Duplicate IDs

ID (times) sex generation

27409 (2) (M, M) F16

27485 (2) (F, F) F16

44463 (2) (F, F) F24

49705 (2) (F, F) F26

2050.1 (2) (F, M) F36

2021.1 (2) (F, M) F36

27409: duplication removed

27485: duplication removed

44463: duplication removed

49705: duplication removed

2050.1 is male, the female is 2005.1 – typo in the breeding record that transferred over to the line info file, which is now corrected. 2005.1 never produced offspring so no changes were made to AI Line Info\_37.

2021.1 is male, the female is 2012.1 – another type on the breeding record. Corrections were made for AI Line Info\_36 and 37.

5/10/2013 NMG

1. [1] "Generation: 3 Parents not present in previous generation: NA"
2. [1] "Generation: 4 Parents not present in previous generation: 1190"
3. [2] "Generation: 4 Parents not present in previous generation: 1182"
4. [3] "Generation: 4 Parents not present in previous generation: 1191"
5. [1] "Generation: 23 Parents not present in previous generation: 37731"
6. [2] "Generation: 23 Parents not present in previous generation: 38016"
7. [1] "Generation: 27 Parents not present in previous generation: 47401"
8. [2] "Generation: 27 Parents not present in previous generation: 47400"

1) Sire for F3 animals 1139 and 1140 missing (?) in AI Line Info\_3. The Dam is #135.

2) 3) 4) These animals are listed as being F2 rather than F3 in AI Line Info\_3. I don’t know whether this is the case or if it’s just a typo. The parent IDs for these animals are 35, 40, 44 and 45, so it seems like they really could be F2 individuals based on the sequence of animal IDs from F1-F3. Instead of editing the entire lineage I’m leaving this as-is, so there will always be an error message produced for these individuals. If the code is edited to output the generation column [,4], you would notice the discrepancy.

5) 6) These animals are the inbred dam (37731) and inbred sire (38016) of individual 41095, who is F1. I think they’re actually B6 because the strain is listed as 15o-C57B. 41095 (a female) was mated with 41841 to produce 1 male (44895) and 1 female (44899) listed as generation F24. Their offspring were used to create further lineages so there is nothing we can do about it at this point.

7) 8) These parents are listed in the F27 file but are actually F25s. So their offspring (51914 and 51916, females) are really F26.

More duplicates found after correcting the errors above:

In Breeder ped file through F56:

Lines 7070 7830 8194 8798

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Duplicated ID | 2nd position | 1st position | 2nd position parents | 1st position parents | Gen |
| 1 | 42059.1 | 7070 | 7011 | 40679-40619 | 40638-41336 | 47 |
| 2 | 45495.1 | 7830 | 7520 | 45471-45467 | 44048-44123 | 49, 50 |
| 3 | 49500.1 | 8194 | 8068 | 48454-48341 | 46031-46188 | 51, 52 |
| 4 | 52184.1 | 8798 | 8737 | 51912-51769 | 50082-51617 | 54,55 |

1 ) AI Line Info 47 file and Breeder Log have two different animals both coded as 42059; 1st is a black female and 2nd is a white male. Based on the breeding record it looks like (Diane?) tagged females with odd numbers and males with even numbers. The numbers are also in order. The true 42059 is the female. The male looks like it was either a typo or a misread of the ear tag. It should be 42050 (there are no other mice with this ID in F47). I have changed this in LGSM AIL (F47).xls and in the AI Line info file.

2 ) AI Line Info 49 file lists only one individual as 45495; its parents are the same as the mouse in position 1. AI Line Info 50 also lists a mouse with this id; it looks like a typo since all the other mice in this generation have IDs that start with 46---. I have changed the F50 mouse to ID 4**6**495 in the breeder log and in the AI LINE info 50 file.

3 ) The mouse in position 1 is in generation 51. The mouse in pos. 2 is in generation 52. I am changing the mouse in gen 51/pos.1 to 49000 in the breeder files and the info file.

4 ) The F54 mouse (pos.1) will remain as 52184. It looks like the mouse in F55 should be 52084 since its female sibling is 52085.

Oddly, 10 duplicates arise when I combine the Breeder ped with the Test Mouse ped.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| position | **ID** | sire | dam | sex |
| 7519 | **45494.1** | 44048.1 | 44123.1 | 2 |
| 9068 | **45494.1** | 45471.1 | 45467.1 | 1 |
| 7825 | **45496.1** | 45471.1 | 45467.1 | 1 |
| 9069 | **45496.1** | 45371.1 | 45301.1 | 1 |
| 7682 | **46017.1** | 45380.1 | 45340.1 | 2 |
| 9088 | **46017.1** | 45314.1 | 45385.1 | 2 |
| 7940 | **49251.1** | 46464.1 | 46039.1 | 2 |
| 9315 | **49251.1** | 46464.1 | 46039.1 | 2 |
| 8154 | **50345.1** | 48400.1 | 49235.1 | 2 |
| 9483 | **50345.1** | 48400.1 | 49235.1 | 2 |
| 8505 | **51360.1** | 50398.1 | 50363.1 | 1 |
| 9651 | **51360.1** | 50424.1 | 50287.1 | 1 |
| 8556 | **51737.1** | 51310.1 | 51325.1 | 2 |
| 9736 | **51737.1** | 50796.1 | 51305.1 | 2 |
| 8808 | **52180.1** | 51942.1 | 51817.1 | 1 |
| 9936 | **52180.1** | 51638.1 | 51279.1 | 1 |
| 8739 | **52182.1** | 51638.1 | 51279.1 | 1 |
| 9937 | **52182.1** | 51912.1 | 51769.1 | 1 |
| 8824 | **54165.1** | 51996.1 | 52037.1 | 2 |
| 9976 | **54165.1** | 51996.1 | 52037.1 | 2 |

1 ) 45494 in LineInfo49 file is female and is listed as a dam in the F50 file. 45494 is also the ID of a male in generation F50. It looks like a typo, like this male should really be 46494 because all of the mice in its generation have the prefix 46---. Unfortunately this error was perpetuated in all of my phenotype spreadsheets and records, so I’m going to leave it as is to prevent confusion. Instead I’m going to change the dam in LineInfo49 so that her ID is 454940.

2 ) This is a similar situation to (1). I am changing the ID of 45496 in generation F50 to 454960 in line info file.

3 ) 46017 is an F50 tester mouse and breeder mouse. The breeder’s ID is a typo; it should read 46107. I changed this in the Breeder Log and line info.

4 ) 49251 is an F51 test mouse and is also listed as a parent of the F52 generation. These are siblings. I’m changing the breeder’s ID to 49253 in the breeder log and line info.

5 ) 50345 is listed as an F52 test *and* breeder mouse, as in (4). I’m changing the breeder ID to 50347 in the breeder log and line info.

6 ) 51360 is listed as an F53 test and breeder mouse. In the case of the breeder, this is a typo (confirmed by the cage card F53-46) and should be 51630 instead. I’m changing this in the breeder logs and line info.

7 ) 51737 is listed as an F54 test and breeder mouse. The breeder mouse was another typo victim; the cage card confirms that her ID is 51837. Changing this in the breeder logs and line info.

8 ) **52180** is an F54 mouse and an F55 breeder. I am changing the F55 breeder ID to 52080 in the breeder logs. He doesn’t show up in the line info log so no change there.

9 ) 52182 is listed as an F55 test mouse and and F54 breeder mouse. It is an F54 breeder mouse. The test mouse should be 52082 instead - I think I had trouble reading the tag because the testing cage card reads 52082, then it has the 0 crossed out and turned into a 1. Since this error will be perpetuated in all of my test sheets and labels, I’m going to leave both of them as is in the breeder sheets, but in the line info file I’ll change the breeder 52182 to 521820, as with items (1) and (2).

10 ) **54165** is listed as an F55 test and breeder mouse. I am changing the breeder ID to 54167.

7/2/15 NMG

I was working in plot.haplotypes.R, extracting any dams and sires from the pedigree which appear in genotyped.samples.txt (expecting to see only animals from F39-43). However two sires from F54 came up (below: 50770 and 51348).

50770 is a typo. In the breeding records he was listed as 50776. I made this change in pedforQTLRel.txt.

51348 will be changed for 51349 in the genotyped.samples.txt file.

9325 51804.1 50770.1 51269.1 1 F54

9330 51811.1 50770.1 51269.1 2 F54

9504 51993.1 51348.1 51421.1 2 F54

9505 51994.1 51348.1 51421.1 1 F54