Tarea lunes 11 de abril

Eduardo Castro T.

April 11, 2016

Como primer paso extraemos los identificadores del archivo generado .xls mediante la línea de codigo siguiente:

```
#cut -f8 *xls|sed '1d' | sed '/^.$/d'|sed 's/`/\n/g'| sed 's/\^.*//'> Pfam.tx
```

```
##              V1
## 1    PF04502.10
## 2    PF01545.18
## 3    PF00763.20
## 4    PF02882.16
## 5    PF01496.16
## 6    PF00248.18
## 7    PF00400.29
## 8    PF00400.29
## 9    PF00400.29
## 10    PF09070.8
## 11    PF08324.8
## 12    PF08611.7
## 13   PF00248.18
## 14   PF01965.21
## 15   PF01115.14
## 16   PF05221.14
## 17   PF05221.14
## 18   PF00670.18
## 19   PF02826.16
## 20    PF07991.9
## 21   PF00025.18
## 22    PF09439.7
## 23   PF08477.10
## 24    PF04670.9
## 25   PF01926.20
## 26   PF00071.19
## 27   PF00687.18
## 28   PF00293.25
## 29   PF03372.20
## 30   PF02179.13
## 31    PF03981.9
## 32   PF00076.19
## 33   PF01480.14
## 34    PF10419.6
## 35    PF10446.6
## 36   PF00956.15
## 37    PF11976.5
## 38   PF00240.20
## 39    PF13881.3
## 40    PF14560.3
## 41   PF01020.14
## 42   PF02861.17
```

```
## 43   PF02861.17
## 44   PF00004.26
## 45   PF00158.23
## 46   PF07724.11
## 47   PF07728.11
## 48   PF00004.26
## 49    PF10431.6
## 50   PF04031.10
## 51   PF07946.11
## 52   PF00988.19
## 53   PF00117.25
## 54   PF00400.29
## 55   PF00400.29
## 56    PF13243.3
## 57   PF00432.18
## 58   PF00432.18
## 59    PF13249.3
## 60   PF00432.18
## 61   PF04130.10
## 62   PF04130.10
## 63   PF01055.23
## 64   PF05023.11
## 65   PF06728.10
## 66   PF00160.18
## 67    PF13181.3
## 68   PF00117.25
## 69   PF07722.10
## 70   PF04715.10
## 71   PF00425.15
## 72    PF15411.3
## 73   PF00258.22
## 74   PF03460.14
## 75   PF01077.19
## 76   PF03460.14
## 77   PF01077.19
## 78    PF10422.6
## 79   PF00012.17
## 80   PF06723.10
## 81    PF10712.6
## 82   PF00012.17
## 83   PF06723.10
## 84   PF00012.17
## 85   PF06723.10
## 86   PF00012.17
## 87    PF13589.3
## 88   PF02518.23
## 89   PF00183.15
## 90    PF12382.5
## 91   PF00078.24
## 92   PF00665.23
## 93   PF00078.24
## 94   PF00078.24
## 95   PF00665.23
## 96    PF12382.5
```

```
## 97   PF02525.14
## 98    PF08585.9
## 99    PF06395.8
## 100   PF04855.9
## 101   PF04855.9
## 102   PF09792.6
## 103   PF09792.6
## 104   PF08558.7
## 105   PF08558.7
## 106   PF16212.2
## 107  PF00122.17
## 108   PF13246.3
## 109  PF00011.18
## 110  PF04969.13
## 111  PF00995.20
## 112  PF00995.20
## 113   PF16651.2
## 114  PF00076.19
## 115  PF00076.19
## 116   PF13893.3
## 117  PF00076.19
## 118   PF13893.3
## 119   PF16651.2
## 120   PF16461.2
## 121  PF07575.10
## 122   PF16413.2
## 123   PF05536.8
## 124   PF12766.4
## 125   PF08159.9
## 126  PF01246.17
## 127  PF07738.10
## 128   PF13793.3
## 129  PF00156.24
## 130   PF14572.3
## 131  PF07500.11
## 132  PF01096.15
## 133   PF08711.8
## 134  PF00153.24
## 135  PF00118.21
## 136  PF02259.20
## 137   PF08771.8
## 138  PF00454.24
## 139  PF02260.17
## 140   PF08771.8
## 141  PF00621.17
## 142  PF00780.19
## 143  PF00780.19
## 144   PF15405.3
## 145  PF04082.15
## 146  PF04083.13
## 147  PF00561.17
## 148  PF03959.10
## 149  PF02364.12
## 150  PF00176.20
```

```
## 151 PF00443.26
## 152  PF13423.3
## 153   PF14533.3
## 154 PF00443.26
## 155 PF07716.12
## 156 PF00324.18
## 157 PF03522.12
## 158  PF10300.6
## 159 PF00128.21
## 160 PF00773.16
## 161 PF00773.16
## 162  PF06198.8
## 163 PF03104.16
## 164 PF00136.18
## 165  PF14260.3
## 166  PF14703.3
## 167 PF02714.12
## 168  PF12621.5
## 169  PF13967.3
## 170  PF09724.6
## 171 PF00400.29
## 172 PF04851.12
## 173 PF00176.20
## 174 PF00271.28
## 175 PF00176.20
## 176 PF00133.19
## 177  PF13603.3
## 178 PF00133.19
## 179 PF00133.19
## 180 PF08264.10
## 181 PF03178.12
## 182 PF00275.17
## 183 PF01202.19
## 184 PF01487.12
## 185  PF08501.8
## 186 PF01761.17
## 187 PF01761.17
## 188  PF08596.7
## 189  PF08596.7
## 190  PF12719.4
## 191  PF13520.3
## 192 PF00324.18
## 193  PF13520.3
## 194 PF00324.18
## 195  PF12471.5
## 196 PF00787.21
## 197  PF08628.9
## 198 PF02194.12
## 199 PF04658.10
## 200  PF12341.5
## 201 PF01544.15
## 202 PF01544.15
## 203 PF01544.15
## 204 PF00171.19
```

```
## 205  PF08449.8
## 206  PF08641.9
## 207   PF13561.3
## 208 PF00106.22
## 209 PF00179.23
## 210 PF03547.15
## 211 PF03547.15
## 212 PF04032.13
## 213   PF14821.3
## 214 PF00291.22
## 215   PF14821.3
## 216 PF00291.22
## 217 PF04930.12
## 218 PF00348.14
## 219 PF00069.22
## 220 PF07714.14
## 221 PF01636.20
## 222 PF02966.13
## 223 PF02714.12
## 224 PF02714.12
## 225 PF00957.18
## 226 PF00339.26
## 227 PF02752.19
## 228   PF07019.9
## 229   PF13920.3
## 230   PF13923.3
## 231   PF13639.3
## 232   PF14634.3
## 233 PF00097.22
## 234   PF13445.3
## 235 PF02190.13
## 236   PF04939.9
## 237 PF00400.29
## 238 PF00400.29
## 239 PF00400.29
## 240 PF00248.18
## 241 PF00962.19
## 242   PF10253.6
## 243   PF10198.6
## 244   PF16881.2
## 245 PF04055.18
## 246 PF00443.26
## 247   PF13423.3
## 248 PF00443.26
## 249   PF13324.3
## 250 PF01217.17
## 251 PF03403.10
## 252 PF03403.10
## 253 PF03061.19
## 254   PF09197.7
## 255   PF09197.7
## 256 PF01704.15
## 257 PF00484.16
## 258 PF00484.16
```
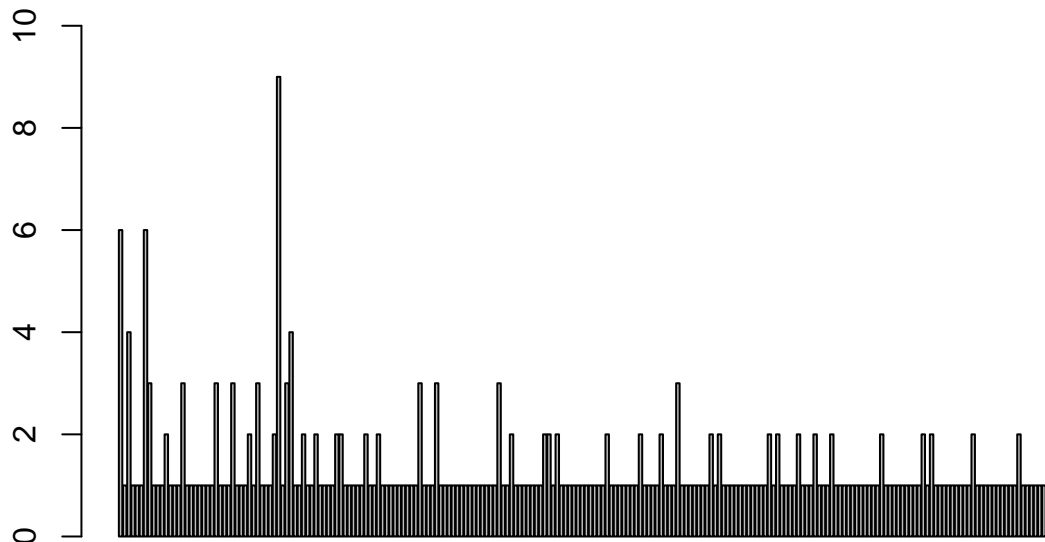
```
## 259 PF00004.26
## 260 PF00004.26
## 261 PF02933.14
## 262 PF07724.11
## 263 PF00004.26
## 264 PF00004.26
## 265 PF02115.14
## 266 PF00696.25
## 267  PF13840.3
## 268 PF01842.22
## 269 PF04427.15
## 270 PF00465.16
## 271  PF13685.3
## 272 PF01761.17
## 273 PF01398.18
## 274 PF03914.14
## 275  PF09073.7
## 276 PF00349.18
## 277 PF03727.13
## 278 PF00164.22
## 279 PF07732.12
## 280 PF00394.19
## 281 PF07731.11
## 282 PF00394.19
## 283 PF07731.11
## 284  PF08783.8
## 285  PF13696.3
## 286  PF13917.3
## 287 PF00098.20
## 288  PF10373.6
## 289  PF07798.8
## 290 PF01841.16
## 291 PF03835.12
## 292 PF02116.12
## 293 PF04146.12
## 294 PF00076.19
## 295 PF00076.19
## 296  PF09768.6
## 297  PF15638.3
## 298  PF04124.9
```

En la tabla podemos ver la lista de los identificadores.

Los graficamos

PF00004.26   PF00687.18   PF03061.19   PF07798.8   PF13324.3

Ordenamos e imprimimos una tabla con los identificadores organizados para observar de forma clara los identificades con una mayor frecuencia.

```
#cat <(cut -f 8 *.xls | sed 's/\^.*//' | sed 's/^\.//') <(cut -f 8 *.xls | grep '`PF' | sed 's/.*`//' |
```

```
print(PfamJE)
```

```
##      V1        V2
## 1     1 PF00011.18
## 2     1 PF00025.18
## 3     1 PF00069.22
## 4     1 PF00071.19
## 5     1 PF00098.20
## 6     1 PF00106.22
## 7     1 PF00118.21
## 8     1 PF00122.17
## 9     1 PF00128.21
## 10    1 PF00153.24
## 11    1 PF00160.18
## 12    1 PF00164.22
## 13    1 PF00171.19
## 14    1 PF00179.23
## 15    1 PF00183.15
## 16    1 PF00258.22
## 17    1 PF00271.28
## 18    1 PF00275.17
## 19    1 PF00293.25
## 20    1 PF00339.26
## 21    1 PF00348.14
## 22    1 PF00349.18
## 23    1 PF00394.19
## 24    1 PF00425.15
## 25    1 PF00454.24
## 26    1 PF00465.16
```

```
## 27     1 PF00561.17
## 28     1 PF00621.17
## 29     1 PF00687.18
## 30     1 PF00696.25
## 31     1 PF00763.20
## 32     1 PF00787.21
## 33     1 PF00956.15
## 34     1 PF00957.18
## 35     1 PF00962.19
## 36     1 PF00988.19
## 37     1 PF01020.14
## 38     1 PF01055.23
## 39     1 PF01077.19
## 40     1 PF01096.15
## 41     1 PF01115.14
## 42     1 PF01217.17
## 43     1 PF01246.17
## 44     1 PF01398.18
## 45     1 PF01480.14
## 46     1 PF01496.16
## 47     1 PF01545.18
## 48     1 PF01636.20
## 49     1 PF01704.15
## 50     1 PF01841.16
## 51     1 PF01842.22
## 52     1 PF01965.21
## 53     1 PF02115.14
## 54     1 PF02116.12
## 55     1 PF02179.13
## 56     1 PF02190.13
## 57     1 PF02194.12
## 58     1 PF02259.20
## 59     1 PF02260.17
## 60     1 PF02364.12
## 61     1 PF02525.14
## 62     1 PF02752.19
## 63     1 PF02861.17
## 64     1 PF02882.16
## 65     1 PF02933.14
## 66     1 PF02966.13
## 67     1 PF03061.19
## 68     1 PF03104.16
## 69     1 PF03178.12
## 70     1 PF03372.20
## 71     1 PF03522.12
## 72     1 PF03727.13
## 73     1 PF03835.12
## 74     1 PF03914.14
## 75     1 PF03959.10
## 76     1  PF03981.9
## 77     1 PF04031.10
## 78     1 PF04032.13
## 79     1 PF04055.18
## 80     1 PF04082.15
```

```
## 81    1 PF04083.13
## 82    1  PF04124.9
## 83    1 PF04146.12
## 84    1 PF04427.15
## 85    1 PF04502.10
## 86    1 PF04658.10
## 87    1 PF04715.10
## 88    1 PF04851.12
## 89    1 PF04930.12
## 90    1  PF04939.9
## 91    1 PF04969.13
## 92    1 PF05023.11
## 93    1 PF05221.14
## 94    1  PF05536.8
## 95    1  PF06198.8
## 96    1  PF06395.8
## 97    1 PF06728.10
## 98    1  PF07019.9
## 99    1 PF07500.11
## 100   1 PF07575.10
## 101   1 PF07716.12
## 102   1 PF07722.10
## 103   1 PF07732.12
## 104   1 PF07738.10
## 105   1  PF07798.8
## 106   1 PF07946.11
## 107   1  PF07991.9
## 108   1  PF08159.9
## 109   1 PF08264.10
## 110   1  PF08324.8
## 111   1  PF08449.8
## 112   1  PF08501.8
## 113   1  PF08585.9
## 114   1  PF08611.7
## 115   1  PF08628.9
## 116   1  PF08641.9
## 117   1  PF08711.8
## 118   1  PF08783.8
## 119   1  PF09073.7
## 120   1  PF09724.6
## 121   1  PF09768.6
## 122   1  PF10198.6
## 123   1  PF10253.6
## 124   1  PF10300.6
## 125   1  PF10373.6
## 126   1  PF10419.6
## 127   1  PF10422.6
## 128   1  PF10431.6
## 129   1  PF10446.6
## 130   1  PF10712.6
## 131   1  PF11976.5
## 132   1  PF12341.5
## 133   1  PF12471.5
## 134   1  PF12621.5
```

```
## 135   1  PF12719.4
## 136   1  PF12766.4
## 137   1  PF13181.3
## 138   1  PF13243.3
## 139   1  PF13246.3
## 140   1  PF13249.3
## 141   1  PF13324.3
## 142   1  PF13561.3
## 143   1  PF13589.3
## 144   1  PF13603.3
## 145   1  PF13793.3
## 146   1  PF13893.3
## 147   1  PF13920.3
## 148   1  PF13967.3
## 149   1  PF14260.3
## 150   1  PF14533.3
## 151   1  PF14572.3
## 152   1  PF14703.3
## 153   1  PF15405.3
## 154   1  PF15411.3
## 155   1  PF15638.3
## 156   1  PF16212.2
## 157   1  PF16413.2
## 158   1  PF16461.2
## 159   1  PF16881.2
## 160   1       Pfam
## 161   2  PF00078.24
## 162   2  PF00117.25
## 163   2  PF00133.19
## 164   2  PF00291.22
## 165   2  PF00432.18
## 166   2  PF00484.16
## 167   2  PF00665.23
## 168   2  PF00773.16
## 169   2  PF00780.19
## 170   2  PF00995.20
## 171   2  PF02714.12
## 172   2  PF03403.10
## 173   2  PF03547.15
## 174   2  PF04130.10
## 175   2   PF04855.9
## 176   2  PF07731.11
## 177   2   PF08558.7
## 178   2   PF08596.7
## 179   2   PF08771.8
## 180   2   PF09197.7
## 181   2   PF09792.6
## 182   2   PF12382.5
## 183   2   PF13423.3
## 184   2   PF13520.3
## 185   2   PF14821.3
## 186   2   PF16651.2
## 187   3  PF00004.26
## 188   3  PF00176.20
```

```
## 189   3 PF00248.18
## 190   3 PF00324.18
## 191   3 PF01544.15
## 192   3 PF01761.17
## 193   3 PF06723.10
## 194   4 PF00012.17
## 195   4 PF00443.26
## 196   5 PF00076.19
## 197   6 PF00400.29
```

observando que los siguientes dominios son los mas frecuentes:

1.- PF00012.17 - 4

2.- PF00443.26 - 4

3.- PF00076.19 - 5

4.- PF00400.29 - 6

Y la ultima es un dominio denominado WD40.