

# Работа с текстом: nano, less, grep, cat

CNE 8972 Биоинформатика · 6B05105 Биотехнология  
Неделя 4 · Жаниязов Ж.А., старший преподаватель

# План лекции

- Почему важен текст
- Просмотр файлов: cat, head, tail, less
- Редактор nano
- Поиск grep
- Перенаправление и конвейеры

4 / 15

неделя · работа с текстом

# Почему важен текст

- Большинство биоинформатических файлов текстовые: FASTA, FASTQ, VCF, GFF
- Команды Linux обрабатывают потоки текста
- Файлы велики, открывать их редакторами нельзя
- Нужны быстрые команды просмотра и поиска

**Текст**  
основа данных

# Просмотр: cat, head, tail

| Команда        | Что делает                                  |
|----------------|---------------------------------------------|
| cat файл       | Выводит файл целиком (осторожно с большими) |
| head -n 5 файл | Первые 5 строк                              |
| tail -n 5 файл | Последние 5 строк                           |
| wc -l файл     | Число строк                                 |
| tail -f файл   | Следит за дописыванием файла                |

# Постраничный просмотр: less

- Открывает большой файл по страницам и не загружает его целиком
- Пробел — страница вперёд, b — назад, G — в конец
- /слово — поиск, n — следующее совпадение
- q — выход

The logo for the 'less' command, featuring the word 'less' in a bold, orange, sans-serif font.

q — выход

# Редактор nano

| Сочетание       | Действие                   |
|-----------------|----------------------------|
| nano файл       | Открыть или создать файл   |
| Ctrl+O          | Сохранить                  |
| Ctrl+X          | Выйти                      |
| Ctrl+W          | Поиск                      |
| Ctrl+K / Ctrl+U | Вырезать строку / вставить |

# Поиск: grep

- Ищет строки, содержащие шаблон: `grep "ATG" файл`
- `-i` игнорировать регистр, `-n` показывать номера строк
- `-s` считать совпадения, `-v` выводить строки без шаблона
- Шаблон лучше заключать в кавычки



**grep**  
поиск по шаблону

# grep в биоинформатике

| Задача                            | Команда                                    |
|-----------------------------------|--------------------------------------------|
| Число последовательностей в FASTA | <code>grep -c "&gt;" file.fa</code>        |
| Заголовки последовательностей     | <code>grep "&gt;" file.fa</code>           |
| Поиск гена в аннотации            | <code>grep -i "recA" annotation.gff</code> |
| Строки VCF без комментариев       | <code>grep -v "^#" file.vcf</code>         |

# Перенаправление ввода и вывода

| Символ | Значение                                    |
|--------|---------------------------------------------|
| >      | Записать вывод в файл (перезаписывает)      |
| >>     | Дописать вывод в конец файла                |
| 2>     | Записать сообщения об ошибках               |
|        | Передать вывод следующей команде (конвейер) |

# Конвейер: пример



Короткие команды объединяются в конвейер для решения задачи.

# Сортировка и подсчёт

| Команда | Назначение                               |
|---------|------------------------------------------|
| sort    | Сортировка строк                         |
| uniq -c | Подсчёт повторяющихся строк (после sort) |
| cut -f1 | Выбор столбца по номеру                  |
| wc      | Подсчёт строк, слов и символов           |

# Типичные ошибки

## Большой файл в nano

Используйте less, head или tail

## Перезапись через >

Для дозаписи нужен >>

## Шаблон без кавычек

Специальные символы обрабатываются оболочкой

## Регистр

Для grep нужна опция -i

# Итоги и контрольные вопросы

- Чем less отличается от cat?
- Что делает grep -с?
- Для чего нужен конвейер?
- Практика: ПР №4 «Поиск и редактирование файлов»

50–70

стр. · [1] Lesk