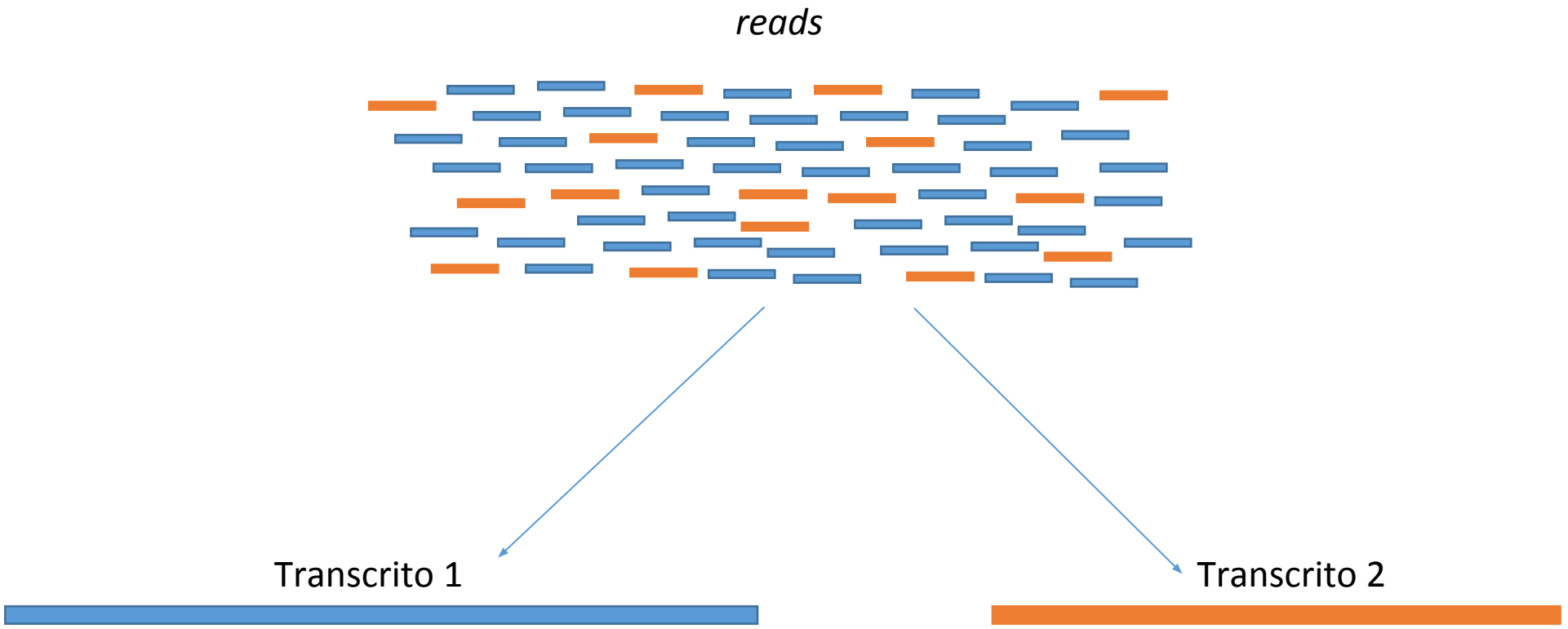
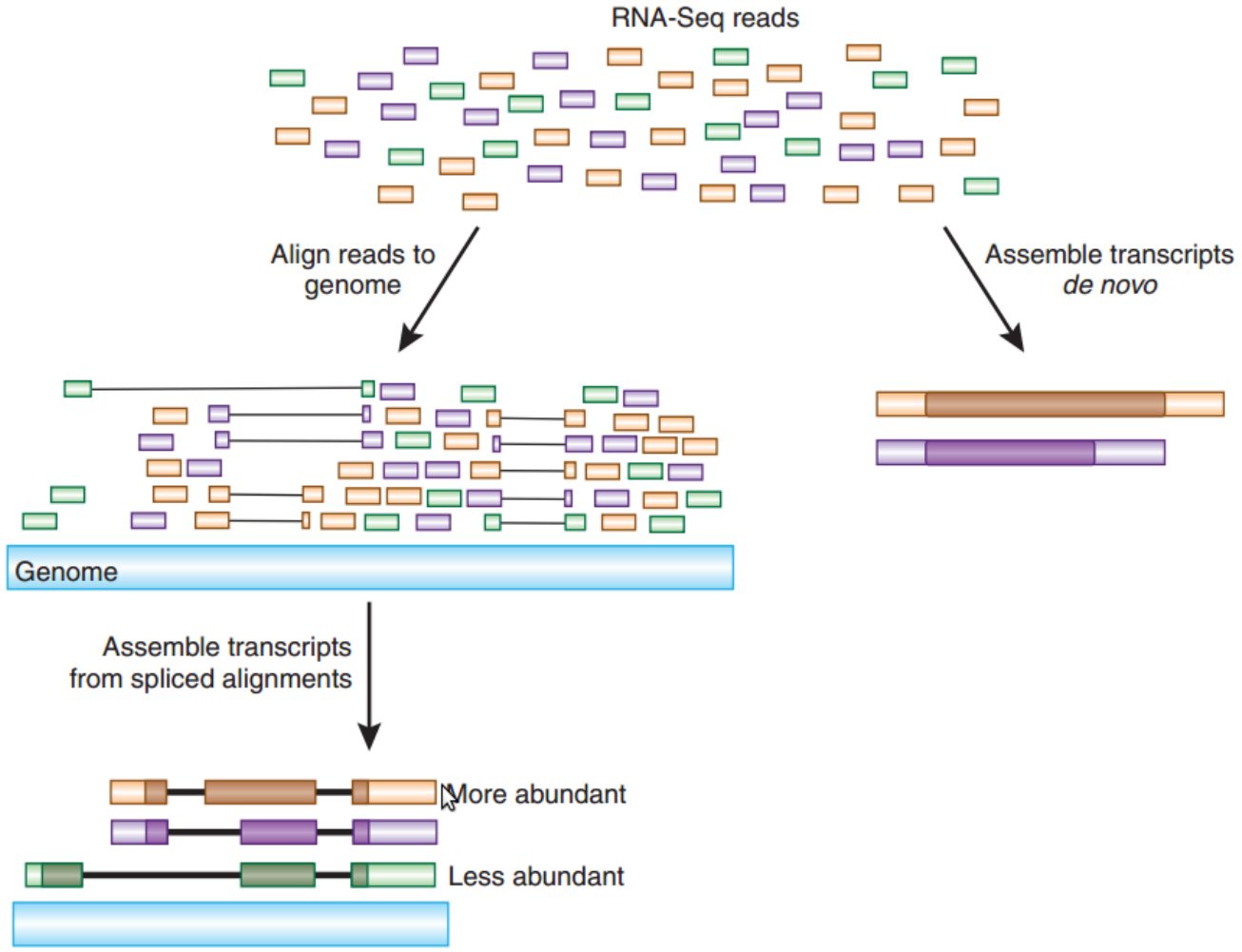
# Montagem *ab initio*

Osvaldo Reis Junior

# Reconstruir transcritos a partir dos *reads*

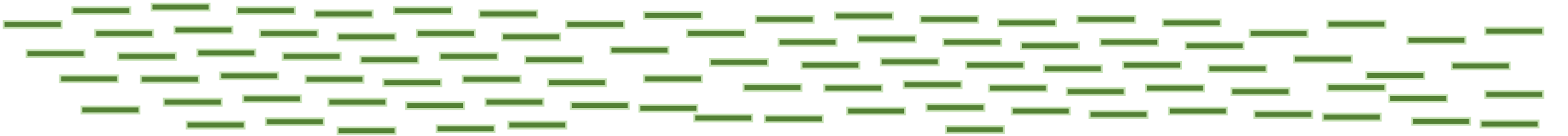


# Dificuldade em montar transcritos pouco expressos

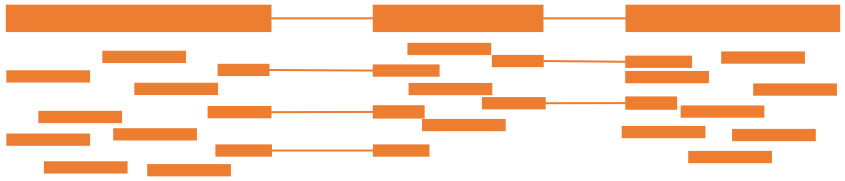


# Cobertura não é uniforme

Genoma

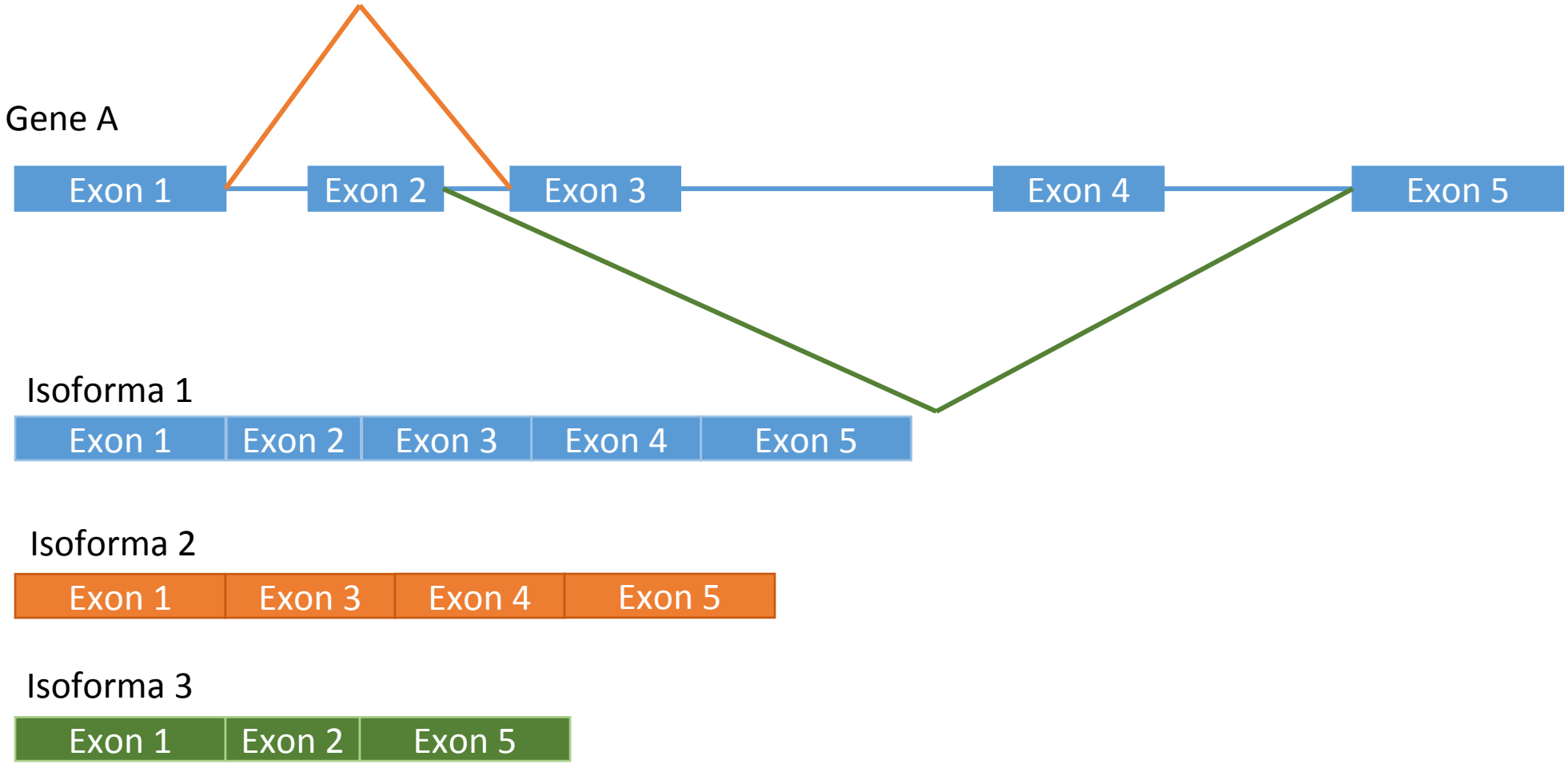


Gene A



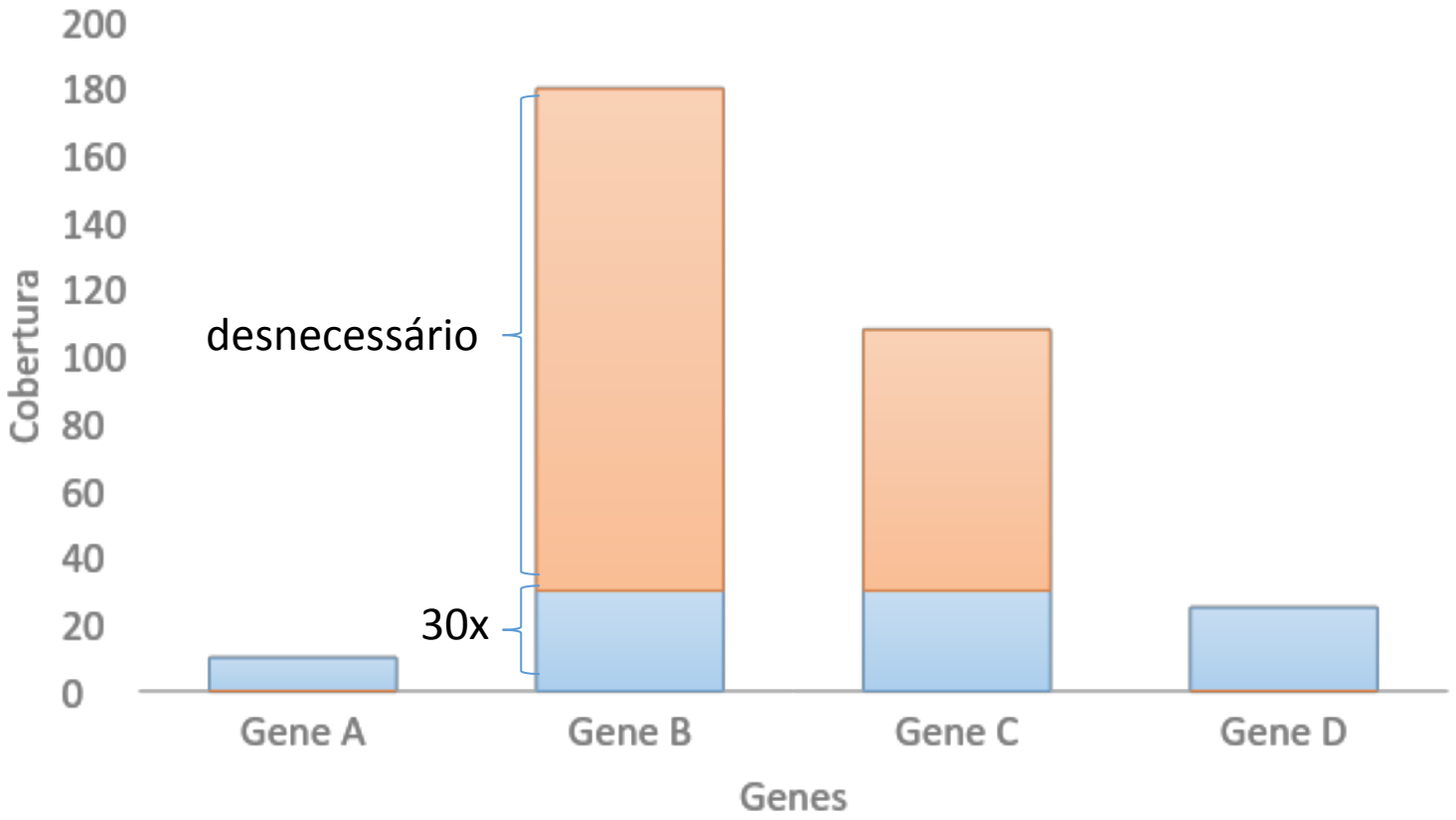
Gene B

# Muitas isoformas de um mesmo gene



# Uso de memória elevado para a montagem

- A solução é normalização *in silico*



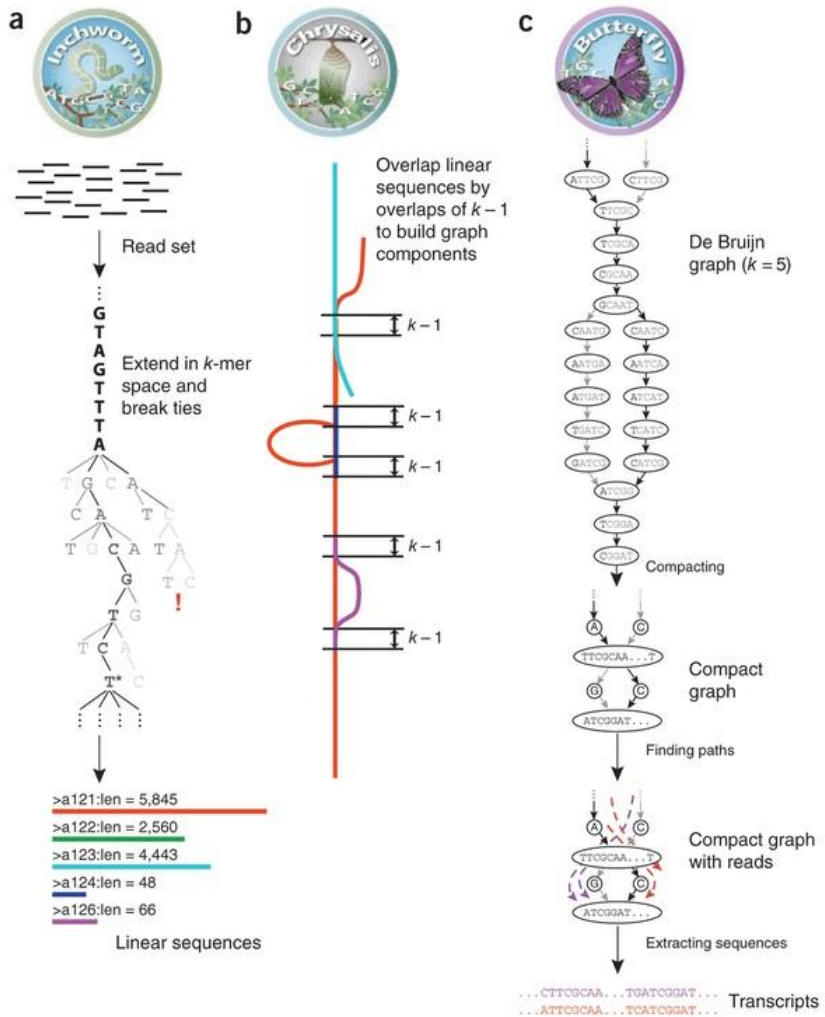
# Montagem *ab initio*

Programas:

- Oases
- Trans-Abyss
- Trinity

# Montagem *ab initio*

Trinity



# Trinity


Required:
**--seqType <string>** :type of reads: ( fa, or fq )

**--max_memory <string>** :suggested max memory to use by Trinity where limiting can be enabled. (jellyfish, sorting, etc) provied in Gb of RAM, ie. '--max_memory 10G'

If paired reads:
**--left <string>** :left reads, one or more (separated by space)
**--right <string>** :right reads, one or more (separated by space)
Or, if unpaired reads:
**--single <string>** :single reads, one or more (note, if single file contains pairs, can use flag: --run_as_paired )

# Trinity


**--SS_lib_type** <string> :Strand-specific RNA-Seq read orientation.
if paired: RF or FR,
if single: F or R. (dUTP method = RF)
See web documentation.

**--CPU <int>** :number of CPUs to use, default: 2

**--min_contig_length <int>** :minimum assembled contig length to report
(def=200)

**--long_reads <string>** :fasta file containing error-corrected or circular consensus (CCS) pac bio reads

**--jaccard_clip** :option, set if you have paired reads and
you expect high gene density with UTR
overlap (use FASTQ input file format
for reads).
(note: jaccard_clip is an expensive
operation, so avoid using it unless
necessary due to finding excessive fusion
transcripts w/o it.)

# Trinity


**--trimmomatic** :run Trimmomatic to quality trim reads
see '--quality_trimming_params' under full usage info for tailored settings.
**--normalize_reads** :run in silico normalization of reads. Defaults to max. read coverage of 50.
see '--normalize_max_read_cov' under full usage info for tailored settings.
**--output <string>** :name of directory for output (will be
created if it doesn't already exist)
default( your current working directory: "/Users/bhaas/GITHUB/trinityrnaseq/trinity_out_dir"
note: must include 'trinity' in the name as a safety precaution! )
**--full_cleanup** :only retain the Trinity fasta file, rename as ${output_dir}.Trinity.fasta