melanogaster* being lethal in homozygosis. The same cannot be true for sex chromosomes, of course. Among the 70% of natural chromosomes that are not recessive lethals, homozygosity reduces survival to adulthood in 84%.

Full-sib mating during 10 generations is expected to produce an inbreeding coefficient of 0.886. This is the probability at any locus that the two copies inherited from the parents are identical by descent; that is, they are exact copies of the same chromosome in one of the recent ancestors. But how is this identity-by-descent distributed along a chromosome? It is not random, but clustered, because of linkage. We expect an alternative sequence of fragments that are either identical by descent or not. That is, the inbreeding itself already produces a mosaic of heterozygosity and homozygosity. The question is how long the pieces are. Too short IBD tracts would require very dense markers. According to I. R. Franklin ?, previous authors (Bennet and Fisher) had supposed that ‘after a long number of generations of inbreeding the number of heterogenic segments will be approximately distributed as a Poisson variable, and that the distribution of the length of each segment ( $x$ ) will tend to a negative exponential  $(1/a)e^{-x/a}$ .’ However, this assumptions are not satisfactory, especially for the first few generations of inbreeding. Taking into account that there is no crossing-over in *Drosophila males*, the

Table 3: Digestion reactions.

Sample	DNA ( $\mu$ l)	H <sub>2</sub> O ( $\mu$ l)	10 $\times$ buffer ( $\mu$ l)	NspI ( $\mu$ l)	Total ( $\mu$ l)
1	80	207	32	1	320
3	80	207	32	1	320
5	80	207	32	1	320
6	80	207	32	1	320

probability that two loci in the same chromosome are homozygous by descent after one generation of sib mating is  $(1/32)(3 + 3\lambda_f + \lambda_f^2 + \lambda_f^3)$  ?, where  $\lambda_f$  is the linkage parameter  $1 - 2y_f$ , representing the probability of recombination in the female.

The suggestion given by Dr. Carazo of using simulations is actually good. I started learning to use simuPOP.

## 4 Enzymatic digestion of DNA samples 1, 3, 5, and 6

I run the digestions as described on table 4, for above 4 hours, at 37°C. Together with Muhammed Andiç, we also cleaned up the reaction with magnetic beads. We used a 1:1 ratio of beads to DNA, and eluted with water (40  $\mu$ l).

## 5 Literature review

From page 489 of ?: "The adaptive significance of achiasmatic meiosis is not entirely clear. Obviously it has been evolved repeatedly in at least ten major groups of animals. In those cases where achiasmatic meiosis is confined to one sex, we have far too little information as to the chiasma frequency in the other sex. If it were generally true that in the 'chiasmatic sex' there is only one chiasma per bivalent, we might assume that selection for a low level of genetic recombination *per se* had led to total loss of chiasmata in one sex. If, on the other hand, absence of chiasmata in one sex is generally accompanied by a fairly high chiasma frequency in the other, this explanation obviously would not hold. Only evidence from species that had acquired achiasmatic mechanisms fairly recently in their phylogeny would be significant from this

point of view (*Tipula caesia* would be relevant – *Drosophila* spp. would not). The only evidence on this point seems to be the grasshopper *Thericles whitei*, with no chiasmata in the male and a very low chiasma frequency in the female; it obviously supports the hypothesis of selection for low levels of recombination."

However, selection for low levels of recombination would not explain why the achiasmatic sex is nearly always the heterogametic sex ?. White also mentions the hypothesis of achiasmatic spermatogenesis being an adaptation to tolerate inversion-heterozygosity. However, no evidence supports this hypothesis.

More recently, ? mentions two more hypotheses: first, that achiasmy may have evolved as a side effect of selection to suppress recombination between the sex chromosomes ??. And second, that it evolved as a way to promote tight linkage without suppressing recombination on the X or Z chromosome ?.

Well, I cannot review now all the hypotheses now. If anything, I should relate the results of the simulations to the theoretical expectations of the variance in inbreeding coefficient ??. It seems that multiple linked loci in sex chromosomes have received relatively little attention in the literature ??.

## 6 Annealing of adapters

On September 28th, Muhammed and I prepared 100 ml of 1× TE buffer, following a recipe online: 1.0 ml 1M Tris pH 8.0, 0.2 ml EDTA 0.5M, and 98.8 ml H<sub>2</sub>O. We also prepared annealing buffer, according to the protocol, and suspended all the oligonucleotides in TE buffer to 100 µM, according to the volume specified on table 1. Finally, we prepared 10 µl of annealed adapters of each type.

## 7 Quantification with Agilent Bioanalyser

Amparo Martínez, from the Sequencing Service, sent the results from the Bioanalyser run. The four samples have a very similar profile (Figure 1), which looks like the result of a successful digestion followed by a more aggressive selection of long fragments that I expected from a 1:1 ratio of magnetic beads to DNA. Next time, I should use a more generous proportion of beads. Table 4 shows some additional details. I note that the amount of DNA available is now an order of magnitude lower than what I had for the *Coregonus* samples.

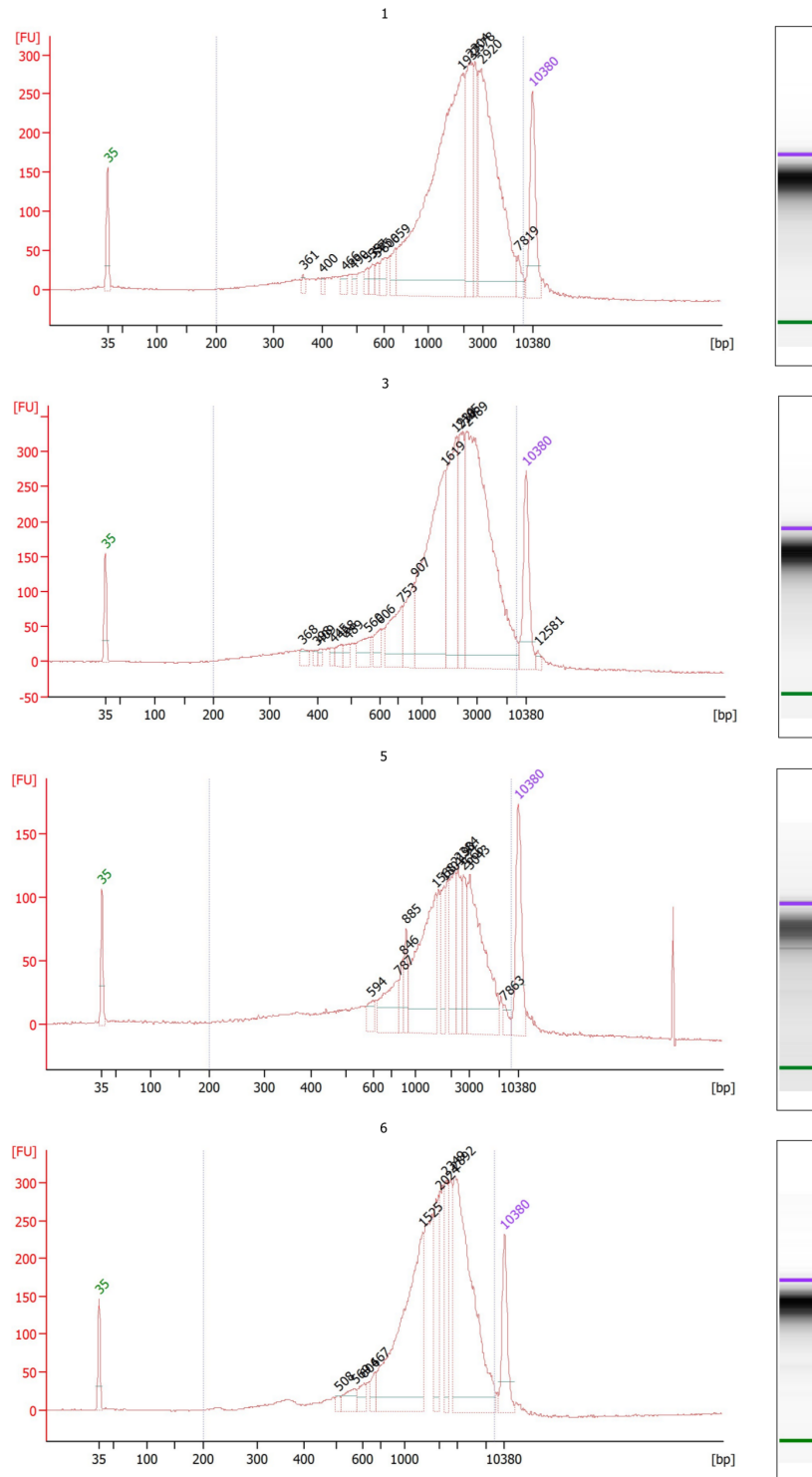


Figure 1: Electropherograms from the high-sensitivity DNA assay of the four samples.

Table 4: Distribution of DNA fragments between 40 and 9000 bp. ‘Size’ refers to the average fragment size. For total mass of DNA left over, a volume of 35  $\mu\text{l}$  is assumed to be left. The average fragment mass (Frag.Mass) is calculated in two ways. First, by dividing the overall concentration by the overall molarity (Frag.Mass.1); and then, by applying the formula for the molecular mass of double stranded DNA to the average length of the fragments (i.e.,  $607.4x + 157.9$  g/mol; Frag.Mass.2).

Sample	Size (bp)	Conc. (pg/ $\mu\text{l}$ )	Molarity (pmol/l)	Left over (ng)	Frag.Mass.1 (ng/fmol)	Frag.Mass.2 (ng/fmol)
1	2206	1425.92	3042.0	49.91	0.468744	1.34008
3	2239	1439.86	2484.9	50.40	0.579444	1.36013
5	2022	1113.97	3509.8	38.99	0.317388	1.22832
6	2376	1523.30	2066.7	53.32	0.737069	1.44334

## 8 Ligation of adapters

In folder 2016-10-04, I calculate the composition of the ligation reactions to have a 10-fold excess of adapters to DNA fragment ends. On table 5 I show the composition and final concentrations of the working solutions of the four adapters, required to prepare the ligation reactions specified on table 6. Importantly, I decide to assign adapters 1.1, 1.2, 1.3, and 1.4 to samples 1, 3, 5, and 6, respectively. See table 7.

Table 5: Composition and final concentration of the working solutions of adapters for the ligation reactions on table 6. The buffer must be  $1\times$  annealing buffer. Amounts are  $\mu\text{l}$ , except the final molarity, in the last column, in pmol/ $\mu\text{l}$ .

Sample	Adapter	$1\times$ buffer	15 $\mu\text{M}$ Stock	Total	pmol/ $\mu\text{l}$
1	1.1	8.580	1.420	10.00	2.13
3	1.2	8.840	1.160	10.00	1.74
5	1.3	8.362	1.638	10.00	2.46
6	1.4	9.036	0.964	10.00	1.45

Table 6: Volumes of reactants in the ligation reactions. All in  $\mu\text{l}$ . T4 ligase is at 2000000 U/ml.

Sample	DNA	10 $\times$ Buffer	Adapter	1.5 M NaCl	T4	H <sub>2</sub> O	Total
1	35	4.40	1.00	1.10	2.00	0.50	44.00
3	35	4.40	1.00	1.10	2.00	0.50	44.00
5	35	4.40	1.00	1.10	2.00	0.50	44.00
6	35	4.40	1.00	1.10	2.00	0.50	44.00

Table 7: Adapters and codewords assigned to the DNA samples.

Sample	Adapter	Codeword
1	1.1	TTGATCCAGT
3	1.2	GATCAGGCAGT
5	1.3	CCAGCTTGT
6	1.4	AGCTGAAT