

# Error simulating large number of reads

I am trying to simulate a full-scale experiment of ~30M, 100-nt reads for Arabidopsis, but I keep getting the same error (also at 16M reads and 8M reads):

```
[INFO] Loading default PCR distribution
```

```
    preparing transcript sequences *****Problems reading 3: 23459833, 74> 23459834 into 100: null
```

```
check for the right species/genome version!
```

I looked at the GTF and FASTA files. The maximum position for chromosome 3 in the GTF file is 23,459,804; the actual chromosome sequence length is 23,459,831. For some reason the simulator is trying to go beyond the end of the last gene in chromosome 3? The full command I'm using is:

```
flux-simulator -t simulator -x -l -s -p a_thaliana_flux.par
```

I have attached the .par file I'm using for 8M reads in case this will help: [a\\_thaliana\\_flux.par](#)

Thanks!