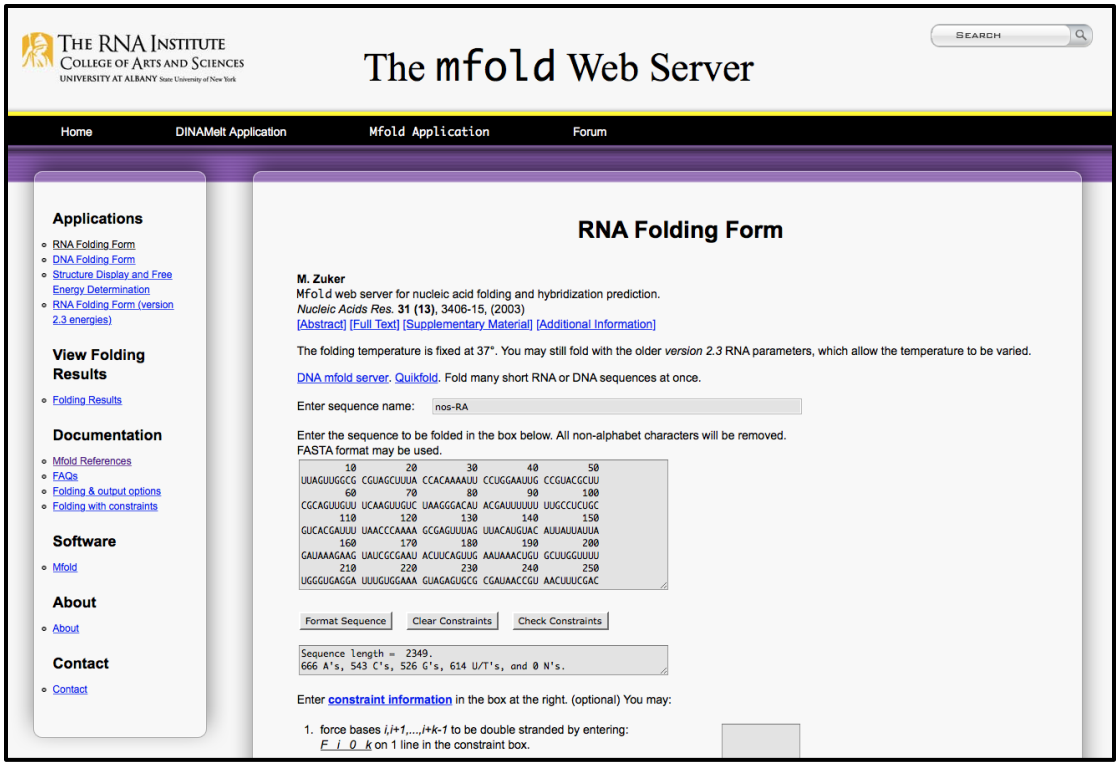
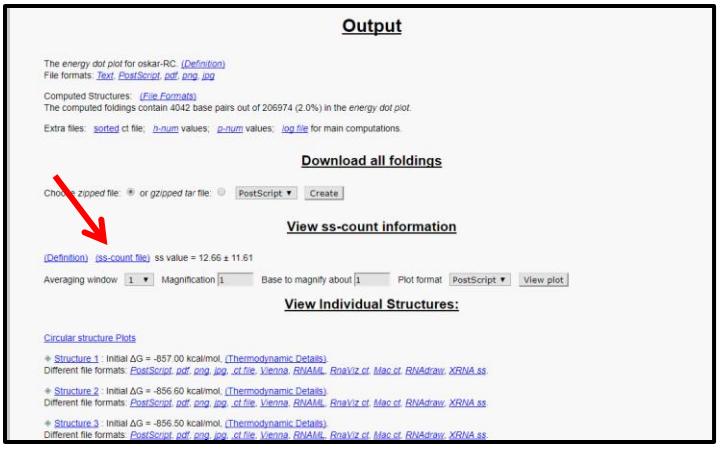
Tutorial for *PinMol* Mac version (contact us by email if you have questions– icatrina@gmail.com):

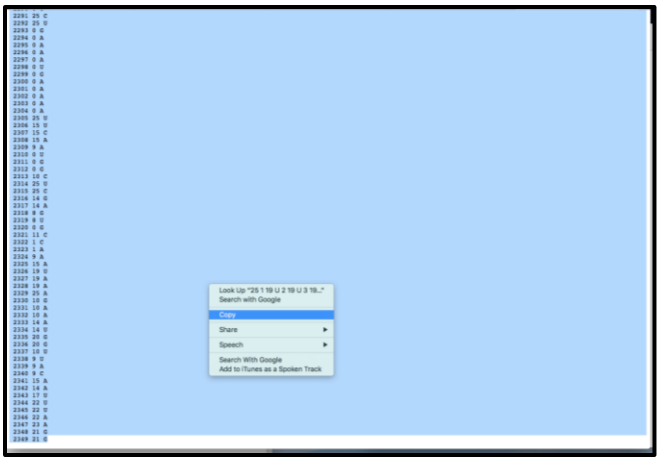
1. **Generate input file using mfold:**
   http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form
   Paste/upload your transcript sequence (*e.g. nanos* mRNA) in FASTA format, and run an immediate (< 800 bases) OR batch job (800-9,000 bases).



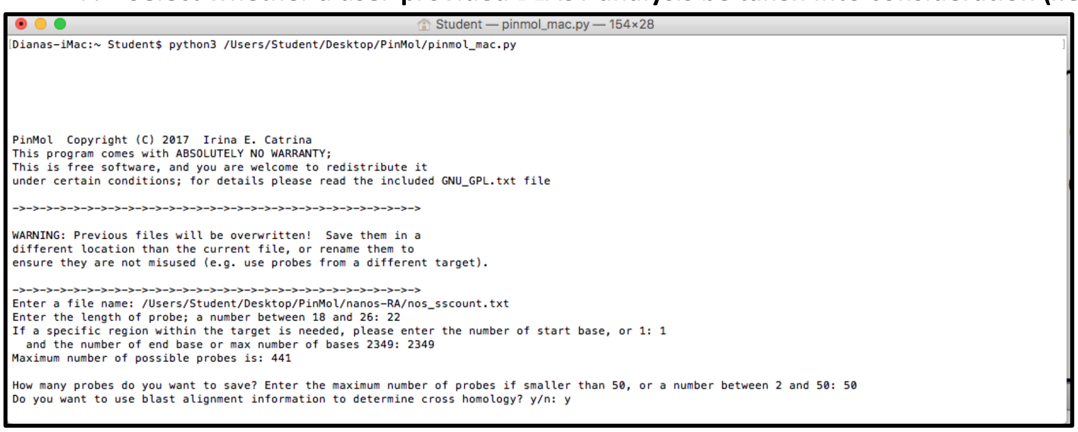
2. **Input file:** Once the *mfold* job is completed, save the ss-count file as an **ASCII/ANSI** text file using the *Notepad*, *Atom* or *MS Word* software. When clicking the "ss-count " link (red arrow), a new window opens with the ss-count information.
   Select the whole text (Win: CTRL+A; Mac: Command+A), copy/paste it into a text editor, and then save it (*e.g.* "nos_sscount.txt").

   *If you use *MS Word*, after the text is pasted, use: *File>Save As>Simple Text>Save>Other encoding>US ASCII*

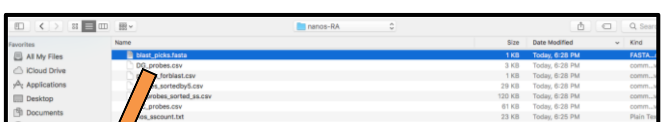


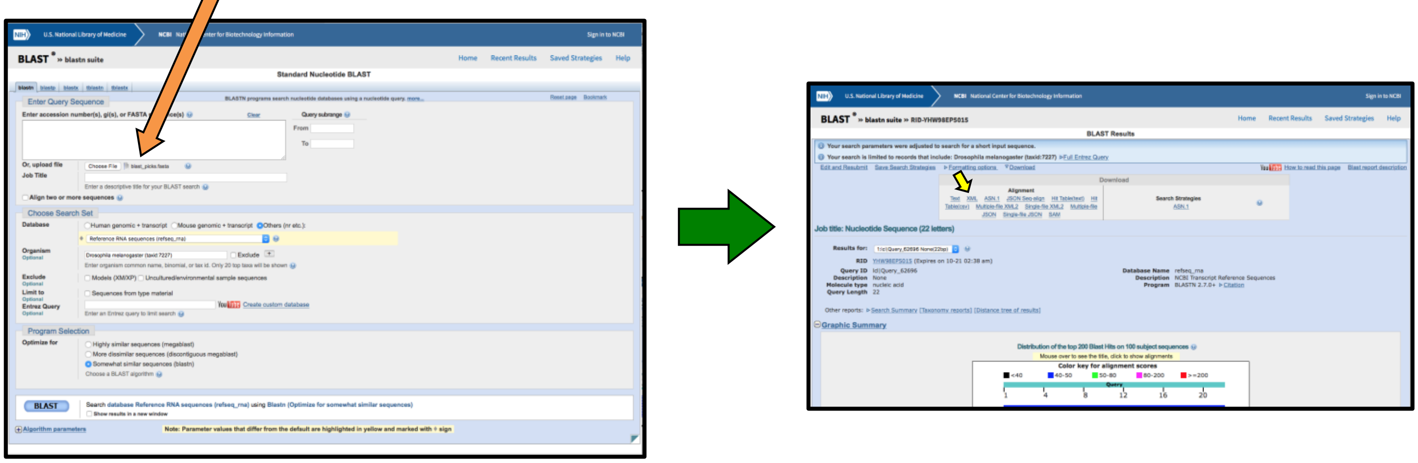


3. **Run the *PinMol* software from a terminal window and follow the instructions.** Each step is described below.
   A. **Run software from command prompt.**
   B. **Input the file path for the ss-count file saved at step 2.**
   Note: Previous files will be overwritten so it is advised to create a new folder if you wish to keep those files.
   C. **Input the length of probe, between 18 and 26 bases (*e.g.* "22").**
   D. **Input target region.** If the full length target OR a limited target region is considered, enter the nucleotide number for the first nucleotide ("1" OR the number for the desired start nucleotide) and last nucleotide (maximum possible – for *nanos-RA* mRNA "2,349" OR the desired number for the end nucleotide).
   Note: *PinMol* provides the total number of nucleotides of the RNA target.
   E. **Input desired number of probes** (*e.g.* "50" or maximum possible if smaller than 50, which is listed in the output – for this *nanos-RA* example is "441", which is the default value if a larger number than the maximum possible is entered).
   F. **Select whether a user-provided BLAST analysis be taken into consideration (*i.e.* "y").**

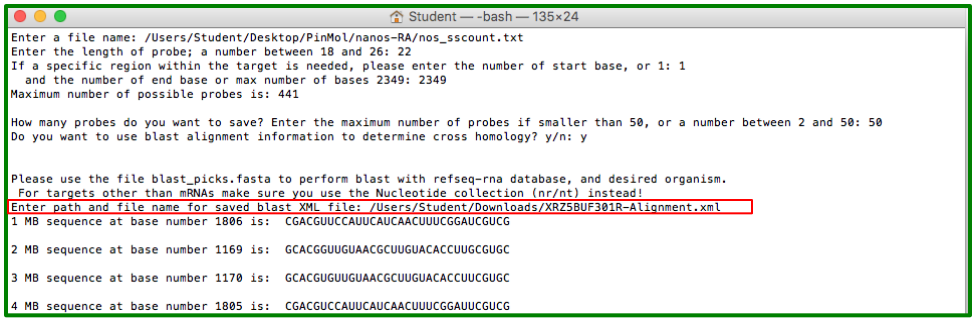


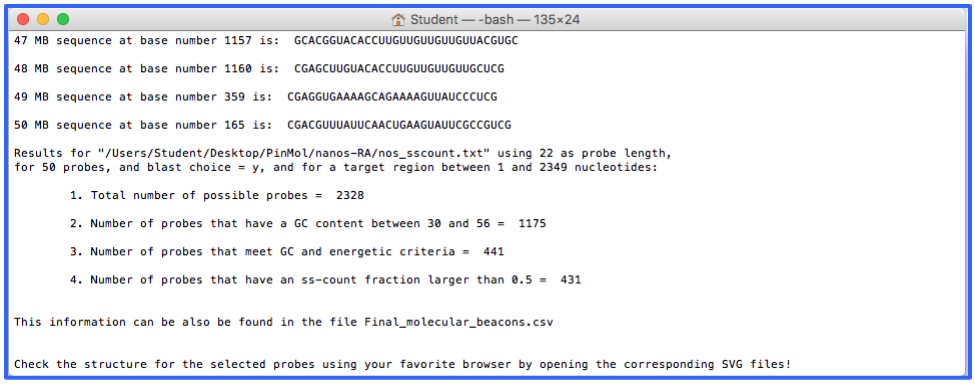
4. **Run blast analysis using probe sequences**. "blast_picks.fasta" file will be created in the folder containing your ss-count file and can be directly inputted into blastn suite with appropriate organism chosen prior to BLAST (*e.g. Drosophila melanogaster*).





Download XML file (yellow arrow) and enter path/file name into *PinMol*. The beginning (green box) and end (blue box) of the output data is presented below.

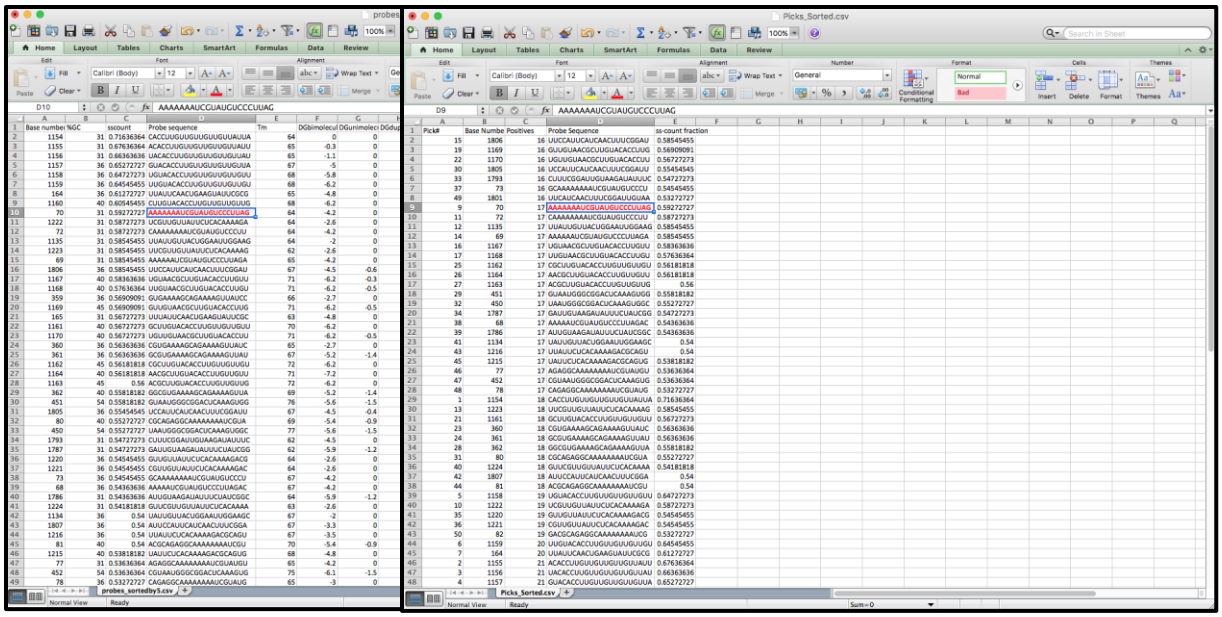




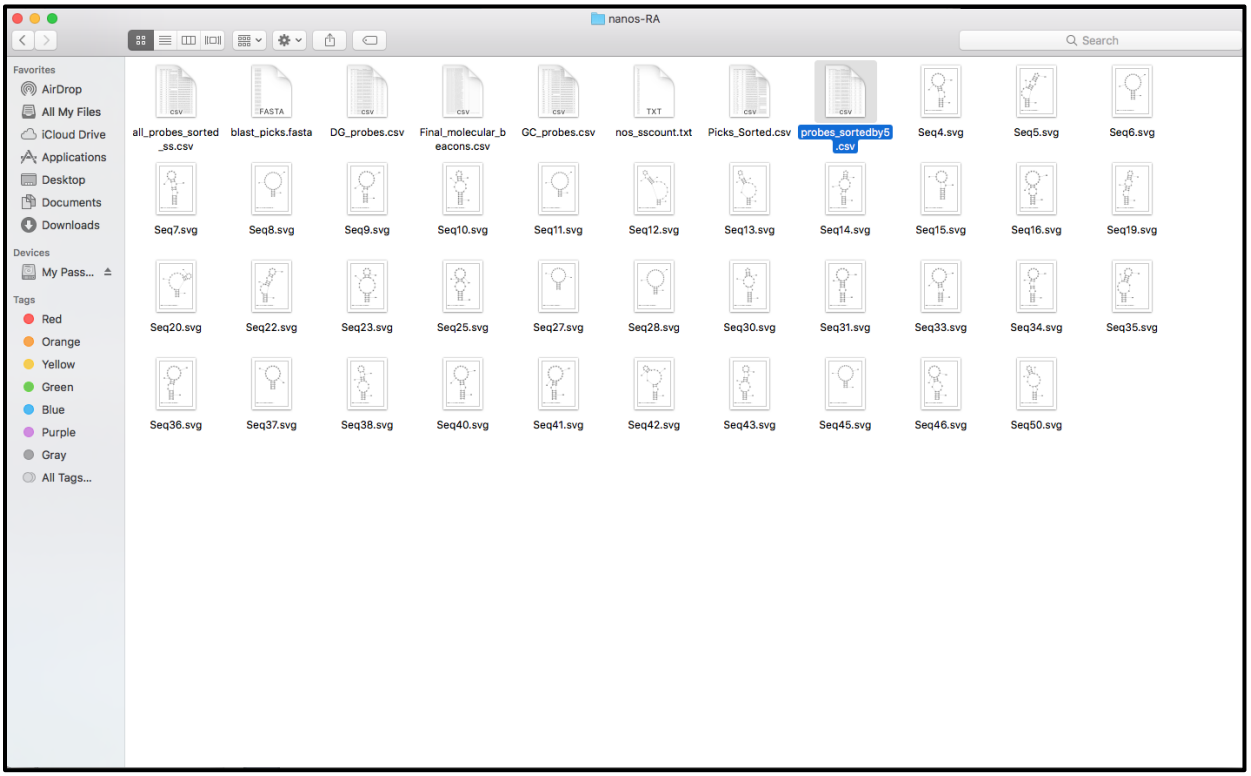
5. **Analyze results.** The output data is saved in the folder containing the input file along with the following files, which are listed in alphabetical order. We used "probe" to refer to the region of the molecular beacon that is complementary to the target region.
    a. **"all_probes_sorted_ss.csv"** – all possible probe sequences sorted in descending order by ss-count fraction – "1" for fully single stranded and "0" for fully double stranded, as predicted in the input file.
    b. **"blast_picks.fasta"** – probe sequences that should be used for BLAST analysis.
    c. **"DG_probes.csv"** – probes that meet both the GC (between 31 and 55%) and the energetic criteria, ΔGunimol > -2.5 kcal/mol and ΔGbimol > -7.5 kcal/mol – see point **f** below and our publication for more information.
    d. **"GC_probes.csv"** – probes that meet the GC content criteria (between 31 and 55%, also see publication for more information).
    e. **"Final_molecular_beacons.csv"** – the output data, as seen in the terminal window.
    f. **"probes_sortedby5.csv"** – probes sorted by five criteria: in descending order by ss-count fraction, ΔGunimol, ΔGbimol and probe's GC percentage, and in ascending order by ΔGduplex. Where the ss-count fraction is described above in point **a**; ΔGunimol is the predicted free energy of uni- or intra-molecular folding of the probe sequence; ΔGbimol is the predicted free energy of bi- or inter-molecular folding of the probe sequence; ΔGduplex is the predicted free energy of hybridization of the probe sequence with a complementary RNA sequence (red highlighted portion of sequence indicates probe).

g. **"Picks_Sorted.csv"** – probes sorted by increasing number of positive hits within genome where 'positives' are defined as number of nucleotide similarities between probe sequence and off-target transcripts (*e.g.* 16 nucleotides). Positive hits are based on the scoring matrix used by BLAST (red highlighted portion of sequence indicates probe sequence).

h. **"Seqi.svg"** – the drawing of the structure of each of the final molecular beacons.



Note: some molecular beacons were discarded (*e.g.* 1 through 3), because they were highly structured or do not fold into a hairpin shape.



6. **Visualize molecular beacon results using a browser.** Open “Seqi.svg” files to view drawing of the final molecular beacons (red highlighted probe sequence and red circle indicate probe region in the final molecular beacon predicted secondary structure). Each “Seqi.svg” file corresponds to the “i” sequence listed for all sorted probes (see point 5g above).