

FLEXIDOT USAGE

Version: 1.0

Citation: Kathrin M. Seibt, Thomas Schmidt, Tony Heitkam (in prep.)
"FlexiDot: Highly customizable ambiguity-aware dotplots for visual sequence analyses"

General usage: \$ python flexidot.py [ARGUMENTS]

GENERAL...

-h, --help Help screen

-v, --verbose Verbose

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

-a, --auto_fas Imports all fasta from current directory
Y = ON
N = OFF [default]

-i, --in_file Input fasta file (fasta file or comma-separated file list)

-o, --output_file_prefix File prefix to be added to the generated filenames
[default = NONE]

-c, --collage_output Multiple dotplots are printed as a collage
Y = ON [default]
N = OFF

-m, --m_col Number of columns per page
[default = 4] (only if --collage_output is ON)

-n, --n_row Number of rows per page [default = 5] (only if --collage_output is ON)

-f, --filetype Output file format
0 = PNG [default]
1 = PDF
2 = SVG

-t, --type_nuc Type of residue is nucleotide
True = nucleotide [default]
False = amino acid

CALCULATION PARAMETERS...

-k, --wordsize Wordsize (kmer length) for dotplot comparison (default = 7)

-p, --plotting_mode Plotting modes
0 = self [default]
1 = paired
2 = poly (matrix with all-against-all dotplots)

-w, --wobble_conversion Ambiguity handling
Y = ON
N = OFF [default]

-M, --mutation_count Number of mismatch mutations (substitutions) allowed in fuzzy matching
[default = 0]

-r, --rc_option Find reverse complementary matches
Y = ON [default]
N = OFF

-s, --alphabetic_sorting Sort sequence identifiers alphabetically
Y = ON
N = OFF [default]

-q, --quick_mode Activate quick calculation mode
True = quicker, lower resolution, no wobble_conversion, no lcs_shading
False [default]

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 |
| -F, --spacing | Spacing between subdotplots (only if --collage_output is ON)) |
| -P, --plot_size | Plotsize [default = 10] |
| -S, --length_scaling | Scale plot size for pairwise comparison (only if --plotting_modes=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 comparison (only if --plotting_modes=0) [default = infinite] |

ANNOTATION SHADING (for -p/--plotting mode 0 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) |

GFF 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 (number of grey shades besides standard white background) [default = 5] |
| -y, --lcs_shading_ref | Choose reference for LCS shading 0 = maximal LCS length [default] 1 = maximally possible length (length of shorter sequence in pairwise comparison) 2 = given interval sizes (length in bp) |
| -Y, --lcs_shading_interval_len | Length of intervals for LCS shading (only if --lcs_shading_ref=2) [default for nucleotides = 100; 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) |

USER MATRIX SHADING (for -p/--plotting mode 2 only)...

| | |
|------------------------------|--|
| -u, --input_user_matrix_file | Shade top triangle according to values in matrix file (tab-delimited matrix with sequence name in column 1 and numbers in columns 2-n, e.g. identity matrix from MEGA) |
| -U, --user_matrix_print | Print matrix entries in the upper right section of all-against-all dotplots Y or 1 = ON N or 0 = OFF [default] |