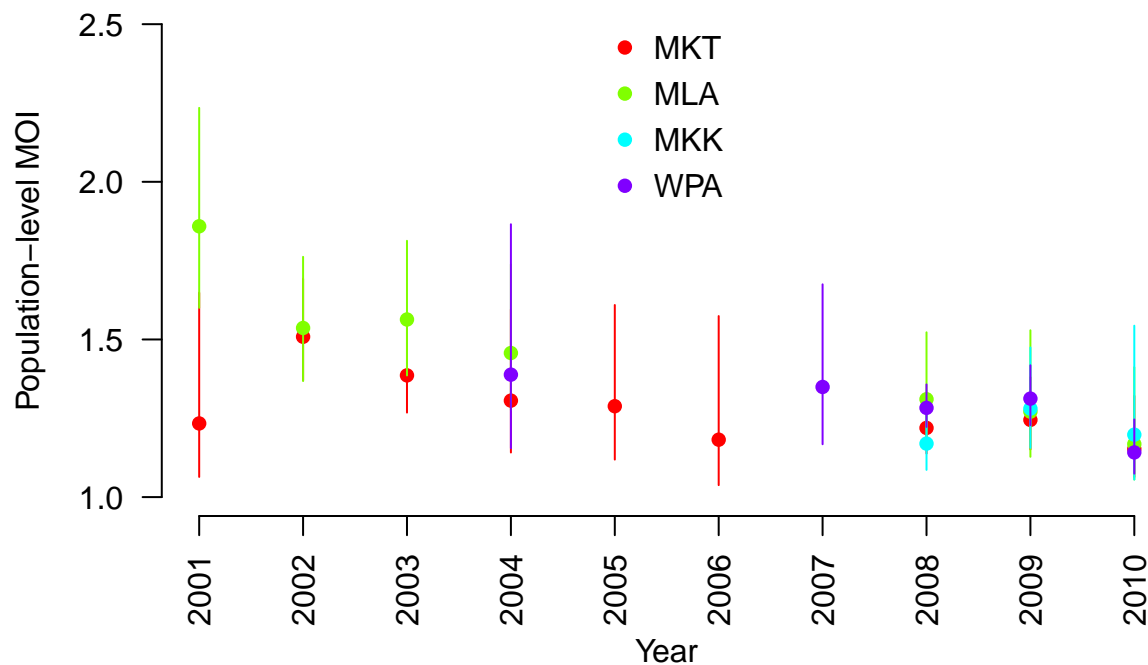


# Visualise results from frequency estimation and compare with proportions barcode data

This script plots results from running COIL extension on TM border data partitioned by site and site and year, compared to proportion frequencies, which were based on data excluding all multiclonal. MCMC fit to each partition using, 10000 iterations,  $\psi = 1$ ,  $\phi = 1$ ,  $\alpha = 1$ ,  $\beta = 1$ ,  $\sigma^2 = 1$ . In the final draft on the manuscript the frequencies inferred under the COIL extension model were deemed superfluous.

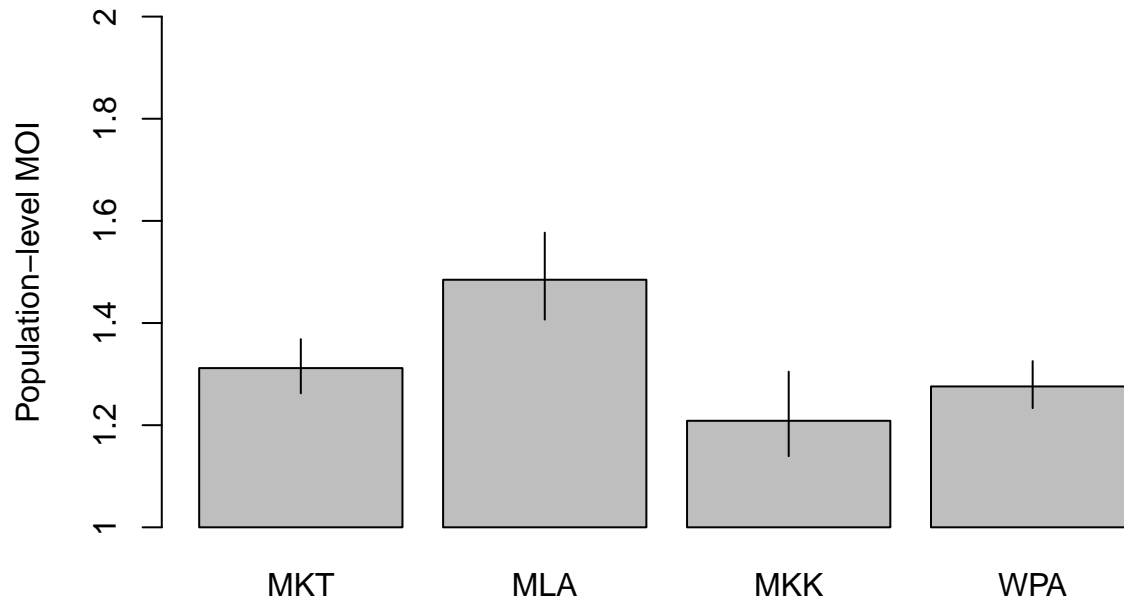
---

## MOI with site and year



```
##
## Call:
## lm(formula = MOI_quants_site_year["50%", ] ~ X_years)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.29284 -0.04597  0.01826  0.05110  0.33251
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.564767   0.053311  29.352  < 2e-16 ***
## X_years      -0.038298   0.007648  -5.007  5.18e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1156 on 22 degrees of freedom
## Multiple R-squared:  0.5326, Adjusted R-squared:  0.5114
```

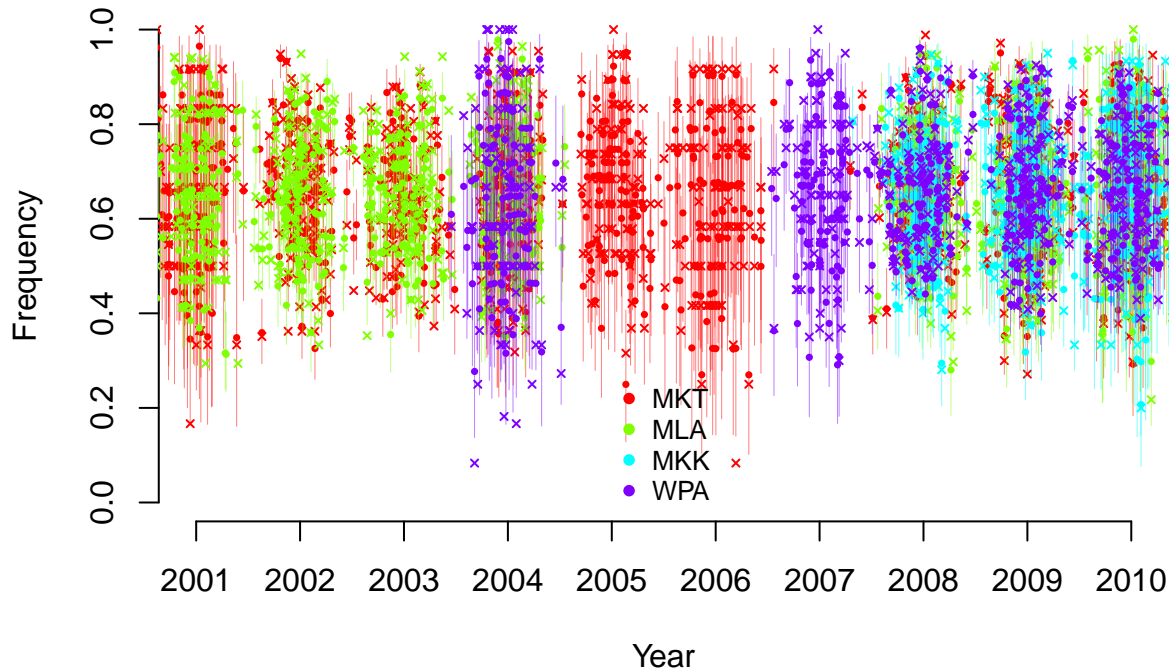
```
## F-statistic: 25.07 on 1 and 22 DF,  p-value: 5.176e-05
##
## Call:
## lm(formula = MOI_quants_site_year["50%", ] ~ X_years + X_sites)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.216462 -0.046135 -0.000627  0.046743  0.238760
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.569013   0.085078  18.442 1.38e-13 ***
## X_years      -0.039213   0.007143  -5.490 2.69e-05 ***
## X_sitesMKT   -0.079711   0.069477  -1.147   0.265
## X_sitesMLA    0.090417   0.071699   1.261   0.223
## X_sitesWPA    0.023960   0.071199   0.337   0.740
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09653 on 19 degrees of freedom
## Multiple R-squared:  0.7186, Adjusted R-squared:  0.6594
## F-statistic: 12.13 on 4 and 19 DF,  p-value: 4.586e-05
```



## MOI over all

Average MOI: 1.3287485, which is pretty similar to averaging over the site-wise mediums: 1.3202423

## Inferred frequencies by site and year



```
##
## Call:
## lm(formula = freq ~ site + year, data = Data_long_format)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.44896 -0.09301  0.00154  0.09724  0.32743
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.466e-01  1.232e-02  52.475  <2e-16 ***
## siteMKT      2.700e-03  1.006e-02   0.268   0.788
## siteMLA      2.224e-04  1.038e-02   0.021   0.983
## siteWPA     -8.707e-05  1.031e-02  -0.008   0.993
## year         9.458e-04  1.035e-03   0.914   0.361
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1348 on 2227 degrees of freedom
## Multiple R-squared:  0.0004532, Adjusted R-squared:  -0.001342
## F-statistic: 0.2524 on 4 and 2227 DF,  p-value: 0.9083
##
## Analysis of Variance Table
##
## Response: freq
##      Df Sum Sq  Mean Sq F value Pr(>F)
## site    3  0.003  0.0010546   0.0580 0.9817
## year    1  0.015  0.0151932   0.8358 0.3607
## Residuals 2227 40.484  0.0181788
##
##      Estimate      Std. Error      t value      Pr(>|t|)
```

```
## -2.685351e-02  4.688386e-03 -5.727666e+00  1.608018e-05
## [1] "11"
```

## Do proportion frequencies vary between sites and years?

This section features in the supplementary of the final draft of the manuscript

\*\*\*\*\*Show in New WindowClear OutputExpand/Collapse Output

```
Call: lm(formula = MOI_quants_site_year["50%", ] ~ X_years)
```

```
Residuals: Min 1Q Median 3Q Max -0.29284 -0.04597 0.01826 0.05110 0.33251
```

```
Coefficients: Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) 1.564767 0.053311 29.352 < 2e-16 X_years -0.038298 0.007648 -5.007 5.18e-05 — Signif.
codes: 0 ‘’ 0.001 ‘’ 0.01 ‘’ 0.05 ‘’ 0.1 ‘’ 1
```

```
Residual standard error: 0.1156 on 22 degrees of freedom Multiple R-squared: 0.5326, Adjusted R-squared:
0.5114 F-statistic: 25.07 on 1 and 22 DF, p-value: 5.176e-05
```

```
Call: lm(formula = MOI_quants_site_year["50%", ] ~ X_years + X_sites)
```

```
Residuals: Min 1Q Median 3Q Max -0.216462 -0.046135 -0.000627 0.046743 0.238760
```

```
Coefficients: Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) 1.569013 0.085078 18.442 1.38e-13 X_years -0.039213 0.007143 -5.490 2.69e-05
```

```
X_sitesMKT -0.079711 0.069477 -1.147 0.265
```

```
X_sitesMLA 0.090417 0.071699 1.261 0.223
```

```
X_sitesWPA 0.023960 0.071199 0.337 0.740
```

```
— Signif. codes: 0 ‘’ 0.001 ‘’ 0.01 ‘’ 0.05 ‘’ 0.1 ‘’ 1
```

```
Residual standard error: 0.09653 on 19 degrees of freedom Multiple R-squared: 0.7186, Adjusted R-squared:
0.6594 F-statistic: 12.13 on 4 and 19 DF, p-value: 4.586e-05
```

R Console

```
Call: lm(formula = MOI_quants_site_year["50%", ] ~ X_years)
```

```
Residuals: Min 1Q Median 3Q Max -0.29284 -0.04597 0.01826 0.05110 0.33251
```

```
Coefficients: Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) 1.564767 0.053311 29.352 < 2e-16 X_years -0.038298 0.007648 -5.007 5.18e-05 — Signif.
codes: 0 ‘’ 0.001 ‘’ 0.01 ‘’ 0.05 ‘’ 0.1 ‘’ 1
```

```
Residual standard error: 0.1156 on 22 degrees of freedom Multiple R-squared: 0.5326, Adjusted R-squared:
0.5114 F-statistic: 25.07 on 1 and 22 DF, p-value: 5.176e-05
```

```
Call: lm(formula = MOI_quants_site_year["50%", ] ~ X_years + X_sites)
```

```
Residuals: Min 1Q Median 3Q Max -0.216462 -0.046135 -0.000627 0.046743 0.238760
```

```
Coefficients: Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) 1.569013 0.085078 18.442 1.38e-13 X_years -0.039213 0.007143 -5.490 2.69e-05
```

```
X_sitesMKT -0.079711 0.069477 -1.147 0.265
```

```
X_sitesMLA 0.090417 0.071699 1.261 0.223
```

```
X_sitesWPA 0.023960 0.071199 0.337 0.740
```

```
— Signif. codes: 0 ‘’ 0.001 ‘’ 0.01 ‘’ 0.05 ‘’ 0.1 ‘’ 1
```

```
Residual standard error: 0.09653 on 19 degrees of freedom Multiple R-squared: 0.7186, Adjusted R-squared:
0.6594 F-statistic: 12.13 on 4 and 19 DF, p-value: 4.586e-05
```

Show in New WindowClear OutputExpand/Collapse Output

Show in New Window Clear Output Expand/Collapse Output

Call: `lm(formula = freq ~ site + year, data = Data_long_format)`

Residuals: Min 1Q Median 3Q Max -0.44896 -0.09301 0.00154 0.09724 0.32743

Coefficients: Estimate Std. Error t value Pr(>|t|)

(Intercept) 6.466e-01 1.232e-02 52.475 <2e-16 \*\*\* siteMKT 2.700e-03 1.006e-02 0.268 0.788

siteMLA 2.224e-04 1.038e-02 0.021 0.983

siteWPA -8.707e-05 1.031e-02 -0.008 0.993

year 9.458e-04 1.035e-03 0.914 0.361

— Signif. codes: 0 ‘’ **0.001** ‘’ 0.01 ‘’ 0.05 ‘’ 0.1 ‘’ 1

Residual standard error: 0.1348 on 2227 degrees of freedom Multiple R-squared: 0.0004532, Adjusted R-squared: -0.001342 F-statistic: 0.2524 on 4 and 2227 DF, p-value: 0.9083

Analysis of Variance Table

Response: freq Df Sum Sq Mean Sq F value Pr(>F) site 3 0.003 0.0010546 0.0580 0.9817 year 1 0.015 0.0151932 0.8358 0.3607 Residuals 2227 40.484 0.0181788

Estimate Std. Error t value Pr(>|t|) -2.685351e-02 4.688386e-03 -5.727666e+00 1.608018e-05 [1] “11” R Console

Call: `lm(formula = freq ~ site + year, data = Data_long_format)`

Residuals: Min 1Q Median 3Q Max -0.44896 -0.09301 0.00154 0.09724 0.32743

Coefficients: Estimate Std. Error t value Pr(>|t|)

(Intercept) 6.466e-01 1.232e-02 52.475 <2e-16 \*\*\* siteMKT 2.700e-03 1.006e-02 0.268 0.788

siteMLA 2.224e-04 1.038e-02 0.021 0.983

siteWPA -8.707e-05 1.031e-02 -0.008 0.993

year 9.458e-04 1.035e-03 0.914 0.361

— Signif. codes: 0 ‘’ **0.001** ‘’ 0.01 ‘’ 0.05 ‘’ 0.1 ‘’ 1

Residual standard error: 0.1348 on 2227 degrees of freedom Multiple R-squared: 0.0004532, Adjusted R-squared: -0.001342 F-statistic: 0.2524 on 4 and 2227 DF, p-value: 0.9083

Analysis of Variance Table

Response: freq Df Sum Sq Mean Sq F value Pr(>F) site 3 0.003 0.0010546 0.0580 0.9817 year 1 0.015 0.0151932 0.8358 0.3607 Residuals 2227 40.484 0.0181788

Estimate Std. Error t value Pr(>|t|) -2.685351e-02 4.688386e-03 -5.727666e+00 1.608018e-05 [1] “11”\*\*\*\*\*

**Frequencies by site (test differences at individual snp-level using the data partitioned by year)**

**Do proportion frequencies fall within CIs of inferred frequencies?**

## [1] 0.7526882

## MKT MLA MKK WPA

## 0.7096774 0.6021505 0.9677419 0.7311828

## [1] 0.859767