Supporting Information

Faster than Ultra-fast Isotope Generation Algorithm for Biomolecules

Rovshan G. Sadygov[1,2,*]

[1]Department of Biochemistry and Molecular Biology

[2]Sealy Center for Molecular Medicine

The University of Texas Medical Branch

Galveston, Texas 77555

*To whom correspondence should be addressed: Tel. 409-772-3287; Fax 409-772-9670; email: rovshan.sadygov@utmb.edu

Keywords: Poisson distribution, isotope envelope, mass spectrometry, stable isotope labeling.

Abbreviations:

Binom – binomial distribution; FFT – fast Fourier Transform; LC-MS, liquid-chromatography-mass spectrometry; multinom – multinomial distribution, MS, mass spectrometry; ppm – parts per million; ppb – parts per billion.

Tale of Contents:

Table S1. Relative Isotope abundance and mass values used for atomic isotopes in this work.

Element	Isotope	Mass	Mass Difference	Relative Abundance
Hydrogen	^{1}H	1.007825		0.99984426
	^{2}H	2.014102	1.006277	0.00015574
Carbon	^{12}C	11.99945		0.988922
	^{13}C	13.00281	1.00336	0.011078
Nitrogen	^{14}N	14.003074		0.996337
	^{15}N	15.000109	0.997035	0.003663
Oxygen	^{16}O	15.99437		0.9976206
	^{17}O	16.99858	1.00421	0.0003790
	^{18}O	17.99861	2.00424	0.0020004
Phosphor	^{31}P	30.973762		1.0
Sulfur	^{32}S	31.97152		0.9504074
	^{33}S	32.97091	0.99939	0.0074869
	^{34}S	33.96732	1.9958	0.0419599
	^{36}S	35.96653	3.99501	0.0001458

```
procedure FourierIsotopes(nH, nC, nN, nO, nS,  p2H, p13C, p15N,
                          (p17O, p18O), (p33S, p34S, p36S))
{
    double pC[16], pH[12], pN[16], pO[16], pS[16];
    pC[0] = 1 -  p13C; pC[0] = p13C;
    pS[0] = 1 - p33S - p34S - p36S; pS[1] = p33S; pS[2] = p34S; pS[4] = p36S;
    pO[0] = 1 – p17O – p18O ; pO[1] = p17O; pO[2] = p18O;
     //repeat the above procedure for H, N atoms.
    F_ array_C ← FFT(array_C);
    F_ array_O ← FFT(array_O);
    F_ array_N ← FFT(array_N);
    F_ array_H ← FFT(array_H);
    F_ array_S ← FFT(array_S);

    F_product ← FFT(array_C)^nC* FFT(array_H)^nH;
    F_product ← F_product * FFT(array_N)^nN* FFT(array_O)^nO;
    F_product ← F_product * FFT(array_S)^nS;
    Isotopes = invFFT(F_product);
    Normalize(Isotopes);
}
```

Figure S1. Pseudo-code of algorithm for computing isotope distributions using fast-Fourier transforms. The elemental isotope distributions are generated using either binomial (H, C, N) or multinomial (O, S) distributions. Each of the elemental distributions is Fourier transformed (FFT). The product of the Fourier transforms is inverse Fourier transformed (invFFT) and normalized to obtain the isotope distribution of a peptide.

```
#the function computes the first 16 isotopes of Insulin if no arguments are given

Isotope_Pattern = function(nH=771, nC=491, nN=130, nO=138, nS=8)
{
        c12 <- rep(0, 16); h1 <- rep(0, 16);  n14 <- rep(0, 16);
        o16 <- rep(0, 16);   s32 <- rep(0, 16);

         c12[1] = 0.988922;  c12[2] = 0.011078;

         h1[1]   = 0.99984426; h1[2] = 0.00015574;

         n14[1] = 0.996337;     n14[2] = 0.003663;

         o16[1] = 0.9976206;   o16[2] = 0.0003790; o16[3] = 0.0020004;

         s32[1] = 0.9504074;    s32[2] = 0.0074869;

         s32[3] = 0.0419599;    s32[5] = 0.0001458;

    Iso_Pattern = Re(fft(fft(c12)^nC*fft(h1)^nH*fft(n14)^nN * fft(o16)^nO * fft(s32)^nS,
inverse=TRUE))/length(c12)


    Iso_Pattern;
}
```

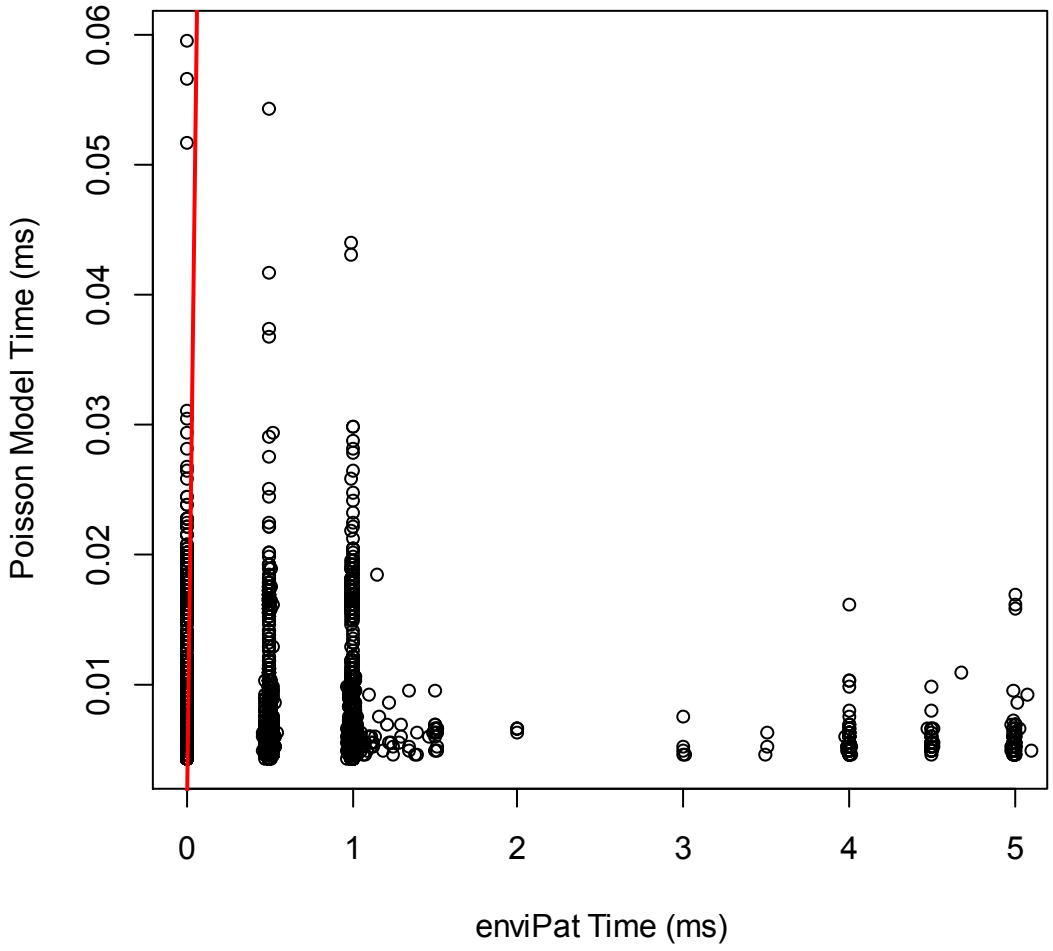

Figure S2. Comparison of Poisson model and enviPat computational times (in milliseconds). The red line is the line of unity.