

Методы электронной таблицы

2020, ААл

Бесплатные ЭТ

- **Google Sheets**

- Вход: Google => Drive => New => Google sheet

- LibreOffice

- OpenOffice

- **Microsoft Excel on-line free**

- «Объедините ваши усилия совершенно бесплатно благодаря интернет-версии Microsoft Excel. Сохраняйте электронные таблицы в OneDrive.» Excel 365.

- Это не благотворительность.** Можно пользоваться ограниченное время.

- Потом Microsoft потребует оплату подписки, и если нет – то деинсталлирует и все пропадет.

- До зачёта, наверное, хватит времени.

- Будьте внимательны и осторожны!**

Excel - платная ЭТ

Учимся на примере Excel 2007

- Ограничения по объему
 - MS Excel 2007: не более 1 млн 49 тыс строк
 - Google Sheets: не более 2 млн ячеек

Другие релизы или программы устроены аналогично.

Разбирайтесь сами!

Что нужно нам от электронной таблицы

- Интерактивная обработка массивов данных среднего размера - до нескольких сотен тысяч строк
 - Сортировка, преобразование, вычисление, слияние данных из разных источников, составление сводных таблиц и т.п.
- Представление результатов для публикаций и отчетов
 - Графики, гистограммы, диаграммы и т.п.

Питон ПРОТИВ Excel

Python Guru or Snake in the Grass?



Python for dummy



Excel - улучшенный гематрат
(определение из словаря)



Изумруд после обработки Excel

Питон ПРОТИВ Excel

	python	Excel
Объём данных	ограничен только памятью компьютера	Не более 1млн строк в таблице
Время работы	Можно оптимизировать алгоритм, переписать его на другом языке программирования, C++, например	При 1млн строк даже некоторые простые функции работают часы.
Гибкость	надо ЗАРАНЕЕ знать (придумать) алгоритм, параметры, которые можно менять, способ представления результата	Можно быстро менять способ обработки и представления тех же самых данных
Квалификация	Выше секретарши вплоть до гуру по программированию	от секретарши до квалифицированного биоинформатика
Доступность графических средств	Доступны, но сложнее в освоении и быстроте адаптации	Просты, доступны, в ограниченном, но достаточном объёме

Как уйти от скучного рассказа?

1. Форматы ячеек. Демонстрация. Таблица размеров геномов.
2. Список функций
3. Распространение данных. Демонстрация: вычисление интеграла
4. Специальная вставка. Демонстрация – pr12 в ведомости
5. Гистограммы. Демонстрация – Все634I
6. Статистика. Демонстрация
7. Импорт с приключениями.
8. Выполнение задания «в классе»

1. Форматы ячеек.

- **Демонстрация** - страница размеров геномов
 - Размеры геномов (найти сумму длин)

Задание: скопировать страницу и отформатировать (числа, заголовки, ширина столбца, переносить ли текст в ячейке, сортировка)

Совет: добавить свою строку (строки). Участи будет учтено в рейтинге.

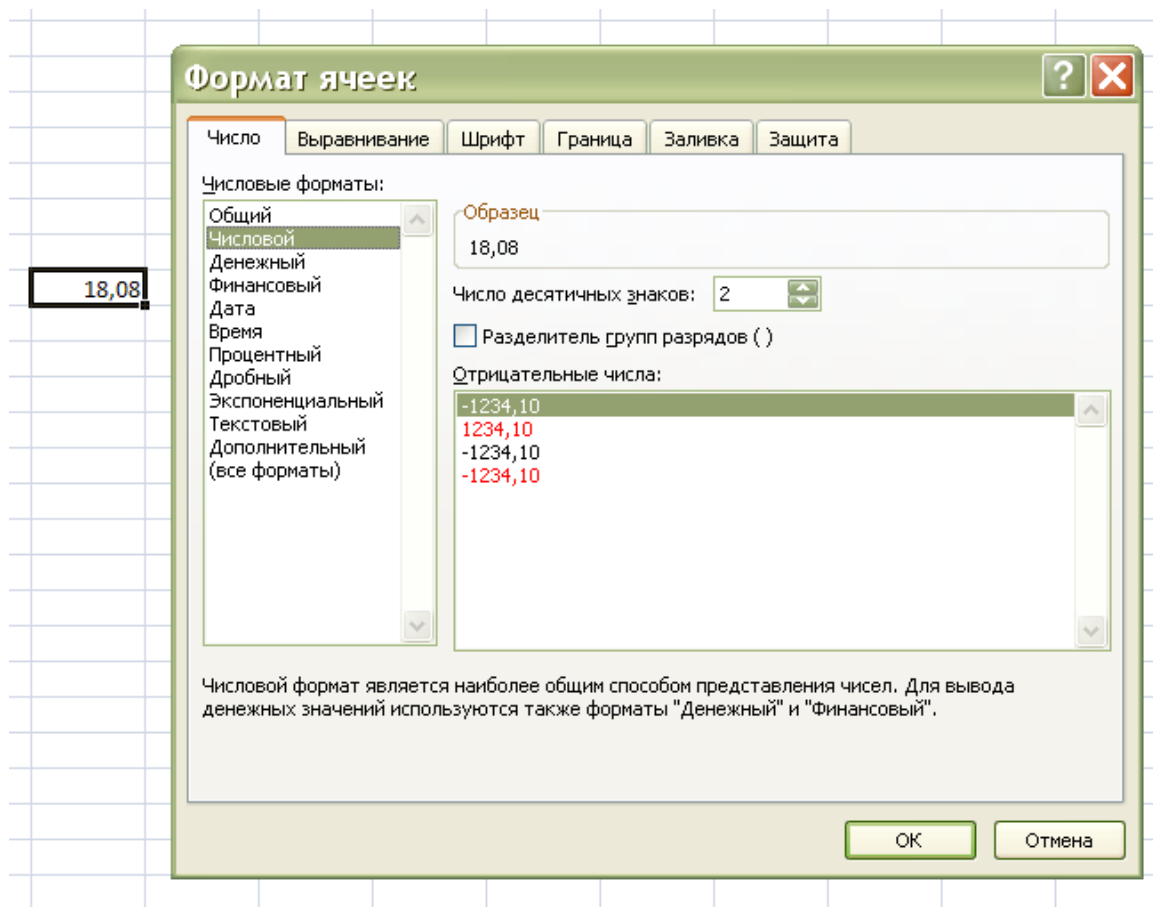
В окне формат ячеек вы можете изменять:

- Числовой формат (чтобы данные воспринимались как число, дата, почтовый индекс и т.д.)
- Положение данных внутри ячейки (выравнивание)
- Шрифт (в т.ч. вводить над/подстрочные символы)
- Границу ячеек (по умолчанию ячейка не имеет границ)
- Цвет заливки
- Защищенность ячейки

Формат ячеек в Excel

Menu – Format – Cells

Меню- Формат - Ячейка



Right click – Формат ячеек

2. Встроенные функции

Математические

- **abs**
- **макс**
- **sin**
- **мин**
- **exp**
- **целое**
- **остат**
- **частное**
- **Log**
- **корень**

Текстовые

- **сжпробелы**
- **левсимв**
- **правсимв**
- **пстр**
- **сцепить**
- **считатьпустоты**
- **длстр**
- **найти**
- **поиск**

Статистические

- **срзнач**
- **стандотклон**
- **медиана**
- **случмежду**

Ссылки

- **впр**
- **гпр**
- **строка**
- **столбец**
- **индекс**

Проверка значений

- **счётесли**
- **счётеслимн**
- **суммесли**
- **суммеслимн**

Логические

- **если**
- **И**
- **ИЛИ**
- **НЕ**
- **еслиошибка**
- **ИСТИНА**
- **ЛОЖЬ**

... и много других

Формула

- Ввод – через “=”
- Встроенные функции
- Мастер формул

	A	B	C	D	E
1	gene	CDS_length	Protein_length		
2	A0007	303	100		
3	A0008	261			
4	A0009	341			
5	A0010	600			

Formula bar: C2 = (B2-3)/3

	A	B	C	D
1	gene	CDS_length	Protein_length	
2	A0007	303	100	
3	A0008	261	86	=
4	A0009	341	112.67	112
5	A0010	600	199	

Formula bar: D3 =

Мастер формул: Категория: Математические. Выберите функцию: ЧАСТНОЕ. ЧАСТНОЕ(числитель, знаменатель). Возвращает целую часть резу...

	A	B	C	D	E	F
1	gene	CDS_length	Protein_length			
2	A0007	303	100			
3	A0008	261	86			
4	A0009	341	112.67	112		
5	A0010	600	199			

Formula bar: D4 = ЦЕЛОЕ((B4-3)/3)

3. Распространение формулы.

Демонстрация

- Задача: вычислить интеграл от 0 до 1 от e^x

Доллар запрещает изменение при распространении

Выделение прямоугольной области

Распространение вниз

(Ctrl+R – распространение вправо)

	A	B	C	D	E
1	x	e^x	интеграл	step =	0.01
2	0	1			
3	0.01				

	A	B	C	D	E	F
1	x	e^x	интеграл	step =	0.01	
2	0	1				
3	0.01					
4	0.02					

	A	B	C	D	E	F
3	0.01					
4						
5						
6						
7						

... продолжение

B2		fx		=EXP(A2)		
	A	B	C	D	E	F
1	x	e^x	интеграл	step =	0.01	
2	0	1				
3	0.01					
4	0.02					
5	0.03					

Формула для
интеграла от 0
до x

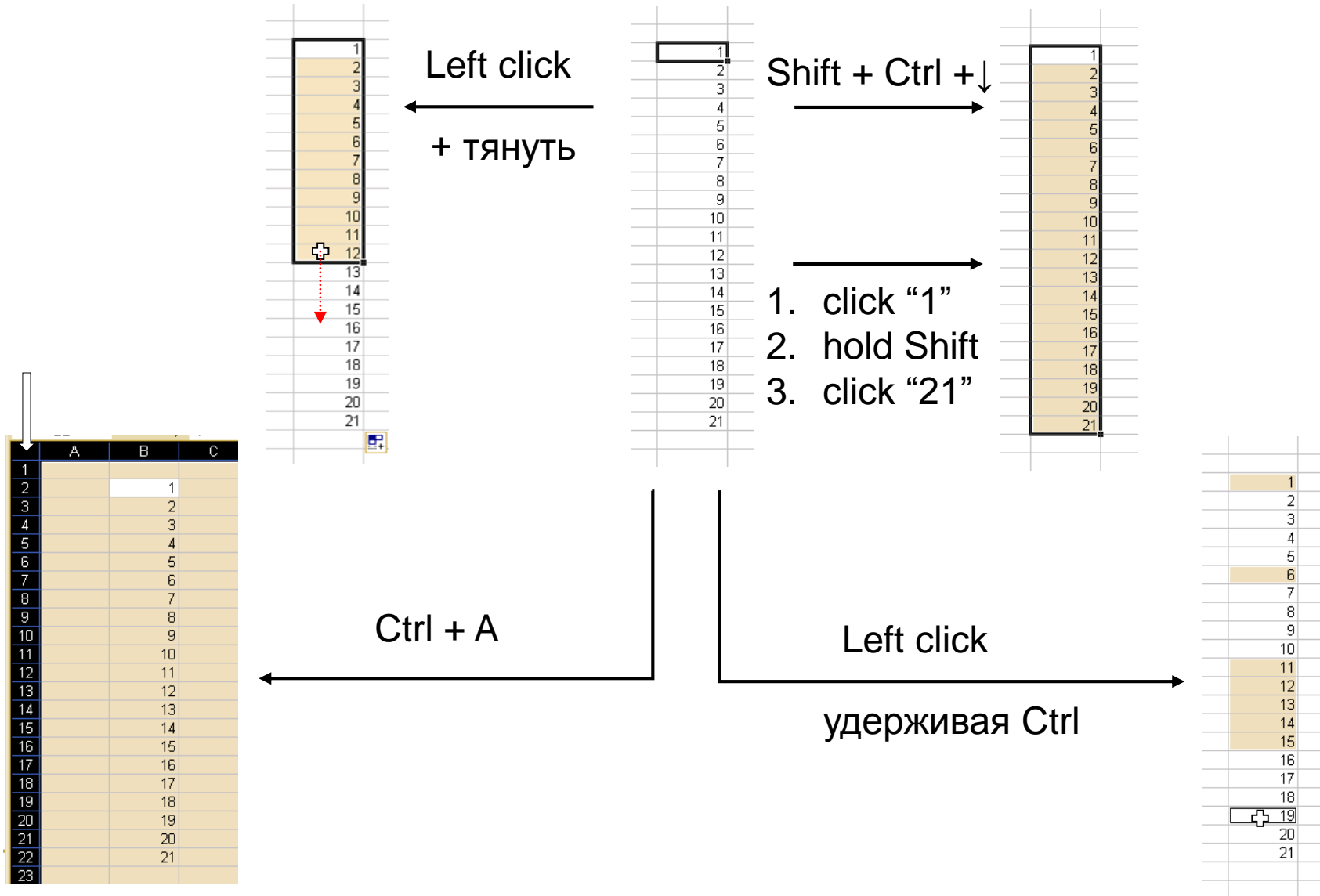
Двойной щелчок приводит к распространению до последней заполненной ячейки в соседнем столбце

C5		fx		=СУММ(B\$2:B5)*E\$1		
	A	B	C	D	E	F
1	x	e^x	интеграл (0,x)	step =	0.01	
2	0	1	0.01			
3	0.01	1.0101	0.020100502			
4	0.02	1.0202	0.030302515			
5	0.03	1.0305	0.04060706			
6	0.04	1.0408	0.051015168			

Работая с большими объемами данных, можно **выделять** их несколькими способами:

- Выделить *весь лист*: нажать на левый верхний угол листа или Ctrl+A
- Выделить *весь столбец/строку*, нажав на имя столбца/строки
- Выделить *часть столбца* или строки:
 - Щелкнуть левой кнопкой мыши в середине ячейки и потянуть до конца диапазона
или
 - Выделить первую ячейку и нажать Shift+Ctrl+↓ (или нужную стрелочку) для выделения до конца непустых ячеек в этом направлении
или
 - Выделить первую ячейку и, удерживая Shift, выделить последнюю нужную ячейку связанного диапазона
- Выделить несвязанный диапазон: удерживая Ctrl, щелкать левой клавишей мышки на нужные ячейки

Выделяем...



Фиксировать можно и столбцы

- Знак \$:

K\$29:L30 или **K\$29:\$L30** или **\$K\$29:\$L\$30**

Все комбинации выделяют ту же область,
но ...

При распространении и **копировании**
формулы модифицируются по разному

4. Специальная вставка

Демонстрация

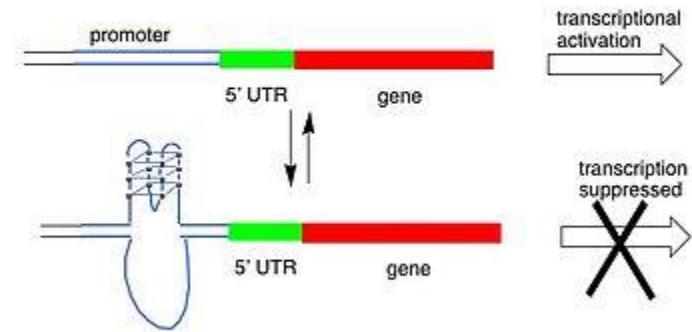
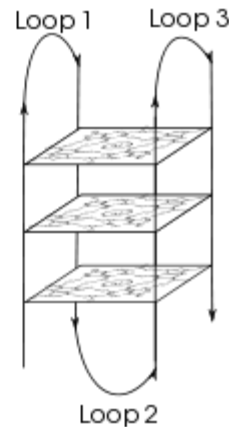
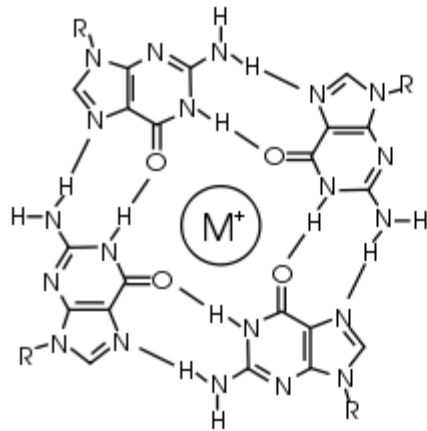
- Ведомость. Вставка pr12

5. Гистограмма

Демонстрация

- Длины белков семейства Vse634I
- Использование фильтра

G-квадруплекс – особая форма ДНК

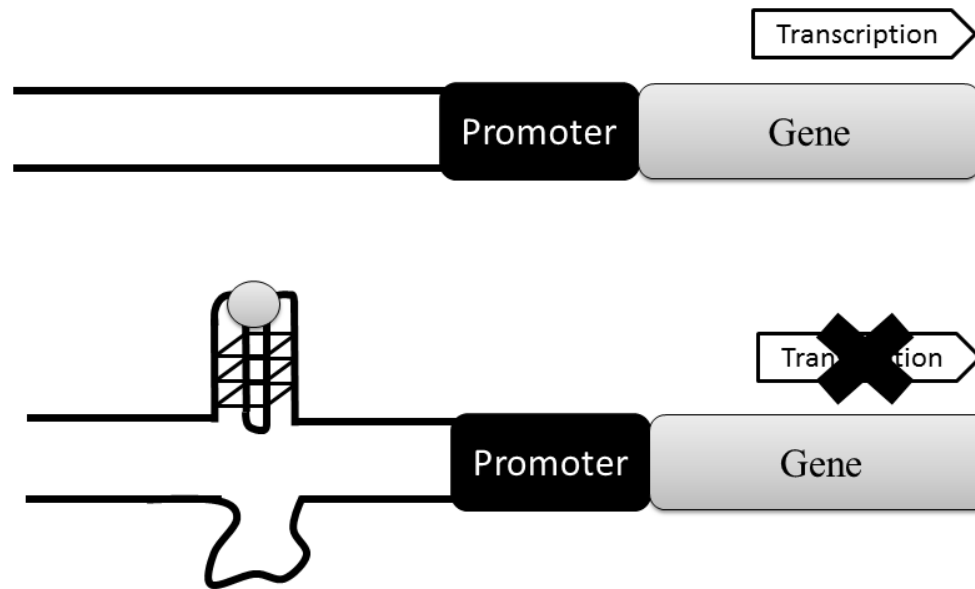


Этажерка – это одноцепочечная ДНК
Слева – вид сверху на одну «полочку»

Известны примеры регуляции трансляции с помощью G-квадруплетов

Yang D, Hurley LH. Structure of the biologically relevant G-quadruplex in the c-MYC promoter.

Nucleosides Nucleotides Nucleic Acids. 2006;25(8):951-68



Биоинформатическое предсказание регуляторных G-квадруплексов

Гистограмма расстояния от вероятных G-квадруплексов до стартов транскрипции генов человека



Вероятный G-квадруплекс: GGGN(1,7)GGGN(1,7)GGGN(1,7)GGG

Объяснение

- G-квадруплекс – особая форма одноцепочечной ДНК или РНК, см. картинки в google
- Необходимым условием образования G-квадруплекса является последовательность из четырех троек гуанинов, разделенных несколькими нуклеотидами.
(Иногда достаточно G-квадруплекс образуется при четырех парах гуанинов)
- Это условие не является достаточным! G-квадруплексная последовательность может складываться в G-квадруплекс только при определенных условиях. Например, в случае ДНК, комплементарная цепь должна быть как-то защищена. Образование G-квадруплекса зависит от ионов и лигандов – регулируется.
- Показано, что образование или необразование G-квадруплекса из G-квадруплексной последовательности регулирует экспрессию некоторых генов.

Есть или может быть подключена команда построения гистограммы

- Меню Данные => Анализ данных => гистограмма
- Если нет Анализа данных, то подключить его так:
 - Меню Большой круг – где сохранение => Параметры => Надстройки => Надстройки Excel Перейти => галочку возле Пакет анализа (не путать с Пакет анализа VBA!)
- Не советую использовать, так как при любом изменении карманов надо все делать заново

6. Какие ответы дает МатСтатистика

- Пять орлов при 20-ти бросаниях: случайность или монета кривая?
- Число “успехов” при многих бросаниях имеет распределение Бернулли, оно же – биномиальное
- Найдем вероятность **P** того, что при двадцати *независимых* бросаниях с вероятностью успеха 0.5 выпадет 5 или менее успехов
- В Excel: fx => статистические => биномраспр:
 - число успехов = 5
 - число испытаний = 20
 - вероятность успеха = 0.5
 - интегральная = ИСТИНА (т.к. ≤ 5 успехов)Ответ: **P = 0.02**
- Интерпретация: вероятность получить такой исход 20-ти бросаний при правильной монете равна 0.02. Очевидно, еще такая же вероятность получить 15 или более решек
- 1000 генов на прямой цепи ДНК из 2200и генов всего:
 - случайность или
 - имеет биологический смысл?

7. Импорт с приключениями

Задание на день занятия

- Импортировать два данных файла в ЭТ

Импорт данных из текста

Microsoft Excel - Crystal

File Edit View Insert Format Tools Data Window Help

Type a question for help

K802 29.06.2004

	A	B	C	D	E	F	G	H	I	J	K	L	M
781	ATOM	781	C	LYS	B	177	-11.358	14.982	2.784	1.00	35.38	C	
782	ATOM	782	O	LYS	B	177	-10.387	15.202	3.509	1.00	33.89	O	
783	ATOM	783	CB	LYS	B	177	-13.415	16.431	2.836	1.00	36.37	C	
784	ATOM	784	CG	LYS	B	177	-13.993	17.791	2.461	1.00	37.76	C	
785	ATOM	785	CD	LYS	B	177	-15.468	17.927	2.812	1.00	42.33	C	
786	ATOM	786	CE	LYS	B	177	-15.717	17.914	4.309	1.00	44.59	C	
787	ATOM	787	NZ	LYS	B	177	-17.165	18.102	4.605	1.00	47.49	N	
788	ATOM	788	N	ASN	B	178	-11.807	13.757	2.533	1.00	36.51	N	
789	ATOM	789	CA	ASN	B	178	-11.162	12.587	3.110	1.00	36.13	C	
790	ATOM	790	C	ASN	B	178	-9.721	12.485	2.633	1.00	34.95	C	
791	ATOM	791	O	ASN	B	178	-8.830	12.129	3.402	1.00	32.72	O	
792	ATOM	792	CB	ASN	B	178	-11.934	11.319	2.739	1.00	38.55	C	
793	ATOM	793	CG	ASN	B	178	-12.971	10.952	3.779	1.00	40.52	C	
794	ATOM	794	OD1	ASN	B	178	-14.111	10.623	3.453	1.00	43.43	O	
795	ATOM	795	ND2	ASN	B	178	-12.573	10.999	5.044	1.00	43.93	N	
796	ATOM	796	N	TRP	B	179	-9.487	12.809	1.366	1.00	32.96	N	
797	ATOM	797	CA	TRP	B	179	-8.134	12.741	0.839	1.00	30.78	C	
798	ATOM	798	C	TRP	B	179	-7.227	13.733	1.567	1.00	33.00	C	
799	ATOM	799	O	TRP	B	179	-6.110	13.390	1.949	1.00	35.23	O	
800	ATOM	800	CB	TRP	B	179	-8.113	13.040	-0.663	1.00	29.30	C	
801	ATOM	801	CG	TRP	B	179	-6.775	12.755	-1.280	1.00	29.83	C	
802	ATOM	802	CD1	TRP	B	179	-6.337	11.560	-1.773	1.00	29.июн	C	
803	ATOM	803	CD2	TRP	B	179	-5.672	13.665	-1.390	1.00	27.85	C	
804	ATOM	804	NE1	TRP	B	179	-5.028	11.667	-2.182	1.00	28.59	N	
805	ATOM	805	CE2	TRP	B	179	-4.599	12.947	-1.962	1.00	26.55	C	
806	ATOM	806	CE3	TRP	B	179	-5.492	15.014	-1.066	1.00	27.59	C	
807	ATOM	807	CZ2	TRP	B	179	-3.351	13.536	-2.205	1.00	27.83	C	
808	ATOM	808	CZ3	TRP	B	179	-4.252	15.600	-1.310	1.00	29.64	C	
809	ATOM	809	CH2	TRP	B	179	-3.200	14.859	-1.877	1.00	27.58	C	
810	ATOM	810	N	VAL	B	180	-7.709	14.956	1.774	1.00	32.28	N	
811	ATOM	811	CA	VAL	B	180	-6.908	15.980	2.442	1.00	30.59	C	

Ready NUM SCRL

Вот что получается при небрежном обращении с Excel. 2004

BMC Bioinformatics

 BioMed Central

Correspondence

Open Access

Mistaken Identifiers: Gene name errors can be introduced inadvertently when using Excel in bioinformatics

Barry R Zeeberg^{†1}, Joseph Riss^{†2}, David W Kane³, Kimberly J Bussey¹, Edward Uchio⁴, W Marston Linehan⁴, J Carl Barrett² and John N Weinstein^{*1}

Abstract

Background: When processing microarray data sets, we recently noticed that some gene names were being changed inadvertently to non-gene names.

Results: A little detective work traced the problem to default date format conversions and floating-point format conversions in the very useful Excel program package. The date conversions affect at least 30 gene names; the floating-point conversions affect at least 2,000 if Riken identifiers are included. These conversions are irreversible; the original gene names cannot be recovered.

Conclusions: Users of Excel for analyses involving gene names should be aware of this problem, which can cause genes, including medically important ones, to be lost from view and which has contaminated even carefully curated public databases. We provide work-arounds and scripts for circumventing the problem.

Проблема не решена. 2016

COMMENT

Open Access



Gene name errors are widespread in the scientific literature

Mark Ziemann¹, Yotam Eren^{1,2} and Assam El-Osta^{1,3*}

Abstract

The spreadsheet software Microsoft Excel, when used with default settings, is known to convert gene names to dates and floating-point numbers. A programmatic scan of leading genomics journals reveals that approximately one-fifth of papers with supplementary Excel gene lists contain erroneous gene name conversions.

Keywords: Microsoft Excel, Gene symbol, Supplementary data

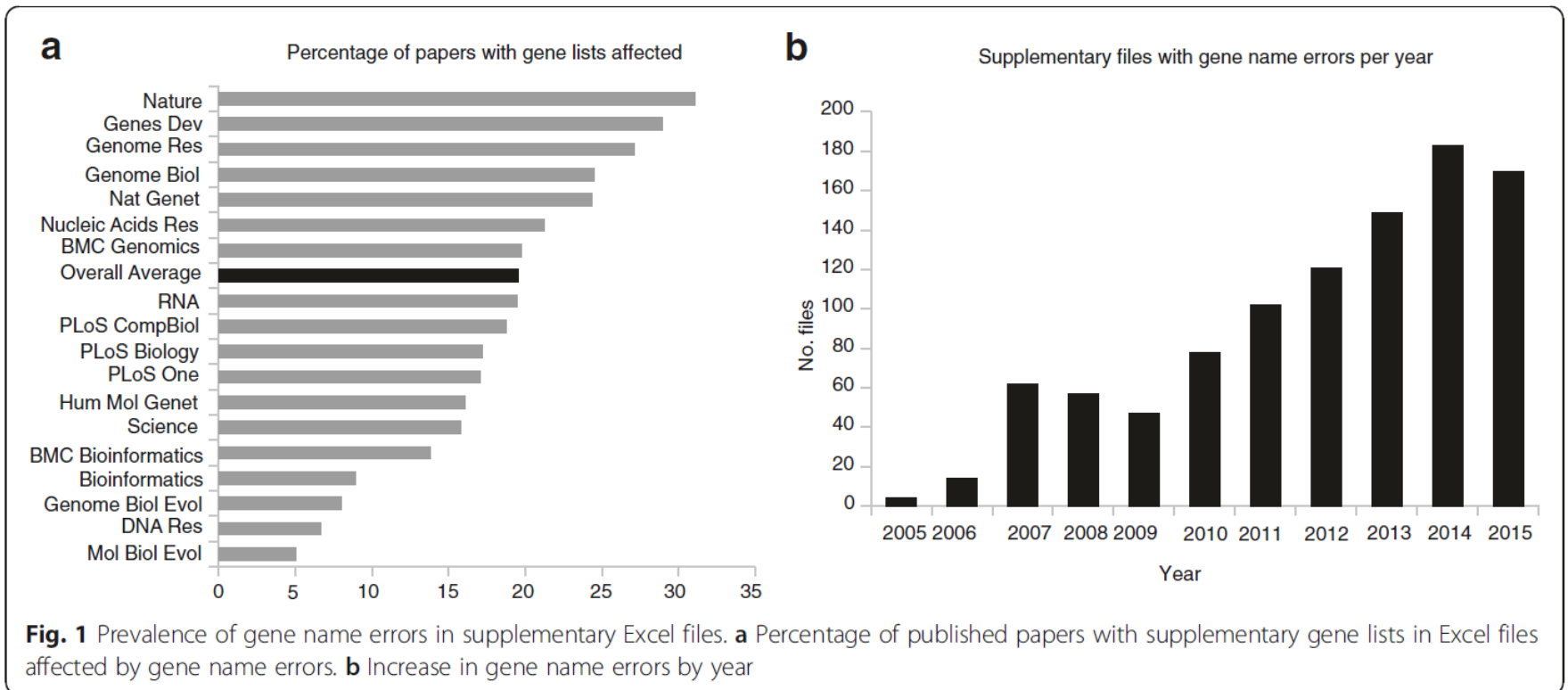
Abbreviations: GEO, Gene Expression Omnibus; JIF, journal impact factor

frequently reused. Our aim here is to raise awareness of the problem.

We downloaded and screened supplementary files from 18 journals published between 2005 and 2015 using a suite of shell scripts. Excel files (.xls and .xlsx suffixes) were converted to tabular separated files (tsv) with `ssconvert` (v1.12.9). Each sheet within the Excel file was converted to a separate tsv file. Each column of data in the tsv file was screened for the presence of gene symbols. If the first 20 rows of a column contained five or more gene symbols, then it was suspected to be a list of gene symbols, and then a regular expression (regex) search of the entire column was applied to identify gene symbol errors. Official gene symbols from Ensembl version 82, accessed November 2015, were obtained for *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Danio rerio*, *Escherichia coli*, *Callis*

The problem of Excel software (Microsoft Corp., Redmond,

Самые “плохие” журналы



И В 2019 ...

«For instance, **gene symbols** and accession numbers **may** inadvertently **be converted into dates or floating point numbers** when -omics researchers copy and paste their results into Microsoft Excel spreadsheets **without tedious error correction by hand**. This is a rather common error, but it is noteworthy that the incidence of **such errors is higher in more prestigious journals** [[46](#)].»

Björn Brembs* “Reliable novelty: New should not trump true”
PLoS Biol. 2019 Feb; 17(2): e3000117.

Удобности

- Выделение области, напр. A5:B120000 (окошко рядом с fx)
- Выделение набора столбцов, набора строк, всего листа (щелчок по именам столбцов, номерам строк, удерживая Ctrl или Shift, если много)
- Переход
 - к заданной ячейке, напр. F10259
 - к краю заполненного диапазона (Ctrl+стрелка)
- Автозаполнение арифметической прогрессии (тянуть за уголок).
- Суммирование непрерывного диапазона чисел (значок Автосуммирование или Alt+=)
- Поиск/замена символа/слова на другой/другое везде, где он встретился в выделенном диапазоне (Ctrl+F/Ctrl+H).
- Отмена последнего действия (Ctrl+Z)
- Введение и редактирование формулы (через fx , в ячейку или окошко формул начиная с “=“)
- Вставка
 - значения в ячейку без изменения формата ячейки (**в окошко формул или F2**)
 - **специальная вставка** (через правую кнопку мыши)
- Изменение числа отображаемых десятичных знаков в числе (формат ячейки через правую кнопку мыши или меню “формат”)
- Распространение формулы
 - Ctrl+D, Ctrl+R
 - Мышью за правый нижний уголок
 - Ctrl+C – выделение диапазона – Ctrl+V
- Вставка пустого столбца/строки (выделить и нажать Вставить в Главном меню)

... и не очень удобства

- Как ввести символы +, −, = как текстовые значения? (написать '+ и т.п.)
- Как бороться с тем, что число вида 10.12 превращается при вводе в “10 декабря” или что-то вроде того? (никак; надо заменить . на , и ввести заново) или в мастере ввода выбрать “текст” для колонки
- Как задать условие в команде countif [счётеесли]? (например так:
=countif(A:A,">="&B1) – считай значения в колонке A большие или равные значению ячейки B1; выражение ">="&B1 есть короткая запись команды concatenate(">=",B1))

8. Задания выполняемые в день занятия

- Создать ЭТ для этого задания
- Скопировать и отформатировать страницу размеров геномов самостоятельно
- Импортировать корректно два данных текстовых файла в ЭТ
- Записаться в очередь (практикум 12.1)

Результаты – на отдельных страницах!

КОНЕЦ

Пока остановимся ...

Плоская таблица

Genomes	superkingdom	no rank	phylum	class
CP001787.1	Archaea		Euryarchaeota	Methanococci
CP002737.1	Archaea	Terrabacteria group	Euryarchaeota	Methanococci
L77117.1	Archaea		Euryarchaeota	Methanococci
CP013050.1	Archaea		Euryarchaeota	Thermococci
BA000011.4	Archaea		Euryarchaeota	Thermoplasmata
DP000238.1	Archaea	TACK group	Thaumarchaeota	
CP003081.1	Bacteria	Bacteroidetes/Chlorobi group	Actinobacteria	Actinobacteria
CP011499.1	Bacteria	Terrabacteria group	Actinobacteria	Actinobacteria
CP001080.1	Bacteria		Aquificae	Aquificae

Требования

- Названия колонок – в первой строке. Если названия длинные, то удобно повернуть текст вертикально и разрешить перенос.
- Остальные строки – с данными. Все ячейки заполнены! Пустая ячейка может быть только если это – предусмотренное значение; например, обозначает «данные неизвестны»
- В первой колонке – уникальный идентификатор строки
- **НЕТ ОБЪЕДИНЕНИЙ ЯЧЕЕК!**

Не плоская таблица. Пример. Как надо было ее составить

Microsoft Excel - Книга1

Файл Правка Вид Вставка Формат Сервис Данные Окно ?

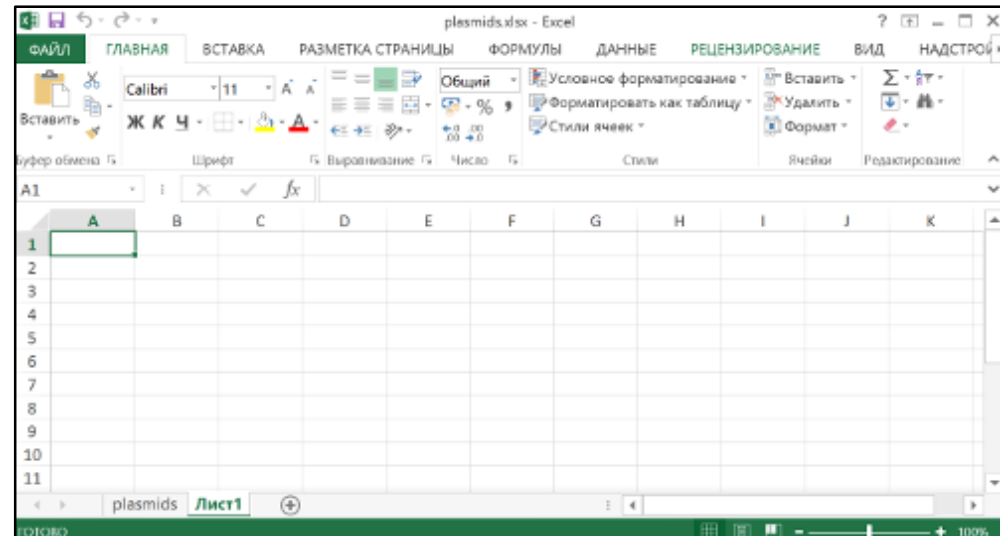
	A	B	C	D
1	Перечень техники, приобретенной кафедрами энергетического факультета ИрГТУ в 2000 г.			
2	Наименование	Кол-во	Цена	Сумма
3	Кафедра теплоэнергетики			
4	Компьютеры	11	17800	195800
5	Сканеры	1	3400	3400
6	Звуковые колонки	2	200	400
7	Итого:			199600
8	Кафедра электрических станций, сетей и систем			
9	Компьютеры	10	17800	178000
10	Принтеры	1	5700	5700
11	Сканеры	1	3400	3400
12	Итого:			187100
13	Кафедра электропривода и автоматизации			
14	Компьютеры	4	18700	74800
15	Принтеры	1	6900	6900
16	Итого:			81700
17				
18	Всего:			468400
19				

Исходная таблица Лист2 Лист3

Готово NUM

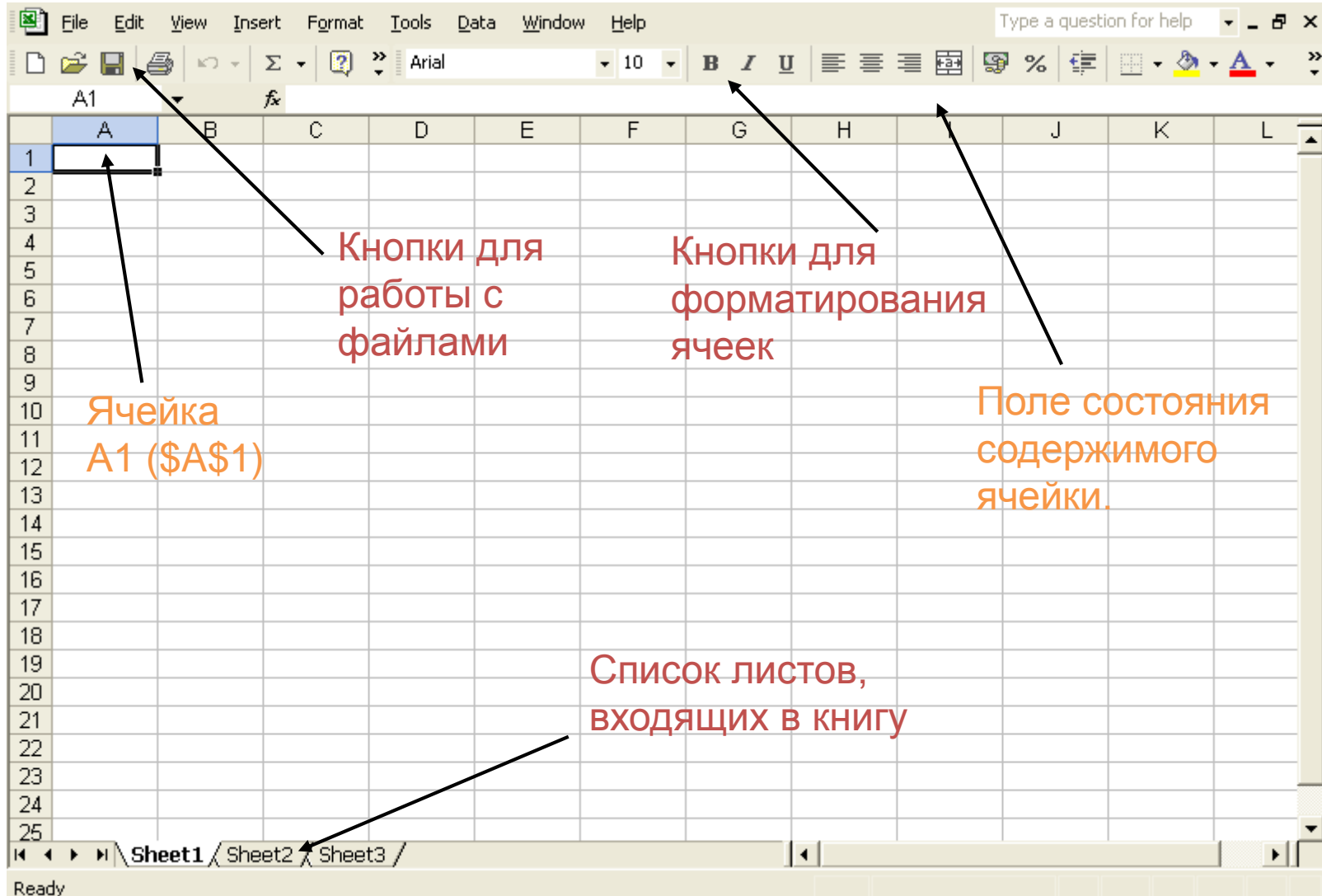
Структура файлов

- Файл Excel имеет расширение .xlsx или .xls
- Файл состоит из листов
- Лист – таблица; состоит из ячеек
- Возможно использование данных из одного .xlsx файла в другом .xlsx файле



Общий вид листа

Skip



Ячейка

Skip

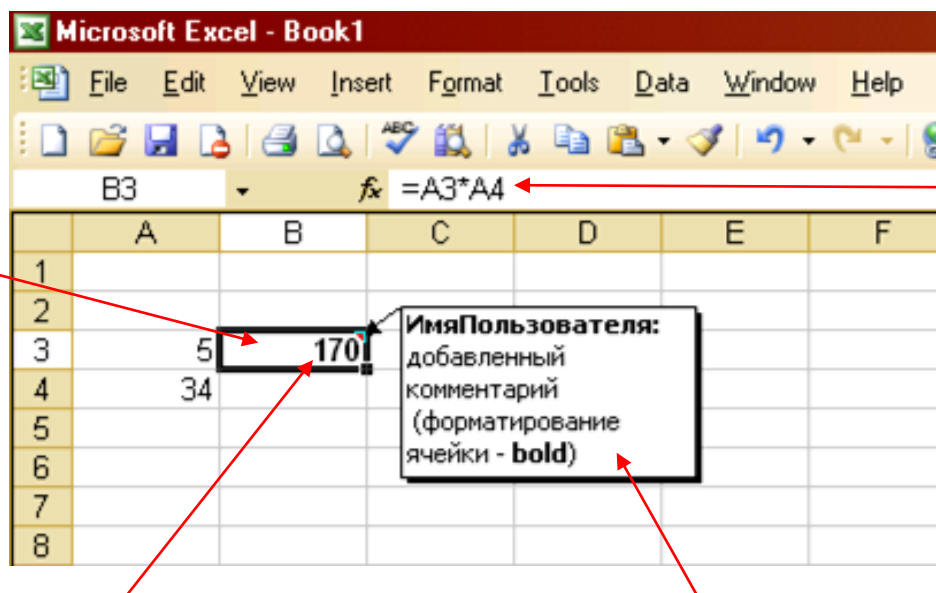
- Четыре атрибута ячейки:

- значение,
- формула,
- форматирование,
- комментарий

- Адрес:

- B2 (в том же листе)
- TestSheet!B2 (в том же файле)
- [TestFile.xlsx]TestSheet!B2 (в другом файле)

Значение:
170



Формула:
=A3*A4

Форматирование:

arial 10, bold,
правое выравнивание,
общий формат числа и т.д

Комментарий

Совместимость таблиц. Функция ВПР

В подсказках есть ссылка на объяснение этой функции

ВПР (VLOOKUP)

Microsoft Excel - Book1

File Edit View Insert Format Tools Data Window Help

D4 fx

	A	B	C
1	Аминокислота, иминокислота пролин	Масса amino-(имино-) кислоты	
2	A	89,09	
3	C	121,16	
4	D	133,1	
5	E	147,13	
6	F	165,19	
7	G	75,07	
8	H	155,1	
9	I	131,17	
10	K	146,19	
11	L	131,17	
12	M	149,21	
13	N	132,12	
14	P	115,13	
15	Q	146,15	
16	R	174,2	
17	S	105,09	
18	T	119,12	
19	V	117,15	
20	W	204,22	
21	Y	181,19	
22			
23			
24			
25			

1

Microsoft Excel - Book1

File Edit View Insert Format Tools Data Window Help

C3 fx

	A	B	C	D
1				
2		А/к последовательность	молекулярная масса а/к остатка	
3		L		
4		V		
5		A		
6		L		
7		V		
8		R		
9		A		
10		S		
11		I		
12		T		
13		D		
14		P		
15		G		
16		R		
17		L		
18		P		
19		E		
20		N		
21		P		
22		K		
23		I		
24		P		
25		H		
26		G		
27		E		
28				

2

ВПР (VLOOKUP)

Искомое значение

Искомое значение (lookup_value)

адрес интервала ячеек,
содержащих идентификатор (в
данном случае, код остатка),
который нужно найти в первой
колонке 1й таблицы

*Ссылка указывает название листа и
интервал данных

The screenshot shows an Excel spreadsheet with a VLOOKUP formula in cell C3. The formula bar displays `=VLOOKUP(Sheet1!B3:B27)`. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I
1									
2		А/к последова тельность	молекулярная масса а/к остатка						
3		L	=VLOOKUP(Sheet1!B3:B27)						
4		V							
5		A							
6		L							
7		V							
8		R							
9		A							
10		S							
11		I							
12		T							
13		D							
14		P							
15		G							
16		R							
17		L							
18		P							
19		E							
20		N							
21		P							
22		K							
23		I							
24		P							
25		H							
26		G							
27		E							

The 'Function Arguments' dialog box for the VLOOKUP function is open, showing the following fields:

- Lookup_value:** Sheet1!B3:B27
- Table_array:** (empty)
- Col_index_num:** (empty)
- Range_lookup:** (empty)

The dialog box also includes a description of the function and buttons for 'OK' and 'Cancel'.

ВПР (VLOOKUP) таблица

таблица (table array)

адрес таблицы (1), в первой колонке которой следует искать заданный идентификатор (а/к)

номер столбца (Col index num)
из табл. 1, который содержит необходимые значения для табл. 2 (второй столбец)

интервальный просмотр (range lookup)

указать ЛОЖЬ, чтобы всё получилось ;)

VLOOKUP X ✓ fx =VLOOKUP(Sheet1!B3:B27;aa_mass!A2:B21;2;FALSE)

	A	B	C	D	E	F	G
	Аминокислота, иминокислота	Масса амино-(имино-) кислоты					
1	пролин						
2	A	89,09					
3	C	121,16					
4	D	133,1					
5	E	147,13					
6	F	165,19					
7	G	75,07					
8	H	155,1					
9	I	131,17					
10	K	146,19					
11	L	131,17					
12	M	149,21					
13	N	132,12					
14	P	115,					
15	Q	146,					
16	R	174,					
17	S	105,					
18	T	119,					
19	V	117,					
20	W	204,					
21	Y	181,					
22							
23							
24							
25							
26							
27							
28							
29							
30							

Function Arguments

VLOOKUP

Look_up_value Sheet1!B3:B27 = "L"

Table_array aa_mass!A2:B21 = {"A";89,09;"C";121,16;"D";133,1;"E";147,13;"F";165,19;"G";75,07;"H";155,1;"I";131,17;"K";146,19;"L";131,17;"M";149,21;"N";132,12;"P";115,;"Q";146,;"R";174,;"S";105,;"T";119,;"V";117,;"W";204,;"Y";181,}

Col_index_num 2 = 2

Range_lookup FALSE = FALSE

= 131,17

Looks for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify. By default, the table must be sorted in an ascending order.

Table_array is a table of text, numbers, or logical values, in which data is retrieved. Table_array can be a reference to a range or a range name.

Formula result = 131,17

[Help on this function](#) OK Cancel

ВПР (VLOOKUP)

результат

распространение

fx =VLOOKUP(Sheet1!B3:B27;aa_mass!A2:B21;2;FALSE)

A/к последовательность	молекулярная масса а/к остатка
L	131,17
V	
A	
L	
V	
R	
A	
S	
I	
T	
D	
P	
G	
R	
L	
P	
E	
N	
P	
K	
I	
P	
H	
G	
E	

fx =VLOOKUP(Sheet1!\$B\$3:\$B\$27;aa_mass!\$A\$2:\$B\$21;2;FALSE)

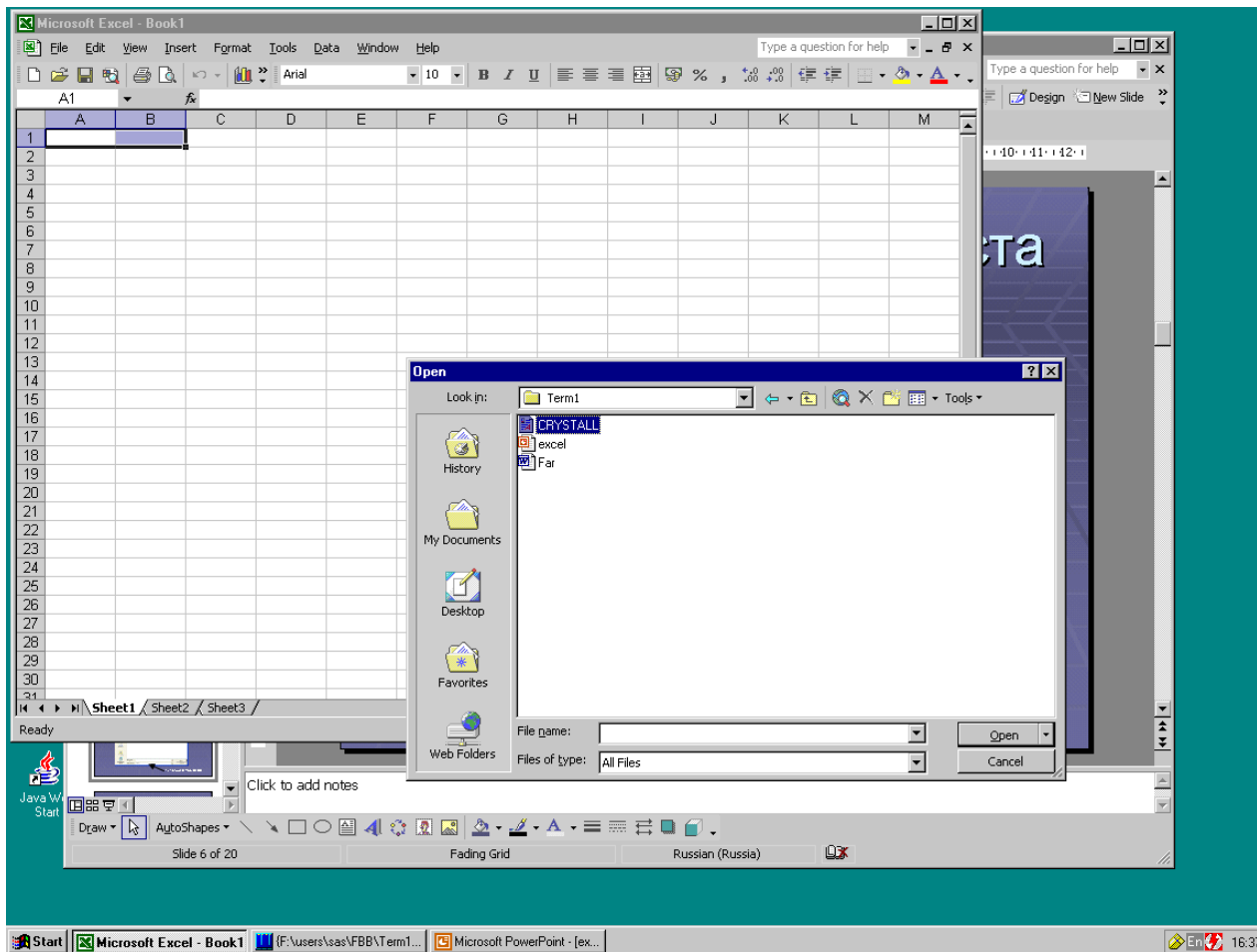
A/к последовательность	молекулярная масса а/к остатка
L	131,17
V	117,15
A	89,09
L	131,17
V	117,15
R	174,2
A	89,09
S	
I	
T	
D	
P	
G	
R	
L	
P	
E	
N	
P	
K	
I	
P	
H	
G	
E	

\$ закрепить области!!!!

Ввод данных

- С помощью мастера импорта
 - Из текстового файла
 - Путем копирования
- Путем автозаполнения
 - Копирование вниз (Ctrl-D) и вправо (Ctrl-R)
 - Арифметическая прогрессия
- Путем вычисления значений. Формулы.
- Разбить тексты в колонке по столбцам

Импорт данных из текста



Импорт данных из текста

Тип текстовой таблицы

Кодировка

Text Import Wizard - Step 1 of 3

The Text Wizard has determined that your data is Fixed Width.
If this is correct, choose Next, or choose the data type that best describes your data.

Original data type
Choose the file type that best describes your data:

- Delimited - Characters such as commas or tabs separate each field.
- Fixed width - Fields are aligned in columns with spaces between each field.

Start import at row: 25 File origin: 1254 : Turkish (Windows)

Preview of file F:\users\sas\FBB\Term1\CRYSTALL.pdb.

23	SCALE2	0.000000	0.021316	0.000000	0.000000						
24	SCALE3	0.000000	0.000000	0.006089	0.000000						
25	ATOM	1	N	LYS	A	74	18.732	8.253	-20.201	1.00	9
26	ATOM	2	CA	LYS	A	74	18.339	9.325	-19.243	1.00	9
27	ATOM	3	C	LYS	A	74	18.492	8.867	-17.798	1.00	9

Cancel < Back Next > Finish

Sheet1 Sheet2 Sheet3

Ready NUM SCRL

Импорт данных из текста

Разделители данных

Text Import Wizard - Step 2 of 3

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters

Tab Semicolon Comma
 Space Other:

Treat consecutive delimiters as one

Text qualifier: " "

Data preview

SCALE2	0.000000	0.021316	0.000000	0.000000				
SCALE3	0.000000	0.000000	0.006089	0.000000				
ATOM	1	N	LYS	A	74	18.732	8.253	-20.2
ATOM	2	CA	LYS	A	74	18.339	9.325	-19.2
ATOM	3	C	LYS	A	74	18.492	8.867	-17.7

Cancel < Back Next > Finish

Импорт данных из текста

Text Import Wizard - Step 3 of 3

This screen lets you select each column and set the Data Format.

'General' converts numeric values to numbers, date values to dates, and all remaining values to text.

Advanced...

Column data format:

- General
- Text
- Date: DMY
- Do not import column (skip)

Data preview

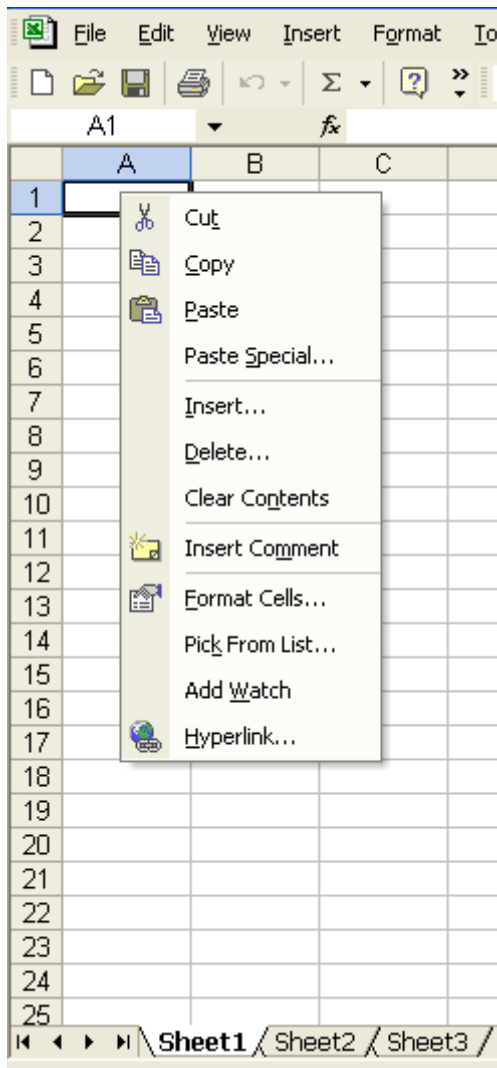
	Gener	Cener	Cener	Cener	Cener	Gener	Gener	General	Gener	Gener	Gener
ATOM 1	1	N	LYS	A	74	18.732	8.253	-20.201	1.00	92.76	N
ATOM 2	2	CA	LYS	A	74	18.339	9.325	-19.243	1.00	92.81	C
ATOM 3	3	C	LYS	A	74	18.492	8.867	-17.798	1.00	93.32	C
ATOM 4	4	O	LYS	A	74	18.009	7.798	-17.422	1.00	93.66	O
ATOM 5	5	CB	LYS	A	74	16.904	9.750	-19.500	1.00	90.24	C

Cancel < Back Next > Finish

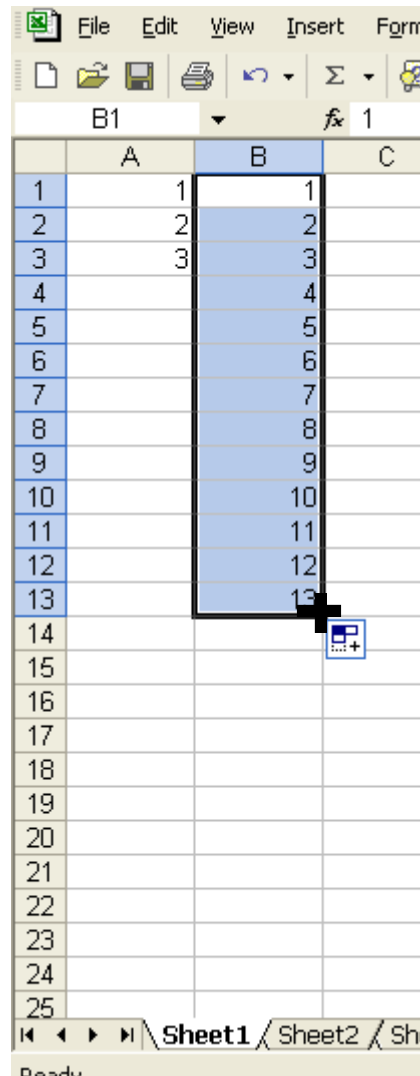
Определение
типа данных
в колонке

Ввод данных

Меню



Курсор

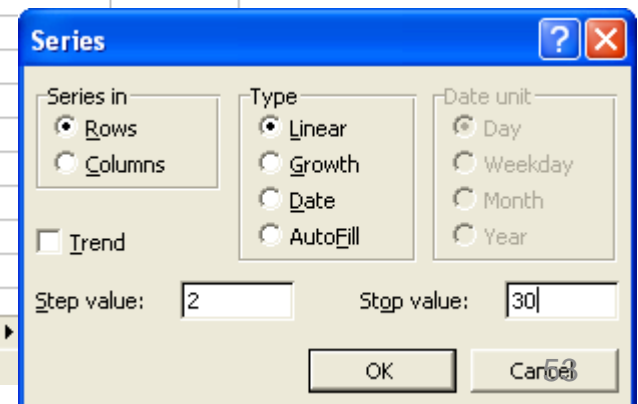


Прогрессия



Меню –
Правка –
Заполнить –
Прогрессия

Menu – Edit –
Fill – Series



Закрепление областей

	A	B	C	D	E	F	G	
1		#Integr8_Gene_ID	Primary_Gene_Name	Ordered_Locus_Name	Biological_Names	Start_position	Offset	Direction
2	IGI01006661	secA_1	CT_141	secA	158147	446	F	
3	IGI00976228	CT_610	CT_610		690426	695	R	
4	IGI00976954	nrdB	CT_828	nrdB	977556	1040	F	
5	IGI00972337	fusA	CT_437	fusA	505508	2084	R	
6	IGI00970720	map	CT_851	map	1000369	875	R	
7	IGI00972391	tuf	CT_322	tuf,tu	361980	1184	R	
8	IGI00972153	dnaK	CT_396	dnaK	451614	1982	F	
9	IGI00973083	gap	CT_505	gap,g	583066	1004	R	
10	IGI00977699	rpoD	CT_615	rpoD	694039	1715	R	
11	IGI00971455	clpB	CT_113	clpB	130617	2603	F	
12	IGI00977797	rpsL	CT_439	rpsL,t	508157	371	R	

	A	B	C	D	E	F	G	
1		#Integr8_Gene_ID	Primary_Gene_Name	Ordered_Locus_Name	Biological_Names	Start_position	Offset	Direction
2	IGI01006661	secA_1	CT_141	secA	158147	446	F	
3	IGI00976228	CT_610	CT_610		690426	695	R	
4	IGI00976954	nrdB	CT_828	nrdB	977556	1040	F	
5	IGI00972337	fusA	CT_437	fusA	505508	2084	R	
6	IGI00970720	map	CT_851	map	1000369	875	R	
7	IGI00972391	tuf	CT_322	tuf,tu	361980	1184	R	
8	IGI00972153	dnaK	CT_396	dnaK	451614	1982	F	
9	IGI00973083	gap	CT_505	gap,g	583066	1004	R	
10	IGI00977699	rpoD	CT_615	rpoD	694039	1715	R	
11	IGI00971455	clpB	CT_113	clpB	130617	2603	F	
12	IGI00977797	rpsL	CT_439	rpsL,t	508157	371	R	

	A	R	S	T	U	
1		#Integr8_Gene_ID	Description	EMBL_AC	NCBI_Gis	Inc_prediction
32	IGI00977983	30S ribosomal protein S4	AE00127	3328959	no	
33	IGI00977907	30S ribosomal protein S5	AE00127	3328961	no	
34	IGI00972472	Enolase (EC 4.2.1.11) (2)	AE00127	3329030	no	
35	IGI01006737	Pyruvate Dehydrogenase	AE00127	3328656	no	
36	IGI00974103	Translation initiation factor	AE00127	6578116	no	
37	IGI00978858	Valyl-tRNA synthetase	AE00127	3328717	no	
38	IGI01006633	GTP Binding Protein	AE00127	3328488	no	
39	IGI00977426	50S ribosomal protein L1	AE00127	#####	no	
40	IGI00972350	Elongation factor P 1 (E)	AE00127	3328521	no	
41	IGI00970586	Acyl carrier protein (AC)	AE00127	3328645	no	
42	IGI00978368	Superoxide dismutase	AE00127	3328709	no	

выделить границу
закрепления

Окно → Закрепить
области

области закреплены
и видны постоянно
при прокрутке листа

Числовая функция

The screenshot displays the Microsoft Excel interface. The main window shows a spreadsheet with the following data in column B:

Row	Column B
1	
2	7
3	8
4	5
5	3
6	4
7	9
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	

The formula bar at the top shows the formula $=A2+1$ and the active cell B2 contains the value 8. The task pane on the left lists several options, with 'Числовая формула' (Number formula) selected. A file dialog box is open in the background, showing a file name field and 'Open' and 'Cancel' buttons. The Windows taskbar at the bottom shows the Start button and several open applications, including Microsoft Excel, Internet Explorer, and Microsoft PowerPoint.

Текстовая функция

The screenshot shows the Microsoft Excel interface. The title bar reads "Microsoft Excel - Book1". The menu bar includes "File", "Edit", "View", "Insert", "Format", "Tools", "Data", "Window", and "Help". The toolbar contains various icons for file operations and formatting. The active cell is C3, and the formula bar shows the formula `=CONCATENATE(A3;B3)`. The spreadsheet grid shows the following data:

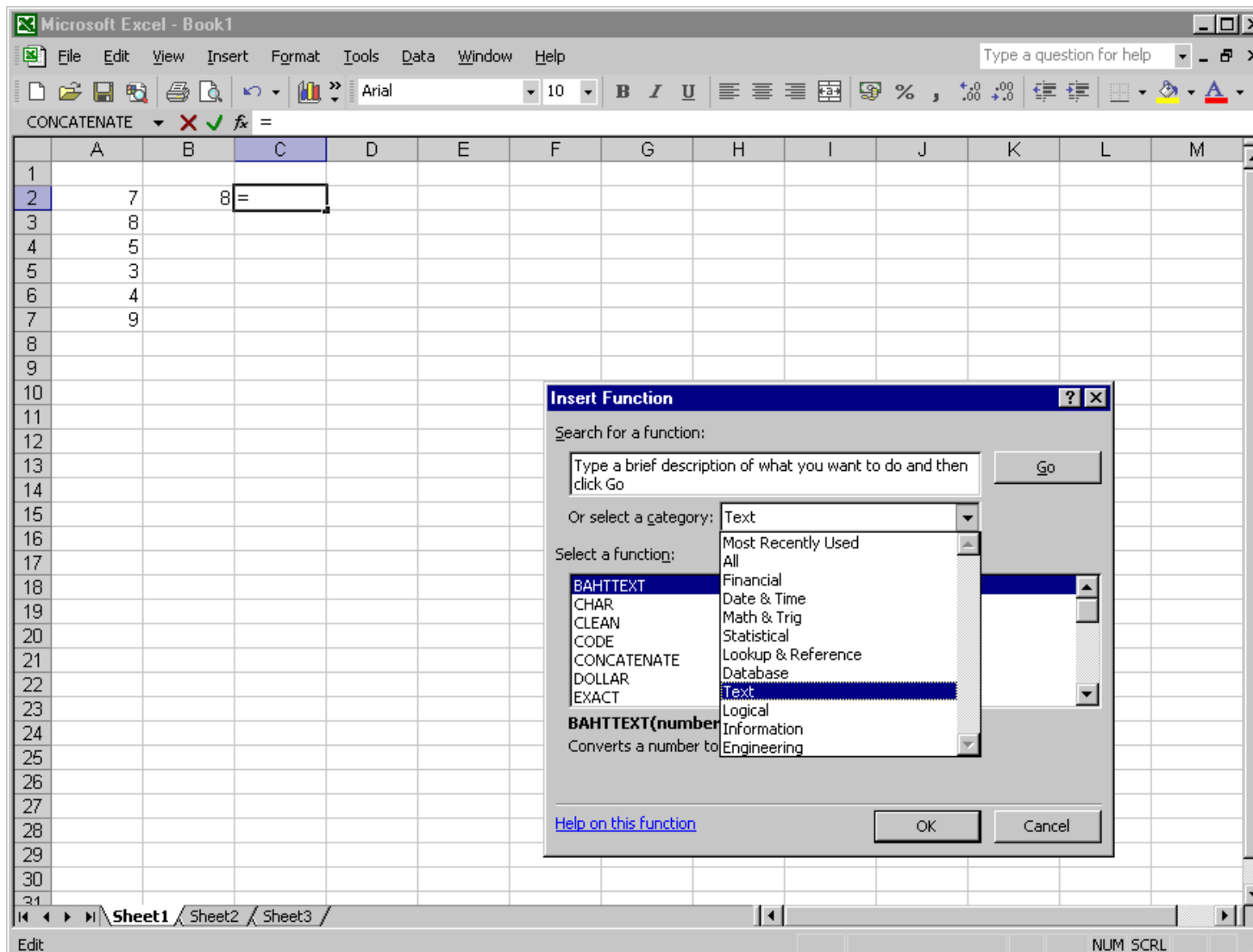
	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2													
3	КОНЬ	ЯК	КОНЬЯК										
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													

The status bar at the bottom shows "Ready" and "NUM SCRL". The sheet tabs at the bottom indicate "Sheet1", "Sheet2", and "Sheet3".

Мастер функций

Menu – Insert – Function

Меню – Вставка – Функция



Мастер функций

The screenshot shows the Microsoft Excel interface with the Function Arguments dialog box open for the ROMAN function. The dialog box is titled "Function Arguments" and contains the following information:

- Function name: ROMAN
- Number: B2 (with a lightning bolt icon indicating a warning or error)
- Form: (empty)
- Result: = 8 (for the Number field) and = number (for the Form field)
- Result: = "VIII"
- Description: Converts an Arabic numeral to Roman, as text.
- Parameter description: **Number** is the Arabic numeral you want to convert.
- Formula result: = VIII
- Buttons: OK, Cancel
- Help link: [Help on this function](#)

The background shows a spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		7	8										
3		8											
4		5											
5		3											
6		4											
7		9											
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													

The status bar at the bottom shows "Point" and "NUM SCRL".