APL Scripts used in Liao, et al., Modularity and selection of nectar traits in the evolution of the selfing syndrome in *Ipomoea lacunosa* (Convolvulaceae)

Variable Definitions

```
  ΔMASTERDATA AND ΔMASTERDATA1

        COL1  TRAIT CODE
        COL2  TOTAL RELATIVE HOMOZYGOUS EFFECT (RHE)   (TOTAL
INCLUDES NEGATIVE RHE'S IN TOTAL)
        COL3  CI LOWER BOUND (CM)
        C0L4  CI UPPER BOUND (CM)
        COL5  IGNORE
        COL6  IGNORE
        COL7  IGNORE
        COL8  ADDITIVE EFFECT?
        COL9  RELATIVE HOMOZYGOUS EFFECT (CC-LL)
        COL10 NUMBER OF LINES WITH CC GENOTYPE
        COL11 NUMBER OF LINES WITH LL GENOTYPE
        COL12 UNIQUE ID FOR QTL

   ΔALLDATA1 AND ΔALLDATA2  FOR ALL QTLS AND GWS QTLS RESPECTI
VELY

        COL1  GENETIC CORRELATION (VARCOMP)
        COL2  PREDICTED GENETIC CORRELATION
        COL3  AVERAGE TOTAL RHE
        COL4  BIAS (= PREDICTED - OBSERVED GENETIC CORRELATIO
NS)
        COL5  QTL OVERLAP
        COL6  CLASS  (1 IF NONSEED, 0 IF SEED,  IF BETWEEN-MO
DULE TRAIT PAIR, 0 IF AT LEAST 1 TRAIT IS SEED TRAIT)
```

I. Permutation Test for Signiicance of Modules

A. Program SHUFFLECORRS

INPUT IS EITHER ΔOVERLAPCORR1 OR ΔOVERLAPCORR2, WHICH
  CORRESPOND TO GENETIC CORRELATIONS CALCULATED FROM RIL
   MEANS OR VARIANCE COMPONENTS

  COLUMNS OF ΔOVERLAPCORR1

  COL1 TRAIT NUMBER FOR TRAIT 1 OF PAIR (CORRESPONDS TO TRAIT
    IN ΔTRAITNAMES)
  COL2 TRAIT NUMBER FOR TRAIT 2 OF PAIR
  COL3 PAIRWISE TRAIT OVERLAP (FROM PROGRAM
    CALCQTLOVERLAP)
  COL4 PAIRWISE TRAIT GENETIC CORRELATIONS
  COL5 1 IF WITHIN-MODULE PAIR, 0 IF BETWEEN-MODULE PAIR

Code:

```
SHUFFLECORRS
⍝ THIS PROGRAM SHUFFLES CORRELATIONS ASSIGNED TO DIFFERENT MOD
   ULES AND DETERMINES PROBABILITY
⍝   THAT DIFFERENCE BETWEEN WITHIN- AND BETWEEN-MODULE AVERAGE
   S IS GREATER THAN THAT OBSERVED.

OBSCORRS←ΔOVERLAPCORR2[;4]          ⍝  ΔOVERLAPCORR1 HAS GENETI
   C CORRELATIONS CALCULATED BY RIL MEANS
OBSTYPE←ΔOVERLAPCORR2[;5]          ⍝  ΔOVERLAPCORR2 HAS GENETI
   C CORRELATIONS CALCULATED BY VARIANCE COMPONENTS

IND1←(OBSTYPE=1)/⍳⍴OBSTYPE      ⍝ INDEX FOR WITHIN-MODULE CORRE
   LATIONS
IND0←(OBSTYPE=0)/⍳⍴OBSTYPE      ⍝ INDEX FOR BETWEEN-MODULE CORR
   ELATIONS

NUM1←⍴IND1
NUM0←⍴IND0

OBSAVE1←(+/OBSCORRS[IND1])÷⍴IND1
OBSAVE0←(+/OBSCORRS[IND0])÷⍴IND0
OBSDIFF←OBSAVE1-OBSAVE0

⎕TCLF
'OBSERVED MEAN WITHIN-MODULE CORRELATION:  ',OBSAVE1
'OBSERVED MEAN BETWEEN-MODULE CORRELATION: ',OBSAVE0
⎕TCLF
'OBSERVED DIFFERENCE: ',OBSDIFF
```

```
SHUFFDIFFS←0ρ0

I←0
RETI:I←I+1
RAND←78?78
NEWTYPE←OBSTYPE[RAND]

IND1←(NEWTYPE=1)/ιρOBSTYPE        ⍝ INDEX FOR WITHIN-MODULE CORRE
    LATIONS
IND0←(NEWTYPE=0)/ιρOBSTYPE

AVE1←(+/OBSCORRS[IND1])÷ρIND1
AVE0←(+/OBSCORRS[IND0])÷ρIND0

SHUFFDIFF←AVE1-AVE0

SHUFFDIFFS←SHUFFDIFFS,SHUFFDIFF
→(I<1000)/RETI

PROP←(+/SHUFFDIFFS≥OBSDIFF)÷1000
⎕TCLF
'PROPORTION OF SHUFFLED DIFFERENCES GREATER THAN OBSERVED: ',P
ROP
⎕TCLF
'PROGRAM SHUFFLECORRS FINISHED.  VECTOR OF SHUFFLED DIFFERENCE
S IN VARIABLE ''SHUFFDIFFS''.'
```

II. QTL overlap permutation test

A. Program CALCMODOVERLAP

PROGRAM CALCMODOVERLAP CALCULATES QTL OVERLAP FOR A MODULE AND
  PERFORMS RANDOM ASSIGNMENT TEST TO DETERMINE WHETHER
  AVERAGE QTL OVERLAP FOR A MODULE IS GREATER THAN EXPECTED

INPUT VARIABLES: ΔMODULE1A, ΔMODULE2A, ΔMODULE3A FOR LESS-
  CONSERVED SET OF QTLS ("ALL QTLs"),
  ΔMODULE1B, ΔMODULE2B, ΔMOULDE3B FOR CONSERVED SET OF QTLS ("GWS
  QTLs")

 INPUT IS AN N X 6 MATRIX OF QTL POSITIONS FOR ALL TRAITS IN A MODULE
   COL1: TRAIT NUMBER (1- X)
   COL2: LINKAGE GROUP
   COL3: LOWER BOUND OF 1.5 LOD CONF INTERVAL (IN CM)
   COL4: UPPER BOUND OF 1.5 LOD CONF INTERVAL (IN CM)
   COL5: LOWER BOUND OF 1.5 LOD CI (IN BP)
   COL6: UPPER BOUND OF 1.5 LOD CI (IN BP)

  OUPTUT IS 1000 AVERAGE OVERLAPS FOR RANDOM ASSIGNMENT OF QTLS TO
    CHROMOSOME POSITIONS
  PROPORTION OF AVERAGES GREATER THAN OBSERVED AVERAGE OVERLAP IS
    ALSO OUTPOUT

PROGRAM RANDOMIZEPOSITIONS CREATES NEW ΔMODULEXA WITH POSITIONS
ON CHROMOSOMES RANDOMIZED FOR EACH QTL IN ΔMODULEXA
  POSITIONS RANDOMIZED USING QTL 1.5 LOD CI IN BP
  (CALLED BY CALCMODOVERLAP)

  NOTE: THIS PROGRAM USES VARIABLE ΔGENEDENSITY, WHICH HAS THREE
COLUMNS:

   COL1 LINKAGE GROUP
   COL2 NUMBER OF 50KB WINDOW
   COL3 NUMBER OF GENES IN WINDOW

ADDITIONAL INPUT VARIABLE IS ΔMASTERDATA OR ΔMASTERDATA1:

   COL1 TRAIT CODE
   COL2 TOTAL RELATIVE HOMOZYGOUS EFFECT (RHE)
         (TOTAL INCLUDES NEGATIVE RHE'S IN TOTAL)
   COL3 CI LOWER BOUND (CM)
   C0L4 CI UPPER BOUND (CM)
   COL5 IGNORE

COL6 IGNORE
COL7 IGNORE
COL8 ADDITIVE EFFECT
COL9   RELATIVE HOMOZYGOUS EFFECT (CC-LL)
COL10 NUMBER OF LINES WITH CC GENOTYPE
COL11 NUMBER OF LINES WITH LL GENOTYPE
COL12 UNIQUE ID FOR QTL

CALLS PROGRAMS:

CONSTRUCTMODOVERLAP
RANDOMIZEPOSITIONS

Code:

```
CALCMODOVERLAP X
⍝ THIS PROGRAM CALCULATES QTL OVERLAPS WITHIN A MODULE FOR JO
  ANNA'S DATA AND PERFORMS BOOTSTRAP ON THEM.

⎕TCLF
'ENTER MODULE NUMBER'
MOD←⎕
DELETE←0 ⍝ 7 MEANS DELETE NOTHING
TRAITQTL←⍋MASTERDATA[;2 3 4]
TRAITCODES←⍋MASTERDATA[;1]
IND←(~TRAITCODES∊(2.4,2.5))/⍳⍴TRAITCODES
TRAITQTL←TRAITQTL[IND;]
TRAITCODES←TRAITCODES[IND]

⍝ GWINDEX←(⍋GENOMEWIDESIGNIFICANT)/⍳⍴⍋GENOMEWIDESIGNIFICANT
⍝ INDEX OF QTL'S THAT ARE SIGNIFICANT GENOME WIDE
⍝ TRAITQTL←TRAITQTL[GWINDEX;]
⍝ TRAITCODES←TRAITCODES[GWINDEX]

⍝ EXECUTE ABOVE THREE STATEMENTS IF WANT TO DO ANALYSIS USING
    JUST GENOMEWIDE QTLS

CONSTRUCTMODOVERLAP X

OBSAVEOVERLAP←+/+/⍋OVERLAPMATALL÷(DIM×(DIM-1))
OBSOVERLAPMATALL←⍋OVERLAPMATALL

⎕TCLF
'STARTING RANDMIZATION TEST'
MODAVEOVERLAP←0⍴0

LL←0
RETLL:LL←LL+1
```

```
RANDOMIZEPOSITIONS

CONSTRUCTMODOVERLAP  NEWTRAITQTL
AVEOVERLAP←+/+/∆OVERLAPMATALL÷(DIM×(DIM-1))
MODAVEOVERLAP←MODAVEOVERLAP,AVEOVERLAP

→(LL<1000)/RETLL

□TCLF
'MODULE',(⍕MOD),' RESULTS'
□TCLF
'OBSERVED AVERAGE OVERLAP = ',OBSAVEOVERLAP
□TCLF
PROP←(+/MODAVEOVERLAP≥OBSAVEOVERLAP)÷1000
'PROPORTION OF RANDOMIZED TRIALS WITH AVERAGE OVERLAP GREATER
THAN OBSERVED = ',PROP
□TCLF

□TCLF
'PROGRAM CALCMODOVERLAP FINISHED. DELETE = ',DELETE
'  OUTPUT IN VARIABLES ''MODAVEOVERLAPP'', ''OBSSEOVERLAP''.'
```

B. Program CONSTRUCTMODOVERLAP

PROGRAM CONSTRUCTMODOVERLAP MAKES MATRIX OF OVERLAP
PROPORTIONS FOR TRAITS WITHIN A MODULE (CALLED BY
CALCMODOVEDRLAP)

PROGRAM CALCMODOVERLAP2 DOES SAME THING BUT AVERAGES OVER ALL
MODULES
   INPUTS ARE ∆MODULE1 - ∆MODULE5

   CALLED FROM: CALCMODOVERLAP
   Output: ∆OVERLAPMATALL

Code:

```
CONSTRUCTMODOVERLAP Y
⍝ THIS PROGRAM CONSTRUCTS QTL OVERLAP MATRIX FOR ALL QTLS
⍝ ENTRY I,J = 0 IF NO OVERLAP, =1 IF OVERLAP
QTLNUMS←⍳⍴TRAITCODES

TRAITQTL←Y
TRAITCODES←Y[;1]

TRSH←TRAITCODES
TRSH2←¯1⌽TRAITCODES
```

6

```
UNIQUE←(TRSH≠TRSH2)/TRAITCODES
NUNIQUE←ρUNIQUE
DIM←ρUNIQUE

SIZE←ρTRAITCODES        ⍝ TOTAL NUMBER OF QTLS
NUMMAT1←NUMMAT2←MAT←(DIM,DIM)ρ0     ⍝ INITIALIZE OVERLAP MATRIX
⍋OVERLAPMATALL←(DIM,DIM)ρ0

LINES2←0 8ρ0

I←0
RETI:I←I+1          ⍝ QTL1 LOOP

J←I
RETJ:J←J+1          ⍝ QTL2 LOOP
→(J>SIZE)/DOWN
LINEI←TRAITQTL[I;]  ⍝ GET QTL DATA FOR QTL1  (CHROMOSOME, CI L
    OW, CI HIGH)
LINEJ←TRAITQTL[J;]  ⍝ GET QTL DATA FOR QTL2

TEST1←LINEI[1]=LINEJ[1]  ⍝ 1 IF BOTH QTLS FOR SAME TRAIT
→(TEST1=1)/RETJ

TEST1←LINEI[2]=LINEJ[2]     ⍝ QTL1 AND QTL2 ON SAME CHROMOSOME?
→(TEST1=0)/RETJ             ⍝ IF NOT, SKIP AND LEAVE ZERO IN  M
    ATRIX

TEST←((LINEI[6]=LINEJ[5])∨(LINEJ[6]=LINEI[5]))
→(TEST=1)/RETJ
CRIT1←(LINEI[5]≥LINEJ[5]) ∧ (LINEI[5]≤LINEJ[6])      ⍝ THESE TW
    O LINES TEST WHETHER THERE IS OVERLAP
CRIT2←(LINEI[6]≥LINEJ[5]) ∧ (LINEI[6]≤LINEJ[6])
CRIT3←(LINEJ[5]≥LINEI[5]) ∧ (LINEJ[5]≤LINEI[6])
CRIT4←(LINEJ[6]≥LINEI[5]) ∧ (LINEJ[6]≤LINEI[6])

TEST2←CRIT1∨CRIT2∨CRIT3∨CRIT4

→(TEST2=0)/RETJ

LINES2←LINES2,[1](LINEI[⍳4],LINEJ[⍳4])

CODE1←TRAITCODES[I]
CODE2←TRAITCODES[J]
INDEX1←(UNIQUE=CODE1)/⍳ρUNIQUE
INDEX2←(UNIQUE=CODE2)/⍳ρUNIQUE

NUM1←+/(TRAITCODES=CODE1)
NUM2←+/TRAITCODES=CODE2
                        ⍝ TEST2 =1 IF THERE IS OVERLAP
```

```
♣(TEST2=1)/'MAT[INDEX1;INDEX2]←MAT[INDEX2;INDEX1]←MAT[INDEX2;I
NDEX1]+1'                          ⍝ IF THERE IS OVERLAP, ENTER 1 IN
   MAT FOR ENTRIES I,J AND J,I
♣(TEST2=1)/'NUMMAT1[INDEX1;INDEX2]←NUMMAT1[INDEX2;INDEX1]←NUM1
'
♣(TEST2=1)/'NUMMAT2[INDEX1;INDEX2]←NUMMAT2[INDEX2;INDEX1]←NUM2
'
→RETJ

DOWN:→(I<(SIZE-1))/RETI

SIZE2←1↑ρMAT

K←0
RETK:K←K+1

L←K
RETL:L←L+1

→(MAT[K;L]=0)/DOWN2

AVE←(MAT[K;L])÷(NUMMAT1[K;L]+NUMMAT2[K;L]-MAT[K;L])     ⍝   THIS
   IS THE JACCARD INDEX
ΔOVERLAPMATALL[K;L]←ΔOVERLAPMATALL[L;K]←AVE

DOWN2:→(L<SIZE2)/RETL
→(K<(SIZE2-1))/RETK

→0

□TCLF
'PROGRAM CONSTRUCTOVERLAP FINISHED. TRAITCODE ',(⍕DELETE),'. M
ATRIX IN VARIABLE ''MAT''.'
```

C. Program RANDOMIZEPOSITIONS

PROGRAM RANDOMIZEPOSITIONS CREATES NEW ΔMODULEXA WITH POSITIONS
ON CHROMOSOMES RANDOMIZED FOR EACH QTL IN ΔMODULEXA
   POSITIONS RANDOMIZED USING QTL 1.5 LOD CI IN BP
   (CALLED BY CALCMODOVERLAP)

   NOTE: THIS PROGRAM USES VARIABLES ΔCHROMDENSITIES AND
ΔCUMDENSITY, PRODUCED BY PROGRAM CALCGENDENSITIES

   CALLED FROM: CALCMODOVERLAP
   Output: NEWTRAITQTL

Code:

```
RANDOMIZEPOSITIONS
A THIS PROGRAM RANDOMIZES THE POSITIONS OF THE QTLS IN A GIVEN
  MODULE

NEWTRAITQTL←0 6ρ0

SIZE←1↑ρTRAITQTL
MAXII←SIZE
II←0
RETII:II←II+1
LINE←TRAITQTL[II;]      A CHOOSE QTL II
NEWLINE←LINE[1]
MAXDEN←∆CHROMDENSITIES[15]         A∆CHROMDENSITIES IS VECTOR
    OF CUMULATIVE CHROMOSOME SIZES TOTAL GENES
RAND←?MAXDEN
LINKGRP←1+(+/RAND≥∆CHROMDENSITIES )
⍟(RAND=MAXDEN)/'LINKGRP←15'

NEWLINE←NEWLINE,LINKGRP

QTLSIZE←LINE[6]-LINE[5]    A SIZE OF QTL 1.5 LOD INTERVAL IN BP
QTLHALFSIZE←QTLSIZE÷2

IND←(∆CUMDENSITY[;1]=LINKGRP)/⍳1↑ρ∆CUMDENSITY
PART←∆CUMDENSITY[IND;]
MAX←⌈/PART[;3]
RAND2←?MAX
WINDOW←1+(+/PART[;3]≤RAND2)
BASEPOS←(WINDOW-1)×50000
RAND3←?50000
POSITION←BASEPOS+RAND3
A POSITION←?(⌈∆LGSIZESBP[LINKGRP])     A CHOSE POSITION OF QTL
ON NEW LG

LOWER←POSITION-QTLHALFSIZE
UPPER←POSITION+QTLHALFSIZE

⍟(LOWER<0)/'POSITION←POSITION-LOWER'
⍟(UPPER>∆LGSIZESBP[LINKGRP])/'DIFF←UPPER-∆LGSIZESBP[LINKGRP] ◊
 POSITION←POSITION-DIFF'
LOWER←POSITION-QTLHALFSIZE
UPPER←POSITION+QTLHALFSIZE

⍟(LOWER<0)/'LOWER←0'

NEWLINE←NEWLINE,LINE[3 4],LOWER,UPPER
NEWTRAITQTL←NEWTRAITQTL,[1]NEWLINE
→(II<MAXII)/RETII
```

D. Program CALCGENEDENSITIES

INPUT VARIABLE ∆GENEDENSITY HAS THREE COLUMNS:

    COL1 LINKAGE GROUP
    COL2 NUMBER OF 50KB WINDOW
    COL3 NUMBER OF GENES IN WINDOW

PRODUCES TWO VARIABLES:

    ∆CHROMDENSITIES: VECTOR WITH 15 ELEMENTS, EACH CORRESPONDING TO TOTAL NUMBER OF GENES ON CHROMOSOMES 1-15

    ∆CUMDENSITY HAS 3 COLUMNS:

    COL1 LINKAGE GROUP
    COL2 NUMBER OF 50KB WINDOW
    COL3 CUMULATIVE NUMBER OF GENES ON CHROMOSOME

Code:

```
CALCGENEDENSITIES
⍝  THIS PROGRAM INPUTS ∆GENEDENSITY AND CALCULATES CUMULATIVE
   GENE DENSITIES PER CHROMOSOME

ROWS←1↑⍴∆GENEDENSITY
∆CUMDENSITY←0 3⍴0
∆CHROMDENSITIES←0⍴0
CUMDEN←0⍴0
I←0
RETI:I←I+1

IND←(∆GENEDENSITY[;1]=I)/⍳ROWS
PART←∆GENEDENSITY[IND;]
LINE←PART[1;]
PREVCUM←LINE[3]
∆CUMDENSITY←∆CUMDENSITY,[1]LINE

MAXJ←1↑⍴PART
J←1
RETJ:J←J+1
LINE←PART[J;]

TEMP←LINE[2]
TEST←TEMP≠⌈TEMP
→(TEST=1)/DOWN

LINE[3]←LINE[3]+PREVCUM
PREVCUM←LINE[3]
```

10

```
△CUMDENSITY←△CUMDENSITY,[1]LINE
DOWN:→(J<MAXJ)/RETJ

TEMP←¯1↑△CHROMDENSITIES
△CHROMDENSITIES←△CHROMDENSITIES,(TEMP+PREVCUM)

→(I<15)/RETI

□TCLF
'PROGRAM CALCTGENEDENSITIES FINISHED.  OUTPUT IN VARIABLES ''△
CUMDENSITY'' AND ''△CHROMDENSITIES''.'
```

III. Predicting genetic correlations from QTL overlap

PROGRAM PREDICTCORR DOES THIS. INPUT FROM ΔMASTERDATA OR ΔMASTERDATA1 FOR LESS CONSERVATIVE ("ALL QTLs") AND CONSERVATIVE ("GWS QTLs")QTL SETS

ADDITIONAL INPUT: ΔMIDPOINTS

COL1: QTL1 ID
COL2: QTL2 ID
COL3: LINKAGE GROUP
COL4: MIDPOINT OF OVERLAP INTERVAL (CM)
COL5: NEAREST MARKER POSITION (CM)
COL6: NUMBER OF LINES WITH CC GENOTYPE AT MARKER POSITION
COL7: NUMBER OF LINES WITH LL GENOTYPE AT MARKER POSITION

PROGRAM CALLS CALCCOV2 WHICH CALCULATES COV AND VARIANCES CONTRIBUTED BY QTLS FROM A PAIR OF TRAITS, THEN SUMS TO GET PREDICTED TOTAL GENETIC COVARIANCE AND GENETIC VARIANCES FOR THE PAIR OF TRAITS.

PROGRAM PRODUCES VARIABLE ΔGENCORRS' (ORDER OF TRAITS AS IN ΔTRAITNAMES)

PROGRAM TRIANGLE RUN ON ΔGENCORRS TO OBTAIN LINEAR VECTOR OF CORRELATIONS

THIS VECTOR WAS COMBINED WITH ΔGENCORRS2 TO PRODUCE A MATRIX WITH TWO COLUMNS:

COL1: GENETIC CORRELATION BETWEEN PAIRS OF TRAITS (COVAR- AND VAR-COMP CORRELATIONS)
COL2: PREDICTED GENETIC CORRELATION FROM GENCORRS
COL3: BIAS (COL2 - COL1)
COL4: AVERAGE TOTAL RHE FOR EACH PAIR OF TRAITS
COL5: QTL OVERLAP (JACCARD INDEX) FOR EACH PAIR OF TRAITS

COMPARISON OF PREDICTED WITH OBSERVED CORRELATIONS ANALYZED BY SAS PROGRAM PREDICTCORR IN WORKSPACE 'IRENE'

PROGRAM SHUFFLEPRED TESTS SIGNIFICANCE OF CORRELATIONS FROM PREDICTCORR

INPUTS ARE ΔMASTERDATA OR ΔMASTERDATA1 AND ΔMIDPOINTS OR ΔMIDPOINTS2

PROGRAM PERMUTES TRAITS FOR INDIVIDUAL QTLS

A. Program PREDICTCORR

   CALLS: CALCCORR, TRIANGLE, CALCCOV2

Code:

```
PREDICTCORR
⍝ THIS PROGRAM PREDICTS PAIRWISE TRAIT GENETIC CORRELATIONS FR
   OM QTL OVERLAP AND EFFECT SIZES

MAXJ←ρUNIQCODES
MAXI←MAXJ-1

MIDPOINTS←0 4ρ0    ⍝ THIS VARIABLE HOLDS MIDPOINTS OF OVERLAP I
   NTERVALS FOR QTLS THAT OVERLAP
                   ⍝ FIRST TWO COLUMNS ARE QTL UNIQUE ID, THIRD
   IS LG AND FOURTH IS MIDPOINT

GENCORRS←((ρUNIQCODES),(ρUNIQCODES))ρ0
AVETRHES←((ρUNIQCODES),(ρUNIQCODES))ρ0
OVERLAPMAT←((ρUNIQCODES),(ρUNIQCODES))ρ0

I←0
RETI:I←I+1  ⍝  FIRST TRAIT LOOP

J←I
RETJ:J←J+1  ⍝  SECOND TRAIT LOOP

TRAITCODEI←UNIQCODES[I]
TRAITCODEJ←UNIQCODES[J]

TOTRHES←,⍋TOTRHEA[;1]
INDI←(TRAITCODEI=TOTRHES)/ιρUNIQCODES
INDJ←(TRAITCODEJ=TOTRHES)/ιρUNIQCODES
TOTRHEI←⍋TOTRHEA[INDI;2]
TOTRHEJ←⍋TOTRHEA[INDJ;2]
AVETOTRHE←(TOTRHEI+TOTRHEJ)÷2

IND1←(TRAITCODES=UNIQCODES[I])/ιρTRAITCODES
PART1←QTLDATA[IND1;]
IND2←(TRAITCODES=UNIQCODES[J])/ιρTRAITCODES
PART2←QTLDATA[IND2;]

OVERLAPS←0 4ρ0    ⍝ NUMBERS ARE QTLNUMBERS FROM PARTS 1 AND 2 F
   OR OVERLAPPING QTL

MAXII←1↑ρPART1
MAXJJ←1↑ρPART2
```

```apl
II←0
RETII:II←II+1      ⍝ PART1 LOOP

JJ←0
RETJJ:JJ←JJ+1      ⍝ PART2 LOOP

LINEII←PART1[II;] ⍝ GET QTL DATA FOR QTL1  (CHROMOSOME, CI LOW
    , CI HIGH)
LINEJJ←PART2[JJ;]  ⍝ GET QTL DATA FOR QTL2

TEST1←LINEII[2]=LINEJJ[2]   ⍝ QTL1 AND QTL2 ON SAME CHROMOSOME?

→(TEST1=0)/DOWN              ⍝ IF NOT, SKIP AND LEAVE ZERO IN
    MATRIX

TEST←((LINEII[4]=LINEJJ[3])∨(LINEJJ[4]=LINEII[3]))
→(TEST=1)/DOWN              ⍝ DO NOT COUNT AS OVERLAP IF CI'S ABUT
CRIT1←(LINEII[3]≥LINEJJ[3]) ∧ (LINEII[3]≤LINEJJ[4])      ⍝ THES
    E TWO LINES TEST WHETHER THERE IS OVERLAP
CRIT2←(LINEII[4]≥LINEJJ[3]) ∧ (LINEII[4]≤LINEJJ[4])
CRIT3←(LINEJJ[3]≥LINEII[3]) ∧ (LINEJJ[3]≤LINEII[4])
CRIT4←(LINEJJ[4]≥LINEII[3]) ∧ (LINEJJ[4]≤LINEII[4])

TEST2←CRIT1∨CRIT2∨CRIT3∨CRIT4
→(TEST2=0)/DOWN

OVERLAPS←OVERLAPS,[1](II,JJ,LINEII[12],LINEJJ[12])


TEMP←LINEII[3 4],LINEJJ[3 4]               ⍝ THIS NEXT SET OF 6 LINE
    S CALCULATES MIDPOINTS OF OVERLAP INTERVALS AND STORES IN V
    ARIABLE MIDPOINTS
TEMP2←⍋TEMP
TEMP3←TEMP[TEMP2]
INTERVAL←TEMP3[2 3]
MIDPOINT←(+/INTERVAL)÷2
MIDPOINTS←MIDPOINTS,[1](LINEII[12],LINEJJ[12],LINEII[2],MIDPOI
NT)

DOWN:→(JJ<MAXJJ)/RETJJ

→(II<MAXII)/RETII

⍝ ⎕TCLF
⍝ 'CALCULATING OVERLAPS FINISHED FOR I,J = ',I,J

MAXK←1↑⍴OVERLAPS

⍝ FOLLOWING VARIABLE HAS COLUMNS: COL1 RHE ON TRAITI  RHE ON T
RAITJ
RHE←0 4⍴0
```

14

```
SIZE←1↑ρOVERLAPS
→(SIZE=0)/DOWN4

K←0
RETK:K←K+1 ⍝ THIS LOOP ADDS RHE FOR OVERLAPPING QTLS

RHEVALS←(PART1[OVERLAPS[K;1];9],PART2[OVERLAPS[K;2];9],OVERLAP
S[K;3 4])

RHE←RHE,[1]RHEVALS

→(K<MAXK)/RETK

DOWN4:MAXL←1↑ρPART1
L←0
RETL:L←L+1

TEST←L∈OVERLAPS[;1]
→(TEST=1)/DOWN2

RHE←RHE,[1](PART1[L;9],0,PART1[L;12],0)

DOWN2:→(L<MAXL)/RETL

MAXLL←1↑ρPART2
LL←0
RETLL:LL←LL+1

TEST←LL∈OVERLAPS[;2]
→(TEST=1)/DOWN3

RHE←RHE,[1](0,PART2[LL;9],0,PART2[LL;12])

DOWN3:→(LL<MAXLL)/RETLL

CALCCOV2

GENCORRS[I;J]←R
AVETRHES[I;J]←AVETOTRHE

IND1←((RHE[;1]≠0)∧(RHE[;2]≠0))/⍳1↑ρRHE
A←ρIND1
B←+/RHE[;1]≠0
C←+/RHE[;2]≠0
OV←A÷(B+C-1)
OVERLAPMAT[I;J]←OV

→(J<MAXJ)/RETJ
→(I<MAXI)/RETI

TRIANGLE GENCORRS
```

```
♣(ALLQTLS≠'N')/'DATA←ΔGENCORRS2,78 1ρVEC'
♣(ALLQTLS='N')/'DATA←ΔGENCORRS2[NONSEEDIND;],45 1ρVEC'
BIAS←DATA[;2]-DATA[;1]
♣(ALLQTLS≠'N')/'DATA←DATA,78 1ρBIAS'
♣(ALLQTLS='N')/'DATA←DATA,45 1ρBIAS'
TRIANGLE AVETRHES
♣(ALLQTLS≠'N')/'DATA←DATA,78 1ρVEC'
♣(ALLQTLS='N')/'DATA←DATA,45 1ρVEC'

TRIANGLE OVERLAPMAT
♣(ALLQTLS≠'N')/'DATA←DATA,78 1ρVEC'
♣(ALLQTLS='N')/'DATA←DATA,45 1ρVEC'

CALCCORR DATA[;1 2]
OBSCORR1←CORR
CALCCORR DATA[;1 3]
OBSCORR2←CORR
CALCCORR DATA[;3 4]
OBSCORR3←CORR

CALCCORR DATA[;2 5]
OBSCORR4←CORR

□TCLF
'CORRELATION OF GENETIC CORRELATIONS AND PREDICTED CORRELATION
S: ',OBSCORR1
'CORRELATION OF BIAS AND GENETIC CORRELATIONS: ',OBSCORR2
'CORRELATION OF AVERAGE TOTAL RHE AND BIAS: ',OBSCORR3
'CORRELATION OF PREDICTED GENETIC CORRELATIONS AND QTL OVERLAP
: ',OBSCORR4
□TCLF


'PROGRAM PREDICTCORR FINISHED.  OUTPUT IN VARIABLES ''GENCORRS
'' AND ''AVETRHES''.'
```

B. Program CALCCORR

Called by PREDICTCORR

<u>Code</u>:

```
CALCCORR  X
ค THIS PROGRAM CALCULATES CORRELATION BETWEN TWO COLUMNS IN X
ROWS←1↑ρX
MEANS←+/[1]X÷ROWS
X1←,X[;1]
X2←,X[;2]
DIFF1←X1-MEANS[1]
```

```
DIFF2←X2-MEANS[2]
COV←+/(DIFF1×DIFF2)÷ROWS
VAR1←+/(DIFF1★2)÷ROWS
VAR2←+/(DIFF2★2)÷ROWS
CORR←COV÷((VAR1×VAR2)★.5)
⍝ ⎕TCLF
⍝ 'THE CORRELATION IS ',CORR
```

C. Program TRIANGLE

CALLED BY PREDICTCORR

Code:

```
TRIANGLE X
⍝ THIS PRPGRAM MAKES A VECTOR OUT OF THE UPPER TRIANGLE (NOT I
    NCLUDING DIAGONAL) OF A MATRIX

VEC←0⍴0
MAXI←¯1+1↑⍴X
I←0
RETI:I←I+1

MAXJ←1↑⍴X
J←I
RETJ:J←J+1
VEC←VEC,X[I;J]

→(J<MAXJ)/RETJ

→(I<MAXI)/RETI
```

D. Program CALCCOV2

Called from PREDICTCORR

Code:

```
CALCCOV2
⍝  THIS PROGRAM CALCULATES GENETIC COVARIANCE AND VARIANCES FO
    R A PAIR OF TRAITS TAKING GENOTYPE FREQS INTO ACCOUNT
⍝  INPUT IS RHE FROM PROGRAM  'PREDICTCORR'

KKMAX←1↑⍴RHE
```

```
PAIRS←∆MIDPOINTSA[;1 2]        ⍝ VARIABLE WITH ALL PAIRS OF QTLS
    THAT OVERLAP; NUMBERS ARE QTL IDS; ∆MIDPOINTS FOR ALL QTL,
    ∆MIDPOINTS1 FOR GWS QTL

VARCOV←0 3⍴0

KK←0
RETKK:KK←KK+1
LINE←RHE[KK;]
EFFECTS←0.5×LINE[1 2]

TEST←(EFFECTS[1]≠0)∧(EFFECTS[2]≠0)        ⍝ TEST FOR WHETHER OVE
    RLAPPING QTLS

COV←0

→(TEST=0)/DOWN
⍝ CALCULATE COVARIANCE FOR QTL PAIR
PAIR←2 1⍴LINE[3 4]    ⍝ THIS IS THE PAIR OF QTL UNIQUE IDS

IND←(,PAIRS∧.=PAIR)/⍳1↑⍴PAIRS        ⍝ ROW INDEX OF QTL PAIR IN
    ∆MIDPOINTS
TEST←0=⍴IND
⍎(TEST=1)/'PAIR←2 1⍴LINE[4 3] ◊ IND←(,PAIRS∧.=PAIR)/⍳1↑⍴PAIRS
'

NUMS←,∆MIDPOINTSA[IND;6 7]
        ⍝ VECTOR OF NUMBER OF LINES WITH CC AND WITH LL GENOTYPE

PROPCC←NUMS[1]÷(NUMS[1]+NUMS[2])    ⍝ PROPORTION OF LINES WITH
    CC GENOTYPE

MEAN1←(PROPCC×EFFECTS[1])-((1-PROPCC)×EFFECTS[1])  ⍝ MEAN EFFE
    CT FOR QTL1
MEAN2←(PROPCC×EFFECTS[2])-((1-PROPCC)×EFFECTS[2])  ⍝ MEAN EFFE
    CT FOR QTL2

PROD1←PROPCC×(EFFECTS[1]-MEAN1)×(EFFECTS[2]-MEAN2)
            ⍝ CALCULATE COVARIANCE
PROD2←(1-PROPCC)×((¯1×EFFECTS[1])-MEAN1)×((¯1×EFFECTS[2])-MEAN
2)
COV←PROD1+PROD2

DOWN:

⍝ CALCULATE VARIANCES
VARI←VARJ←0
PAIR2←,PAIR
→(EFFECTS[1]=0)/DOWN2            ⍝ SKIP FOR QTL1 IF EFFECT IS 0
⍝ CALCULATE VARIANCE FOR QTL1
```

```
IND←(,PAIR2[1]=QTLDATA[;12])/ι1↑ρQTLDATA
NCC←QTLDATA[IND;10]
NLL←QTLDATA[IND;11]
PROPCC1←NCC÷(NCC+NLL)

MEAN3←(PROPCC1×EFFECTS[1])-((1-PROPCC1)×EFFECTS[1])
VARI←(PROPCC1×((EFFECTS[1]-MEAN3))⋆2)+((1-PROPCC1)×(((¯1×EFFEC
TS[1])-MEAN3))⋆2)

DOWN2:  →(EFFECTS[2]=0)/DOWN3   ⍝ SKIP FOR QTL2 IF EFFECT IS 0

⍝ CALCULATE VARIANCE FOR QTL2

IND←(,PAIR2[2]=QTLDATA[;12])/ι1↑ρQTLDATA
NCC←QTLDATA[IND;10]
NLL←QTLDATA[IND;11]
PROPCC2←NCC÷(NCC+NLL)

MEAN4←(PROPCC2×EFFECTS[2])-((1-PROPCC2)×EFFECTS[2])
VARJ←(PROPCC2×((EFFECTS[2]-MEAN4))⋆2)+((1-PROPCC2)×(((¯1×EFFEC
TS[2])-MEAN4))⋆2)

DOWN3:

VARCOV←VARCOV,[1](COV,VARI,VARJ)

→(KK<KKMAX)/RETKK

SUM←+/[1]VARCOV
R←SUM[1]÷((SUM[2]×SUM[3])⋆.5)
```

E. Program SHUFFLEPRED

Calls PREDICTCORR, PREDICTCORR2

Code:

```
SHUFFLEPRED
⍝ THIS PROGRAM SHUFFLES QTLS TO CALCULATE PROBABILITY OF CORRE
LATION BETWEEN GENETIC CORRELATIONS AND PREDICTED CORRELATIONS

⍝ CAN RUN THIS FOR COMBINATIONS OF GWS OR ALL QTLS AND ALL TRA
   ITS OR JUST FLOWER AND NECTAR TRAITS
⍝ CHANGE TITLE APPROPRIATELY BELOW
⍝ IN PREDICTCORR2, INDICATE ALLQTLS = 'Y' OR 'N'    (NO MEANS
   JUST FLOWER AND NECTAR TRAITS)
```

```
⍝ IN PREDICTCORR2 USE EITHER ∆MIDPOINTS OR ∆MIDPOINTS1, ∆MASTE
    RDATA OR ∆MASTERDATA1, AND ∆TOTRHE OR ∆TOTRHE1, FOR LESS OR
    MORE CONSERVED QTLS
⍝ IN PREDICTCORR2 USE EITHER ∆MIDPOINTS OR ∆MIDPOINTS1

ALLQTLS←'N'     ⍝  THIS VARIABLE SHOULD REALLY BE NAMED 'ALLTRA
    ITS'
TITLE←'GWS QTLS, FLORAL AND NECTAR TRAITS'

NONSEEDIND←1 2 3 4 5 6 10 11 12 13 14 15 16 17 21 22 23 24 25
 26 27 31 32 33 34 35 36 40 41 42 43 44 48 49 50 51 55 56 57 61
 62 63 76 77 78
∆MIDPOINTSA←∆MIDPOINTS1      ⍝ USE ∆MIDPOINTS FOR LESS-CONSERVA
    TIVE QTLS, ∆MIDPOINTS1 FOR GWS QTLS
∆TOTRHEA←∆TOTRHE1             ⍝ USE ∆TOTRHE FOR LESS-CONSERVATIV
    E QTLS, ∆TOTRHE1 FOR GWS QTLS
QTLDATAA←QTLDATA←∆MASTERDATA1      ⍝ USE ∆MASTERDATA FOR LES
    S-CONSERVATIVE QTL SET, ∆MASTERDATA1 FOR CONSERVATIVE DATAS
    ET
TRAITCODESA←TRAITCODES←∆MASTERDATA1[;1]       ⍝  AND USE ∆MA
    STERDATA FOR LESS-CONSERVATIVE QTL SET, ∆MASTERDATA1 FOR CO
    NSERVATIVE DATASET

IND←(~TRAITCODES∊(2.4,2.5))/⍳⍴TRAITCODES   ⍝ LIST TRAITS TO EXC
    LUDE   NOTE: 2.4 AND 2.5 ARE ALTERNATES FOR 2.3 (NECTARY SIZ
E)

⍙(ALLQTLS='N')/'IND←(~TRAITCODES∊(3.1,3.2,3.3,2.4,2.5))/⍳⍴TRAI
TCODES'  ⍝ USE THIS FOR JUST FLOWER AND NECTAR TRAIT ANALYSIS

TRAITCODES←TRAITCODES[IND]
QTLDATA←QTLDATA[IND;]
QTLDATAB←QTLDATA

UNIQCODES←1.1,1.2,1.3,1.4,1.5,1.6,1.7,3.1,3.2,3.3,2.1,2.2,2.3

⍙(ALLQTLS='N')/'UNIQCODES←1.1,1.2,1.3,1.4,1.5,1.6,1.7,2.1,2.2,
2.3'     ⍝ USE THIS FOR JUST FLOWER AND NECTAR TRAIT ANALYSIS
⍙(ALLQTLS='N')/'∆TOTRHEA←∆TOTRHEA[1 2 3 4 5 6 7 11 12 13;]'
⍝  USE THIS FOR JUST FLOWER AND NECTAR TRAIT ANALYSIS

SCORRS1←SCORRS2←SCORRS3←SCORRS4←0⍴0
PREDICTCORR

⎕TCLF
'OBSERVED CORRELATIONS CALCULATED'
⎕TCLF
'BEGINNING SHUFFLE ANALYSIS'

IJ←0
RETIJ:IJ←IJ+1
```

```
TEST←(IJ÷100)=⌊(IJ÷100)
⍙(TEST=1)/'''REPS COMPLETED '',IJ'
PREDICTCORR2

SCORRS1←SCORRS1,SHUFCORR1
SCORRS2←SCORRS2,SHUFCORR2
SCORRS3←SCORRS3,SHUFCORR3
SCORRS4←SCORRS4,SHUFCORR4

→(IJ<1000)/RETIJ

SUM1←+/SCORRS1≥OBSCORR1
SUM2←+/SCORRS2≥OBSCORR2
SUM3←+/SCORRS3≥OBSCORR3
SUM4←+/SCORRS4≥OBSCORR4
□TCLF
TITLE
□TCLF
'CORRELATION BETWEEN GENETIC CORRELATIONS AND PREDICTED CORREL
ATION'
'    OBSERVED CORRELATION: ',OBSCORR1
'    PROPORTION OF SHUFFLED CORRELATIONS ≥ OBSERVED: ',SUM1÷10
00
'CORRELATION BETWEEN GENETIC CORRELATIONS AND BIAS'
'    OBSERVED CORRELATION: ',OBSCORR2
'    PROPORTION OF SHUFFLED CORRELATIONS ≥ OBSERVED: ',SUM2÷10
00
'CORRELATION BETWEEN BIAS AND AVE TOT RHE: '
'    OBSERVED CORRELATION: ',OBSCORR3
'    PROPORTION OF SHUFFLED CORRELATIONS ≥ OBSERVED: ',SUM3÷10
00
'CORRELATION BETWEEN QTL OVERLAP AND PREDICTED CORRELATION'
'    OBSERVED CORRELATION: ',OBSCORR4
'    PROPORTION OF SHUFFLED CORRELATIONS ≥ OBSERVED: ',SUM4÷10
00
```

F. Program PREDICTCORR2

Called from SHUFFLEPRED

NOTE: This program is essentially the same as PREDICTCORR, slightly modified to be called multiple times by SHUFFLEPRED

Code:

```
PREDICTCORR2
⍝ THIS PROGRAM PREDICTS PAIRWISE TRAIT GENETIC CORRELATIONS FR
   OM QTL OVERLAP AND EFFECT SIZES FOR RANDOM ASSIGMNENT OF TR
   AITS TO QTLS
```

```
QTLDATA←QTLDATAA         ⍝ USE ∆MASTERDATA FOR LESS-CONSERVATIVE
   QTL SET, ∆MASTERDATA1 FOR CONSERVATIVE DATASET
TRAITCODES←TRAITCODESA           ⍝  AND USE ∆MASTERDATA FOR LESS
  -CONSERVATIVE QTL SET, ∆MASTERDATA1 FOR CONSERVATIVE DATASET
MIDNUMS←∆MIDPOINTSA[;1 2]
IND←(~TRAITCODES∊(2.4,2.5))/⍳⍴TRAITCODES   ⍝ LIST TRAITS TO EXC
 LUDE   NOTE: 2.4 AND 2.5 ARE ALTERNATES FOR 2.3 (NECTARY SIZE)

⍕(ALLQTLS='N')/'IND←(~TRAITCODES∊(3.1,3.2,3.3,2.4,2.5))/⍳⍴TRAI
TCODES'  ⍝ USE THIS FOR JUST FLOWER AND NECTAR TRAIT ANALYSIS

TRAITCODES←TRAITCODES[IND]
QTLDATA←QTLDATA[IND;]
CODES←,QTLDATA[;1]
NUM←⍴CODES
ORDER←NUM?NUM

QTLDATA[;1]←CODES[ORDER]
TRAITCODES← CODES[ORDER]

UNIQCODES←1.1,1.2,1.3,1.4,1.5,1.6,1.7,3.1,3.2,3.3,2.1,2.2,2.3

⍕(ALLQTLS='N')/'UNIQCODES←1.1,1.2,1.3,1.4,1.5,1.6,1.7,2.1,2.2,
2.3'     ⍝ USE THIS FOR JUST FLOWER AND NECTAR TRAIT ANALYSIS
⍝ ⍕(ALLQTLS='N')/'∆TOTRHEA←∆TOTRHEA[1 2 3 4 5 6 7 11 12 13;]'
   ⍝  USE THIS FOR JUST FLOWER AND NECTAR TRAIT ANALYSIS

MAXJ←⍴UNIQCODES
MAXI←MAXJ-1

MIDPOINTS←0 4⍴0    ⍝ THIS VARIABLE HOLDS MIDPOINTS OF OVERLAP I
   NTERVALS FOR QTLS THAT OVERLAP
                  ⍝ FIRST TWO COLUMNS ARE QTL UNIQUE ID, THIRD
   IS LG AND FOURTH IS MIDPOINT

GENCORRS←((⍴UNIQCODES),(⍴UNIQCODES))⍴0
AVETRHES←((⍴UNIQCODES),(⍴UNIQCODES))⍴0
OVERLAPMAT←((⍴UNIQCODES),(⍴UNIQCODES))⍴0

I←0
RETI:I←I+1  ⍝  FIRST TRAIT LOOP

J←I
RETJ:J←J+1  ⍝  SECOND TRAIT LOOP

TRAITCODEI←UNIQCODES[I]
TRAITCODEJ←UNIQCODES[J]

IND1←(TRAITCODES=UNIQCODES[I])/⍳⍴TRAITCODES
```

```
PART1←QTLDATA[IND1;]
IND2←(TRAITCODES=UNIQCODES[J])/ιρTRAITCODES
PART2←QTLDATA[IND2;]


OVERLAPS←0 4ρ0    ⍝ NUMBERS ARE QTLNUMBERS FROM PARTS 1 AND 2 F
OR OVERLAPPING QTL

MAXII←1↑ρPART1
MAXJJ←1↑ρPART2

II←0
RETII:II←II+1    ⍝ PART1 LOOP

JJ←0
RETJJ:JJ←JJ+1    ⍝ PART2 LOOP

LINEII←PART1[II;] ⍝ GET QTL DATA FOR QTL1  (CHROMOSOME, CI LOW
  , CI HIGH)
LINEJJ←PART2[JJ;]  ⍝ GET QTL DATA FOR QTL2

TEST1←LINEII[2]=LINEJJ[2]   ⍝ QTL1 AND QTL2 ON SAME CHROMOSOME?

→(TEST1=0)/DOWN        ⍝ IF NOT, SKIP AND LEAVE ZERO IN  MATRIX

TEST←((LINEII[4]=LINEJJ[3])∨(LINEJJ[4]=LINEII[3]))
→(TEST=1)/DOWN            ⍝ DO NOT COUNT AS OVERLAP IF CI'S ABUT
CRIT1←(LINEII[3]≥LINEJJ[3]) ∧ (LINEII[3]≤LINEJJ[4])     ⍝ THES
  E TWO LINES TEST WHETHER THERE IS OVERLAP
CRIT2←(LINEII[4]≥LINEJJ[3]) ∧ (LINEII[4]≤LINEJJ[4])
CRIT3←(LINEJJ[3]≥LINEII[3]) ∧ (LINEJJ[3]≤LINEII[4])
CRIT4←(LINEJJ[4]≥LINEII[3]) ∧ (LINEJJ[4]≤LINEII[4])

TEST2←CRIT1∨CRIT2∨CRIT3∨CRIT4
→(TEST2=0)/DOWN

OPAIR←2 1ρ(LINEII[12],LINEJJ[12])
TESTPAIR←+/,(MIDNUMS∧.=OPAIR)
⍘(TESTPAIR=0)/'TEMP←OPAIR[2; 1],OPAIR[1;1] ◊ OPAIR←2 1ρTEMP'
TESTPAIR2←+/,(MIDNUMS∧.=OPAIR)

 ⍘(TESTPAIR2=0)/'''TESTPAIR2 NOT IN DATASET'' ◊ →0'

OVERLAPS←OVERLAPS,[1](II,JJ,LINEII[12],LINEJJ[12])

DOWN:→(JJ<MAXJJ)/RETJJ

→(II<MAXII)/RETII

MAXK←1↑ρOVERLAPS
```

```apl
⍝ FOLLOWING VARIABLE HAS COLUMNS: COL1 RHE ON TRAITI  RHE ON T
    RAITJ
RHE←0 4⍴0

SIZE←1↑⍴OVERLAPS
→(SIZE=0)/DOWN4

K←0
RETK:K←K+1 ⍝ THIS LOOP ADDS RHE FOR OVERLAPPING QTLS

RHEVALS←(PART1[OVERLAPS[K;1];9],PART2[OVERLAPS[K;2];9],OVERLAP
S[K;3 4])

RHE←RHE,[1]RHEVALS

→(K<MAXK)/RETK

DOWN4:MAXL←1↑⍴PART1
L←0
RETL:L←L+1

TEST←L∊OVERLAPS[;1]
→(TEST=1)/DOWN2

RHE←RHE,[1](PART1[L;9],0,PART1[L;12],0)

DOWN2:→(L<MAXL)/RETL

MAXLL←1↑⍴PART2
LL←0
RETLL:LL←LL+1

TEST←LL∊OVERLAPS[;2]
→(TEST=1)/DOWN3

RHE←RHE,[1](0,PART2[LL;9],0,PART2[LL;12])

DOWN3:→(LL<MAXLL)/RETLL

CALCCOV2

GENCORRS[I;J]←R

IND1←(RHE[;1]≠0)/⍳1↑⍴RHE
TRHE1←+/RHE[IND1;1]
IND2←(RHE[;2]≠0)/⍳1↑⍴RHE
TRHE2←+/RHE[IND2;2]
AVETOTRHE←(TRHE1+TRHE2)÷2
AVETRHES[I;J]←AVETOTRHE
```

```
IND1←((RHE[;1]≠0)∧(RHE[;2]≠0))/ι1↑ρRHE
A←ρIND1
B←+/RHE[;1]≠0
C←+/RHE[;2]≠0
OV←A÷(B+C-1)
OVERLAPMAT[I;J]←OV

→(J<MAXJ)/RETJ
→(I<MAXI)/RETI

TRIANGLE GENCORRS
♣(ALLQTLS≠'N')/'DATA←ΔGENCORRS2,78 1ρVEC'
♣(ALLQTLS='N')/'DATA←ΔGENCORRS2[NONSEEDIND;],45 1ρVEC'
BIAS←DATA[;2]-DATA[;1]
♣(ALLQTLS≠'N')/'DATA←DATA,78 1ρBIAS'
♣(ALLQTLS='N')/'DATA←DATA,45 1ρBIAS'
TRIANGLE AVETRHES
♣(ALLQTLS≠'N')/'DATA←DATA,78 1ρVEC'
♣(ALLQTLS='N')/'DATA←DATA,45 1ρVEC'

TRIANGLE OVERLAPMAT
♣(ALLQTLS≠'N')/'DATA←DATA,78 1ρVEC'
♣(ALLQTLS='N')/'DATA←DATA,45 1ρVEC'

CALCCORR DATA[;1 2]
SHUFCORR1←CORR
CALCCORR DATA[;1 3]
SHUFCORR2←CORR
CALCCORR DATA[;3 4]
SHUFCORR3←CORR

CALCCORR DATA[;2 5]
SHUFCORR4←CORR

⍝ 'PROGRAM PREDICTCORR FINISHED.  OUTPUT IN VARIABLES ''GENCOR
RS'' AND ''AVETRHES''.'
```

IV. Ascertainment Bias Programs

PROGRAMS ASCERTAIN11, ASCERTAIN10, AND ASCERTAIN10A
MODEL EFFECTS OF ASCERTAINMENT BIAS.

A. Program ASCERTAIN11

Code:

```
ASCERTAIN11
⍝THIS PROGRAM CALCULATES ASCERTAINMENT BIAS FOR 11-LOCUS TRAIT
⍝ CALCULATES FALSE POSITIVE RATE FOR N CANDIDATE TRAITS, WHERE
   N × 1, 10, 25 50 AND 100.
⍝ CALCULATES FPR FOR CHOOSING MOST DIVERGE TRAIT FROM CANDIDA
   TE LIST, AND FOR CHOOSING RANDOMLY FROM TOP HALF OF DIVERGE
   D TRAITS.

NTRAITS←50  ⍝  TOTAL NUMBER OF TRAITS TO CHOOSE FROM

SCORE1←SCORE10TOP←SCORE10MID←SCORE25TOP←SCORE25MID←SCORE50TOP←
SCORE50MID←RAND1←RAND10←RAND25←RAND50←RAND10TOP←RAND25TOP←RAND
50TOP←0⍴0
SCORE100TOP←SCORE100MID←RAND100←RAND100TOP←0⍴0

NLOCI←11

I←0
RETI:I←I+1  ⍝ REPLICATE LOOP

SCORES←0⍴0   ⍝  NUMBER OF 2 ALLELES
J←0
RETJ:J←J+1  ⍝ TRAIT LOOP

ALLELES←?(NLOCI⍴2)

⍎(16≥(+/ALLELES))/'ALLELES←|(ALLELES-3)'      ⍝ CHANGE IF CHANG
   E NUMBER OF TRIALS

SCORE←+/(ALLELES=2)
SCORES←SCORES,SCORE

→(J<100)/RETJ

SCORE1←SCORE1,SCORES[1]
RAND1←RAND1,SCORES[?50]

PART10←SCORES[⍳10]
PART10←PART10[⍒PART10]
SCORE10TOP←SCORE10TOP,PART10[1]
```

```
SCORE10MID←SCORE10MID,PART10[5]
RAND10←RAND10,PART10[?10]
RAND10TOP←RAND10TOP,PART10[?5]

PART25←SCORES[ι25]
PART25←PART25[▼PART25]
SCORE25TOP←SCORE25TOP,PART25[1]
SCORE25MID←SCORE25MID,PART25[13]
RAND25←RAND25,PART25[?25]
RAND25TOP←RAND25TOP,PART25[?13]

PART50←SCORES[ι50]
PART50←PART50[▼PART50]

SCORE50TOP←SCORE50TOP,PART50[1]
SCORE50MID←SCORE50MID,PART50[25]
RAND50←RAND50,PART50[?50]
RAND50TOP←RAND50TOP,PART50[?25]

SCORES←SCORES[▼SCORES]
SCORE100TOP←SCORE100TOP,SCORES[1]
SCORE100MID←SCORE100MID,SCORES[50]
RAND100←RAND100,SCORES[?100]
RAND100TOP←RAND100TOP,SCORES[?50]

→(I<10000)/RETI

□TCLF
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 1        ',(+/SCORE1
≥10)÷ρSCORE1
'FALSE POSITIVE PROPORTION FOR RANDOM SAMPLE OF 1  ',(+/SCORE1
≥10)÷ρSCORE1

□TCLF
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 10 TOP     ',(+/SCORE1
0TOP≥10)÷ρSCORE10TOP
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 10 MID     ',(+/SCORE1
0MID≥10)÷ρSCORE10MID
'FALSE POSITIVE PROPORTION FOR RANDOM SAMPLE OF 10 ',(+/RAND10
≥10)÷ρRAND10
'FALSE POSIT PROPORTION OFR RANDOM SAMPLE OF TOP 5 ',(+/RAND10
TOP≥10)÷ρRAND10TOP

□TCLF
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 25 TOP     ',(+/SCORE2
5TOP≥10)÷ρSCORE25TOP
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 25 MID     ',(+/SCORE2
5MID≥10)÷ρSCORE25MID
'FALSE POSITIVE PROPORTION FOR RANDOM SAMPLE OF 25 ',(+/RAND25
≥10)÷ρRAND25
```

```
'FALSE POSI PROPORTION OFR RANDOM SAMPLE OF TOP 13 ',(+/RAND25
TOP≥10)÷ρRAND25TOP

⎕TCLF
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 50 TOP    ',(+/SCORE5
0TOP≥10)÷ρSCORE50TOP
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 50 MID    ',(+/SCORE5
0MID≥10)÷ρSCORE50MID
'FALSE POSITIVE PROPORTION FOR RANDOM SAMPLE OF 50 ',(+/RAND50
≥10)÷ρRAND50
'FALSE POSI PROPORTION OFR RANDOM SAMPLE OF TOP 25 ',(+/RAND50
TOP≥10)÷ρRAND50TOP

⎕TCLF
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 100 TOP   ',(+/SCORE1
00TOP≥10)÷ρSCORE100TOP
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 100 MID   ',(+/SCORE1
00MID≥10)÷ρSCORE100MID
'FALSE POSITIVE PROPORTION FOR RANDOM SAMPLE OF 100 ',(+/RAND1
00≥10)÷ρRAND100
'FALSE POSI PROPORTION OFR RANDOM SAMPLE OF TOP 50 ',(+/RAND10
0TOP≥10)÷ρRAND100TOP

⎕TCLF
⎕TCLF
'PROGRAM ASCERTAIN11 FINISHED.'
```

B. Program ASCERTAIN10

Code:

```
ASCERTAIN10
⍝ THIS PROGRAM CALCULATES ASCERTAINMENT BIAS

NTRAITS←50  ⍝  TOTAL NUMBER OF TRAITS TO CHOOSE FROM

SCORE1←SCORE10TOP←SCORE10MID←SCORE25TOP←SCORE25MID←SCORE50TOP←
SCORE50MID←RAND1←RAND10←RAND25←RAND50←RAND10TOP←RAND25TOP←RAND
50TOP←0ρ0

NLOCI←10

I←0
RETI:I←I+1  ⍝ REPLICATE LOOP

SCORES←0ρ0   ⍝   NUMBER OF 2 ALLELES
J←0
RETJ:J←J+1  ⍝ TRAIT LOOP
```

28

```
ALLELES←?(NLOCIρ2)

♣(15≥(+/ALLELES))/'ALLELES←|(ALLELES-3)'     ⍝ CHANGE IF CHANGE
   NUMBER OF TRIALS

SCORE←+/(ALLELES=2)
SCORES←SCORES,SCORE

→(J<50)/RETJ

SCORE1←SCORE1,SCORES[1]
RAND1←RAND1,SCORES[?50]

PART10←SCORES[ι10]
PART10←PART10[▼PART10]
SCORE10TOP←SCORE10TOP,PART10[1]
SCORE10MID←SCORE10MID,PART10[5]
RAND10←RAND10,PART10[?10]
RAND10TOP←RAND10TOP,PART10[?5]

PART25←SCORES[ι25]
PART25←PART25[▼PART25]
SCORE25TOP←SCORE25TOP,PART25[1]
SCORE25MID←SCORE25MID,PART25[13]
RAND25←RAND25,PART25[?25]
RAND25TOP←RAND25TOP,PART25[?13]

SCORES←SCORES[▼SCORES]
SCORE50TOP←SCORE50TOP,SCORES[1]
SCORE50MID←SCORE50MID,SCORES[25]
RAND50←RAND50,SCORES[?50]
RAND50TOP←RAND50TOP,SCORES[?25]

→(I<10000)/RETI

□TCLF
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 1          ',(+/SCORE1
≥9)÷ρSCORE1
'FALSE POSITIVE PROPORTION FOR RANDOM SAMPLE OF 1  ',(+/SCORE1
≥9)÷ρSCORE1

□TCLF
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 10 TOP    ',(+/SCORE1
0TOP≥9)÷ρSCORE10TOP
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 10 MID    ',(+/SCORE1
0MID≥9)÷ρSCORE10MID
'FALSE POSITIVE PROPORTION FOR RANDOM SAMPLE OF 10 ',(+/SCORE1
≥9)÷ρSCORE1
'FALSE POSIT PROPORTION OFR RANDOM SAMPLE OF TOP 5 ',(+/RAND10
TOP≥9)÷ρRAND10TOP
```

```
□TCLF
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 25 TOP    ',(+/SCORE2
5TOP≥9)÷ρSCORE25TOP
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 25 MID    ',(+/SCORE2
5MID≥9)÷ρSCORE25MID
'FALSE POSITIVE PROPORTION FOR RANDOM SAMPLE OF 25 ',(+/SCORE1
≥9)÷ρSCORE1
'FALSE POSI PROPORTION OFR RANDOM SAMPLE OF TOP 13 ',(+/RAND25
TOP≥9)÷ρRAND25TOP


□TCLF
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 50 TOP    ',(+/SCORE5
0TOP≥9)÷ρSCORE50TOP
'FALSE POSITIVE PROPORTION FOR SAMPLE OF 50 MID    ',(+/SCORE5
0MID≥9)÷ρSCORE50MID
'FALSE POSITIVE PROPORTION FOR RANDOM SAMPLE OF 50 ',(+/SCORE1
≥9)÷ρSCORE1
'FALSE POSI PROPORTION OFR RANDOM SAMPLE OF TOP 25 ',(+/RAND50
TOP≥9)÷ρRAND50TOP

□TCLF
□TCLF
'PROGRAM ASCERTAIN10 FINISHED.'
```

C. Program ASCERTAIN11A

Code:

```
ASCERTAIN11A
A THIS PROGRAM CALCULATES ASCERTAINMENT BIAS FOR CHOOSING RAND
  OMLY FROM X MOST DIVERGED TRAITS OUT OF 100 TRAITS.

NTRAITS←50  A  TOTAL NUMBER OF TRAITS TO CHOOSE FROM

SCORE1←SCORE10TOP←SCORE10MID←SCORE25TOP←SCORE25MID←SCORE50TOP←
SCORE50MID←RAND1←RAND10←RAND25←RAND50←RAND10TOP←RAND25TOP←RAND
50TOP←0ρ0
SCORE100TOP←SCORE100MID←RAND100←RAND100TOP←0ρ0

TSCORE100←TSCORE90←TSCORE80←TSCORE70←TSCORE60←TSCORE50←TSCORE4
0←TSCORE30←TSCORE20←TSCORE10←0ρ0

NLOCI←11

I←0
RETI:I←I+1  A REPLICATE LOOP

SCORES←0ρ0   A  NUMBER OF 2 ALLELES
```

```
J←0
RETJ:J←J+1  ⍝ TRAIT LOOP

ALLELES←?(NLOCIρ2)    ⍝ GENERATE ALLELES FOR TRAIT (1 OR 2 AT E
   ACH LOCUS)

⍎(16≥(+/ALLELES))/'ALLELES←|(ALLELES-3)'     ⍝ SWITCH 1 AND 2
   ALLELES IF MORE 1 THAN 2 ALLELES

SCORE←+/(ALLELES=2)       ⍝ NUMBER OF 2 ALLELES IN GENOTYPE
SCORES←SCORES,SCORE

→(J<100)/RETJ

SCORES←SCORES[⍒SCORES]              ⍝ ORDER TRAITS BY NUMBER OF
   2 ALLELES, HIGHEST TO LOWEST

TSCORE100←TSCORE100,SCORES[?100]   ⍝ PICK STUDY TRAIT RANDOMLY
   FROM AMONG ALL CANDIDATE TRAITS
TSCORE90←TSCORE90,SCORES[?90]      ⍝ PICK STUDY TRAIT RANDOMLY
   FROM AMONG MOST-DIVERGED 90 PERCENT OF CANDIDATE TRAITS
TSCORE80←TSCORE80,SCORES[?80]      ⍝  ETC.
TSCORE70←TSCORE70,SCORES[?70]
TSCORE60←TSCORE60,SCORES[?60]
TSCORE50←TSCORE50,SCORES[?50]
TSCORE40←TSCORE40,SCORES[?40]
TSCORE30←TSCORE30,SCORES[?30]
TSCORE20←TSCORE20,SCORES[?20]
TSCORE10←TSCORE10,SCORES[?10]

→(I<10000)/RETI

MAT←0 10000ρ0       ⍝ MAKE A MATRIX OUT OF THE TSCORE VARIABLES
MAT←MAT,[1]TSCORE100
MAT←MAT,[1]TSCORE90
MAT←MAT,[1]TSCORE80
MAT←MAT,[1]TSCORE70
MAT←MAT,[1]TSCORE60
MAT←MAT,[1]TSCORE50
MAT←MAT,[1]TSCORE40
MAT←MAT,[1]TSCORE30
MAT←MAT,[1]TSCORE20
MAT←MAT,[1]TSCORE10


VEC←(+/MAT≥10)÷10000       ⍝  COUNT PROPORTION OF STUDY TRAITS
   IN MATRIX SIGNIFICANT AT NOMINAL LEVEL (E.G. HAS EITHER 10
   OR 11 2 ALLELES)--THIS IS FALSE DISCOVERY RATE
OBSRATIO←VEC÷VEC[1]      ⍝  CALCULATE RATIO OF FDR FOR A GIVEN
   BETA TO FDR FOR BETA=1      ('FDR RATIO')
Q←⍳9
```

```
QQ←1÷(1−.1×Q)                    ⍝ EXPECTED FDR RATIO (EQUAL TO 1÷BETA)
QQ←1,QQ

OUT←(10 1ρVEC),10 1ρQQ
OUT←OUT,10 1ρOBSRATIO
OUT←OUT,10 1ρ(100,90,80,70,60,50,40,30,20,10)

⎕TCLF
'FALSE POSITIVE RATOS '
⎕TCLF
' OBS   OBSRATIO EXPRATIO    PROPORTION'
OUT

⎕TCLF
⎕TCLF
'PROGRAM ASCERTAIN11A FINISHED.'
```

IV. Counting antagonistic QTLs for assessing selection

PROGRAM CONTRA-DIRECTIONAL QTL COUNTS, FOR EACH FLORAL AND NECTAR TRAIT (NOT TRANSFORMED):

 (1) NUMBER OF CONTRA QTLS (QTLS WITH EFFECTS IN OPPOSITE DIRECTION
    AS SPECIES DIFFERENCE)
(2) NUMBER OF THESE OVERLAPPING CONSISTENT FLORAL QTLS (CONSISTENT
    MEANS EFFECT IN SAME DIRECTION AS SPECIES DIFFERENCE)
(3) NUMBER OF THESE OVERLAPPING CONSISTENT NECTAR QTLS
(4) NUMBER OF THESE OVERLAPPING EITHER CONSISTENT FLORAL QTLS,
    CONSISTENT NECTAR QTLS, OR BOTH

A. Program ANTAGQTL

Code:

```
ANTAGQTL
⍝ THIS PROGRAM IDENTIFIES QTL'S WITH ANTAGONISTIC EFFECTS ON
    DIFFERENT TRAITS
THREES←3.1,3.2,3.3
MATCHES←0 10ρ0
MAXJ←1↑ρ⍋MASTERDATA
MAXI←MAXJ-1
I←0
RETI:I←I+1    ⍝ LOOP FOR FIRST QTL

J←I
RETJ:J←J+1

LINEI←⍋MASTERDATA[I;]
LINEJ←⍋MASTERDATA[J;]

TEST←LINEI[1]=LINEJ[1]
→(TEST=1)/DOWN            ⍝ SKIP IF QTLS AFFECT SAME TRAIT

TEST←LINEI[2]≠LINEJ[2]
→(TEST=1)/DOWN            ⍝ SKIP ON DIFFERENT CHROMOSOME


TEST←((LINEI[4]=LINEJ[3])∨(LINEJ[4]=LINEI[3]))
→(TEST=1)/DOWN
CRIT1←(LINEI[3]≥LINEJ[3]) ∧ (LINEI[3]≤LINEJ[4])     ⍝ THESE
    TWO LINES TEST WHETHER THERE IS OVERLAP
CRIT2←(LINEI[4]≥LINEJ[3]) ∧ (LINEI[4]≤LINEJ[4])
CRIT3←(LINEJ[3]≥LINEI[3]) ∧ (LINEJ[3]≤LINEI[4])
CRIT4←(LINEJ[4]≥LINEI[3]) ∧ (LINEJ[4]≤LINEI[4])
```

33

```
TEST2←CRIT1∨CRIT2∨CRIT3∨CRIT4
→(TEST2=0)/DOWN

TEST1←((LINEI[9]>0)∧(LINEJ[9]<0)) ∨ ((LINEI[9]<0)∧(LINEJ[9]>0)
)

→(TEST1=0)/DOWN

TEST←(LINEI[1]∈THREES)∨(LINEJ[1]∈THREES)
→(TEST=1)/DOWN

NEWLINE←LINEI[ι4],LINEJ[ι4],LINEI[9],LINEJ[9]
MATCHES←MATCHES,[1]NEWLINE

DOWN:→(J<MAXJ)/RETJ

→(I<MAXI)/RETI
```