FLEXIDOT USAGE

Version: **1.04**

Citation: Kathrin M. Seibt, Thomas Schmidt, Tony Heitkam (2018)
"FlexiDot: Highly customizable, ambiguity-aware dotplots for visual sequence analyses"
Bioinformatics, doi: 10.1093/bioinformatics/bty395

General usage:
$ python flexidot.py -a [ARGUMENTS]
$ python flexidot.py -i <fasta_file_name> [ARGUMENTS]

GENERAL...

-h, --help	Help screen
-v, --verbose	Verbose

INPUT/OUTPUT OPTIONS... required are [-a] OR [-i]

-a, --auto_fas	Imports all fasta files from current directory (*.fasta, *.fas, *.fa, *.fna) (-i is not necessary) [inactive by default]
-i, --in_file	Input fasta file (fasta file name or comma-separated file list) > Provide multiple files: Recall -i or provide comma-separated file names)
-o, --output_file_prefix	File prefix to be added to the generated filenames [default = NONE]
-c, --collage_output	Multiple dotplots are combined in a collage Y or 1 = ON [default] N or 0 = OFF
-m, --m_col	Number of columns per page (only if --collage_output=y) [default = 4]
-n, --n_row	Number of rows per page (only if --collage_output=y) [default = 5]
-f, --filetype	Output file format 0 = PNG [default] 1 = PDF 2 = SVG
-s, --alphabetic_sorting	Sort sequences alphabetically according to titles Y or 1 = ON N or 0 = OFF [default]

CALCULATION PARAMETERS...

-k, --wordsize	Wordsize (kmer length) for dotplot comparison [default = 7]
-p, --plotting_mode	Mode of FlexiDot dotplotting 0 = self [default] 1 = paired 2 = poly (matrix with all-against-all dotplots) > Run multiple plotting modes: Recall -p or provide comma-separated numbers
-t, --type_nuc	Type of residue is nucleotide Y or 1 = nucleotide [default] N or 0 = amino acid
-w, --wobble_conversion	Ambiguity handling for relaxed matching Y or 1 = ON N or 0 = OFF [default]
-S, --substitution_count	Number of substitutions (mismatches) allowed per window for relaxed matching [default = 0]
-r, --rc_option	Find reverse complementary matches (only if type_nuc=y) Y or 1 = ON [default] N or 0 = OFF

FLEXIDOT USAGE

GRAPHIC FORMATTING...

-A, --line_width	Line width [default = 1]
-B, --line_col_for	Line color [default = black]
-C, --line_col_rev	Reverse line color [default = green]
-D, --x_label_pos	Position of the X-label Y or 1 = top [default] N or 0 = bottom
-E, --label_size	Font size [default = 10]
-F, --spacing	Spacing between all-against-all dotplots (only if --plotting_mode=2) [default = 0.04]
-M, --mirror_y_axis	Flip y-axis bottom to top (cartesian coordinate system) Y or 1 = y-axis bottom to top N or 0 = y-axis top to bottom [default]
-P, --plot_size	Plotsize [default = 10]
-R, --representation	Region of plot to display (only if --plotting_mode=2) 0 = full [default] 1 = upper 2 = lower
-L, --length_scaling	Scale plot size for pairwise comparison (only if --plotting_mode=1) Y or 1 = Scaling ON (axes scaled according to sequence length) N or 0 = Scaling OFF (squared plots) [default]
-T, --title_length	Limit title length for dotplot comparisons [default = 20] Position of selection can be specified by appending a letter (e.g. -T 20E) B = beginning [default] E = end

ANNOTATION SHADING (for -p/--plotting_mode=0,2 only)...

-g, --input_gff_files	GFF3 file used for markup in self-dotplots (provide multiple files: Recall -g or provide comma-separated file names)
-G, --gff_color_config_file	Tab-delimited config file for custom gff shading column 1: feature type column 2: color column 3: alpha column 4: zoom factor (for small regions)

LCS SHADING (for -p/--plotting_mode=2 only)...

-x, --lcs_shading	Shade subdotplot based on the length of the longest common substring (LCS) Y or 1 = ON N or 0 = OFF [default]
-X, --lcs_shading_num	Number of shading intervals (hues) for LCS (-x) and user matrix shading (-u) [default = 5]
-y, --lcs_shading_ref	Reference for LCS shading 0 = maximal LCS length [default] 1 = maximally possible length (length of shorter sequence in pairwise comparison) 2 = given interval sizes - DNA [default 50 bp] or proteins [default 10 aa] - see –Y
-Y, --lcs_shading_interval_len	Length of intervals for LCS shading (only if --lcs_shading_ref=2) [default for nucleotides = 50; default for amino acids = 10]
-z, --lcs_shading_ori	Shade subdotplots according to LCS on 0 = forward [default], 1 = reverse, or 2 = both strands (forward shading above diagonal, reverse shading on diagonal and below; if --input_user_matrix_file used, best LCS is used below diagonal)

USER MATRIX SHADING (for -p/--plotting_mode=2 only)...

-u, --input_user_matrix_file	Shading above diagonal according to values in matrix file specified by the user (tab-delimited or comma-separated matrix with sequence name in column 1 and numbers in columns 2-n, e.g. identity matrix from sequence multiple alignment - strings are ignored)
-U, --user_matrix_print	Display matrix entries in the fields above diagonal of all-against-all dotplots Y or 1 = ON N or 0 = OFF [default]