



# PSODA

## Open Source Phylogenetic Search and DNA Analysis

Quinn Snell, Mark Clement, Kenneth Sundberg

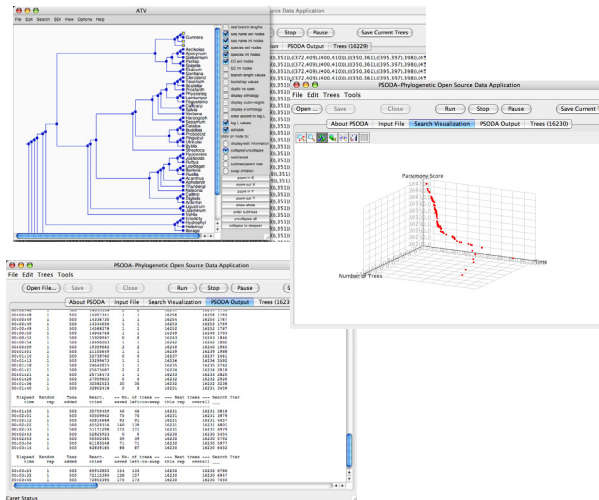
Brigham Young University

{snell,clement,kasundberg}@cs.byu.edu

<http://dna.cs.byu.edu/psoda> --- svn co <http://dna.cs.byu.edu/opensvn/psoda>

### Abstract

PSODA (sō-də) is an open source (GPL v2) sequence analysis package that implements sequence alignment using biochemical properties, phylogeny search with parsimony or maximum likelihood criteria and selection detection using biochemical properties (TreeSAAP). PSODA is compatible with PAUP\* and the search algorithms are competitive with those in PAUP\*. PSODA also adds a basic scripting language to the PAUP block, making it possible to easily create advanced meta-searches. Because PSODA is open-source, we have also been able to easily add in advanced search techniques and characterize the benefits of various optimizations. PSODA is available for Macintosh OS X, Windows, and Linux.

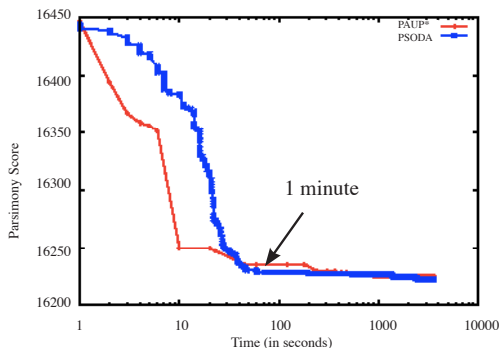


### Advantages of PSODA

- High Performance
- Open Source (-: FREE :-)
- Modular Design (easy algorithm development)
- Advanced Scripting Language
  - makes advanced meta-searches simple
- Reads and Executes PAUP nexus files
- PSODA is competitive with PAUP\*

### Advanced Scripting Language

- Added functionality for PAUP blocks.
- Decision Statements & Loops
- Advanced Functions & User-defined Functions
- Easily Extensible
- Easy scripting of advanced meta-searches such as:
  - Ratchet (Parsimony and Likelihood)
  - DCM and more.



### PAUP\* Ratchet

```

BEGIN RATCHET;
  SET CRITERION=parsimony;
  SET REWEIGHT;
  INCREASE = 10;
  SEARCH START=current swap%50;
  WEIGHTS 2 1 14 17 20 27 31 34 40 50 52 54 57 62 63 64 65 70 77 84 91 92 100 103 107 110 117 121;
  125 126 127 131 133 134 140 142 150 162 165 173 176 183 187 191 197 198 202 204 209 216 222;
  225 226 230 237 238 240 241 244 248 252 254 255 258 269 276 278 284 284 291 294 295 297 309 310 315;
  319 321 322 325 326 332 340 346 350 352 354 355 359 360 362 365 370 372 376 380 384 390 394 400 408 412;
  416 422 425 428 432 434 437 441 444 448 452 454 455 458 462 465 468 472 475 478 482 485 489 493 496 498;
  SEARCH START=current swap%50;
  WEIGHTS 1;
  SEARCH START=current swap%50;
  WEIGHTS 2 1 17 20 22 24 26 28 30 32 34 36 40 43 45 47 50 52 54 56 60 62 64 65 70 72 84 90 92 100 107 110 112;
  116 122 124 126 130 132 134 140 142 150 162 165 173 176 183 187 191 197 198 202 204 207 214 217 220 222;
  224 226 234 236 238 240 244 248 252 254 255 258 269 276 278 284 284 291 294 295 297 309 310 315 319 322 324 329;
  335 347 348 352 352 354 356 360 362 365 370 372 376 380 384 390 394 399 402 408 412 416 422 425 428 432 434 437 441 444 448 452 454 455 458 462 465 468 472 475 478 482 485 489 493 496;
  SEARCH START=current swap%50;
  WEIGHTS 1;
  SEARCH START=current swap%50;
  WEIGHTS 2 1 1 17 20 22 24 26 28 30 32 34 36 40 43 45 47 50 52 54 56 60 62 64 65 70 72 84 90 92 100 107 110 112;
  116 122 124 126 130 132 134 140 142 150 162 165 173 176 183 187 191 197 198 202 204 207 214 217 220 222;
  224 226 234 236 238 240 244 248 252 254 255 258 269 276 278 284 284 291 294 295 297 309 310 315 319 322 324 329;
  335 347 348 352 352 354 356 360 362 365 370 372 376 380 384 390 394 399 402 408 412 416 422 425 428 432 434 437 441 444 448 452 454 455 458 462 465 468 472 475 478 482 485 489 493 496;
  SEARCH START=current swap%50;
  WEIGHTS 1;
  SEARCH START=current swap%50;
  WEIGHTS 2 1 3 9 20 22 24 26 28 30 32 34 36 40 43 45 47 50 52 54 56 60 62 64 65 70 72 84 90 92 100 107 110 112;
  116 122 124 126 130 132 134 140 142 150 162 165 173 176 183 187 191 197 198 202 204 207 214 217 220 222;
  224 226 234 236 238 240 244 248 252 254 255 258 269 276 278 284 284 291 294 295 297 309 310 315 319 322 324 329;
  335 347 348 352 352 354 356 360 362 365 370 372 376 380 384 390 394 399 402 408 412 416 422 425 428 432 434 437 441 444 448 452 454 455 458 462 465 468 472 475 478 482 485 489 493 496;
  SEARCH START=current swap%50;
  WEIGHTS 1;
  SEARCH START=current swap%50;
  WEIGHTS 2 1 3 9 20 22 24 26 28 30 32 34 36 40 43 45 47 50 52 54 56 60 62 64 65 70 72 84 90 92 100 107 110 112;
  116 122 124 126 130 132 134 140 142 150 162 165 173 176 183 187 191 197 198 202 204 207 214 217 220 222;
  224 226 234 236 238 240 244 248 252 254 255 258 269 276 278 284 284 291 294 295 297 309 310 315 319 322 324 329;
  335 347 348 352 352 354 356 360 362 365 370 372 376 380 384 390 394 399 402 408 412 416 422 425 428 432 434 437 441 444 448 452 454 455 458 462 465 468 472 475 478 482 485 489 493 496;
  SEARCH START=current swap%50;
  WEIGHTS 1;
  SEARCH START=current swap%50;
  WEIGHTS 2 1 5 6 19 24 27 31 34 36 40 43 45 47 50 52 54 56 60 62 64 65 70 72 84 90 92 100 107 110 112;
  116 122 124 126 130 132 134 140 142 150 162 165 173 176 183 187 191 197 198 202 204 207 214 217 220 222;
  224 226 234 236 238 240 244 248 252 254 255 258 269 276 278 284 284 291 294 295 297 309 310 315 319 322 324 329;
  335 347 348 352 352 354 356 360 362 365 370 372 376 380 384 390 394 399 402 408 412 416 422 425 428 432 434 437 441 444 448 452 454 455 458 462 465 468 472 475 478 482 485 489 493 496;
  SEARCH START=current swap%50;
  WEIGHTS 1;
  SEARCH START=current swap%50;
  REPEATED TEXT = continue. However, it is unclear when to stop.
  END;
  
```

### PSODA Likelihood Ratchet

```

BEGIN PAUP;
  BEGIN RANDOMREWEIGHT;
    NUMCHARS = GETWEIGHTSLENGTH();
    NUMWEIGHTS = NUMCHARS / PERCENT;

    J = 0;
    WHILE (J < NUMWEIGHTS)
      WEIGHT = RANDOM(MAX = RANGE);
      COL = RANDOM(MAX = NUMCHARS) + 1;
      WEIGHTS[COL] = J++;
    ENDWHILE;

    SET MAXTREES = 1 NREPS = 5;
    WHILE (TRUE)
      SET CRITERION=likelihood;
      HEARCH START = CURRENT SWAP = TBR;
      RANGE = 3;

      WEIGHTS RESET;

      SET CRITERION=parsimony;
      HEARCH START = CURRENT SWAP = TBR;
      RANGE = 3;
    ENDWHILE;
  END;
  
```

### PSODA Features

- Parsimony and Likelihood (RAxML) search
- Bayesian methods (Mr. Bayes)
- Consensus (strict and majority rules)
- Selection Detection (TreeSAAP)
- Graphical User Interface
- Binaries for Mac OS X, Windows and Linux
- Object-oriented C++
  - easy to contribute to new algorithm development
- Available via subversion
  - svn checkout <http://dna.cs.byu.edu/opensvn/psoda>



<http://dna.cs.byu.edu/psoda>

