RNA Editing

Definition: any process, other than splicing, that results in a change in the sequence of a RNA transcript such that it differs from the sequence of the DNA template

First considered a bizarre relic; now recognized as widespread

RNA editing has been reported in:

protozoa, plants and mammals, not yet fungi or prokaryotes; nuclear, mitochondrial, chloroplast, and viral RNAs; mRNA, tRNA, rRNA

RNA editing

Two general types

Base modification (deaminase)

A to I double-stranded mechanism, seen in viruses, human genes

C to U, U to C seen in chloroplasts, plant mitochondria, human genes

Insertion/deletion

U insertion/deletion, seen in kinetoplastid protozoa

mono/di nucleotide insertion, seen in Physarum

nucleotide replacement, seen in Acanthamoeba tRNAs

Where were the real functional genes?

- Investigators generated cDNA clones to some of the kinetoplast mRNAs and sequenced them
- Sequences were partially complementary to pseudogenes on maxicircle DNA

cytochrome oxidas subunit II

COX II DNA: •••GTATAAAAGTAGA G A ACCTGG•••

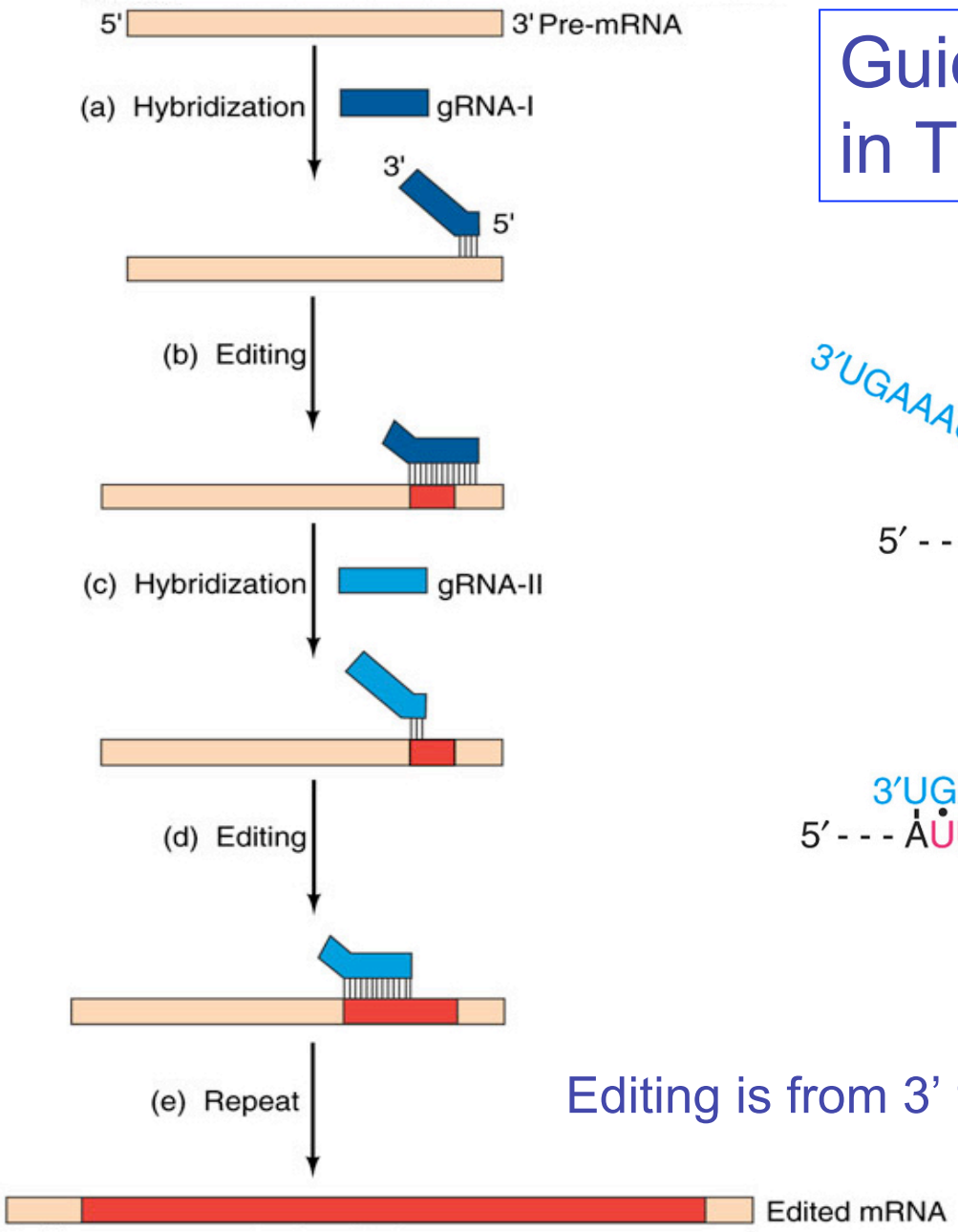
COX II RNA: •••GUAUAAAAGUAGAUUGUAUACCUGG•••

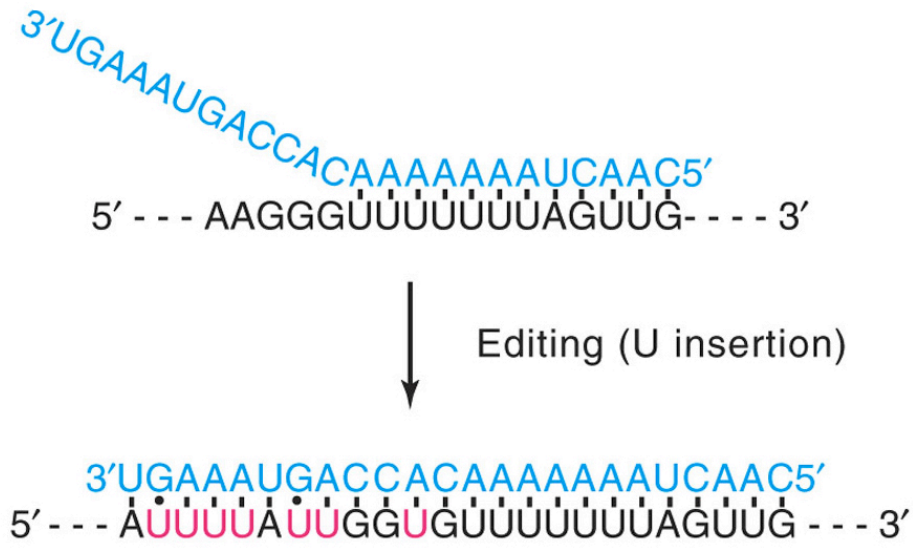
  - the COXII DNA sequence above is missing 4 Us found in the mRNA
- Called this “Editing” because it produced functional mRNAs and proteins from pseudogenes

Editing Mechanism

- **Post-transcriptional**
- **Guide RNAs** (gRNAs) direct editing
  - gRNAs are small and complementary to portions of the edited mRNA
  - Base-pairing of gRNA with unedited RNA gives mismatched regions, which are recognized by the editing machinery
  - Machinery includes an Endonuclease, a Terminal UridylylTransferase (TUTase), and a RNA ligase
- Editing is **directional**, from 3' to 5'

Guide RNAs Direct Editing in Trypanosomes.





Editing is from 3’ to 5’ along an unedited RNA.

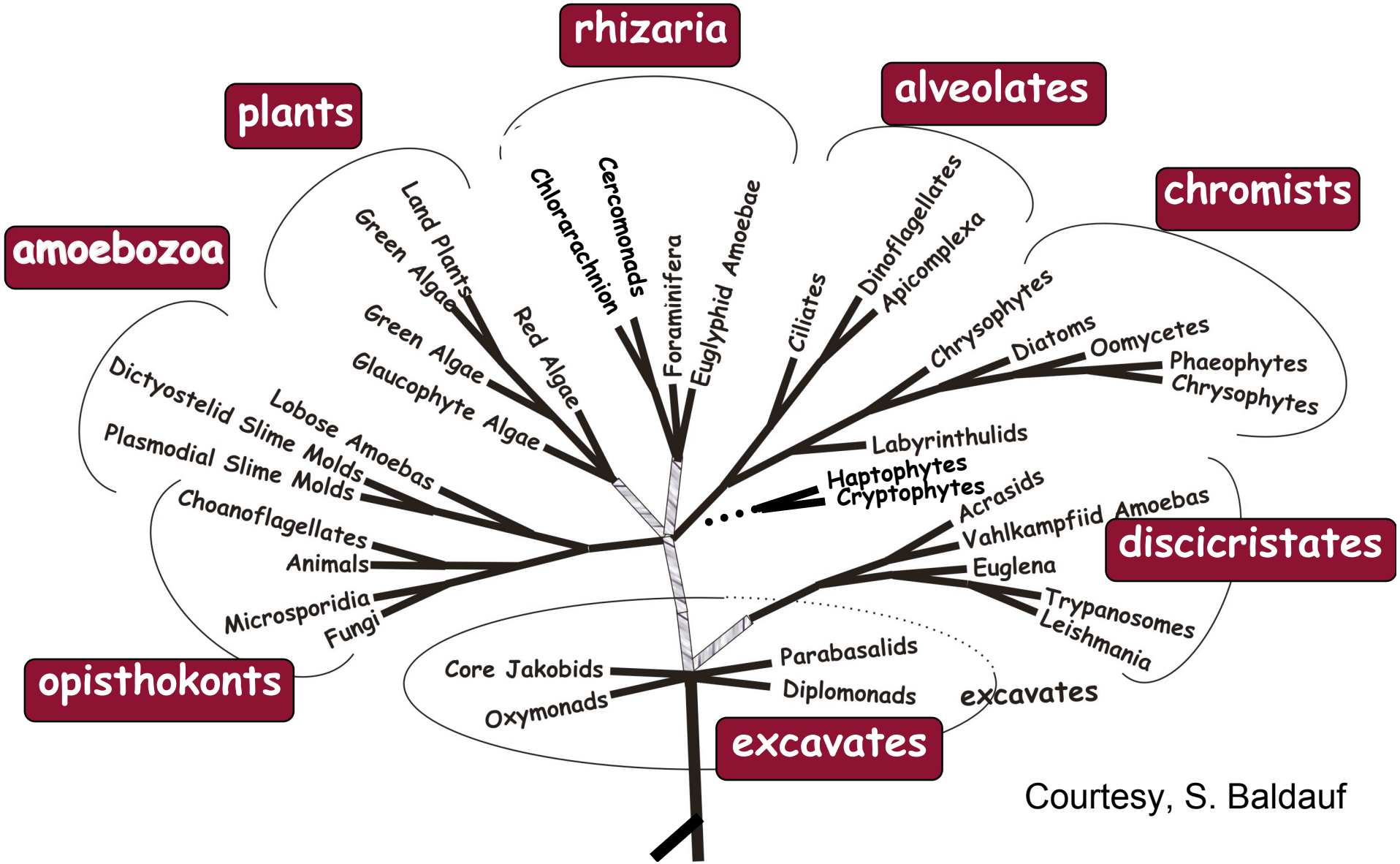
RNAs THAT FUNCTION IN RNA PROCESSING

rRNA	
snoRNAs	form complexes with protein, direct nt modifications snoRNAs also modify tRNAs, and likely other RNAs
tRNA	
RNase P	has both RNA and protein components
mRNA	
snRNPs	U1,2,4,5,6 form spliceosomes with many proteins
gRNAs	provide sequence information for RNA editing
miRNAs	important for regulating gene expression
siRNAs	important for regulating gene expression

Issue #1

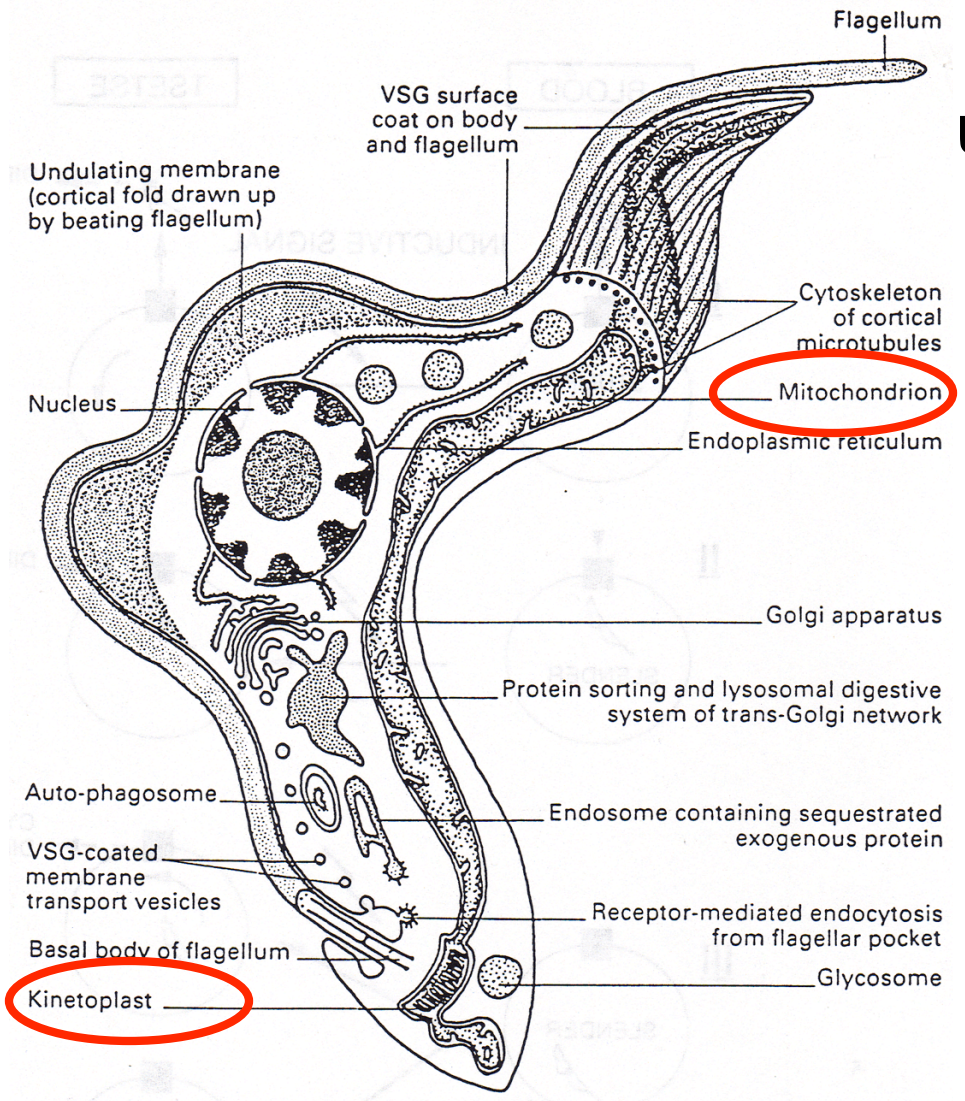
- Do we need to describe multiple RNA types, in this case:
  - mRNA (with features, e.g edits)
    - Pre-edited mRNA
    - Intermediate, partially edited mRNAs
    - Fully edited mature mRNA
  - gRNA (Exists)
    - Anchor site region (Exists), including G:U mismatch pair
    - Editing template region
    - polyU tail (post-transcriptional modification)

Major Groups of Eukaryotes



Courtesy, S. Baldauf

Trypanosomes

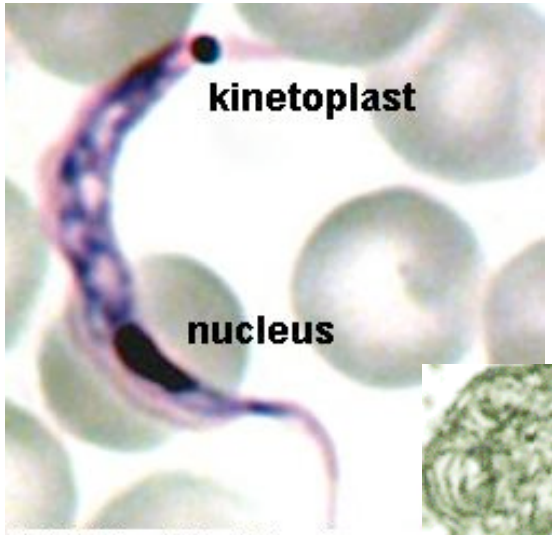


Unusual organelles:

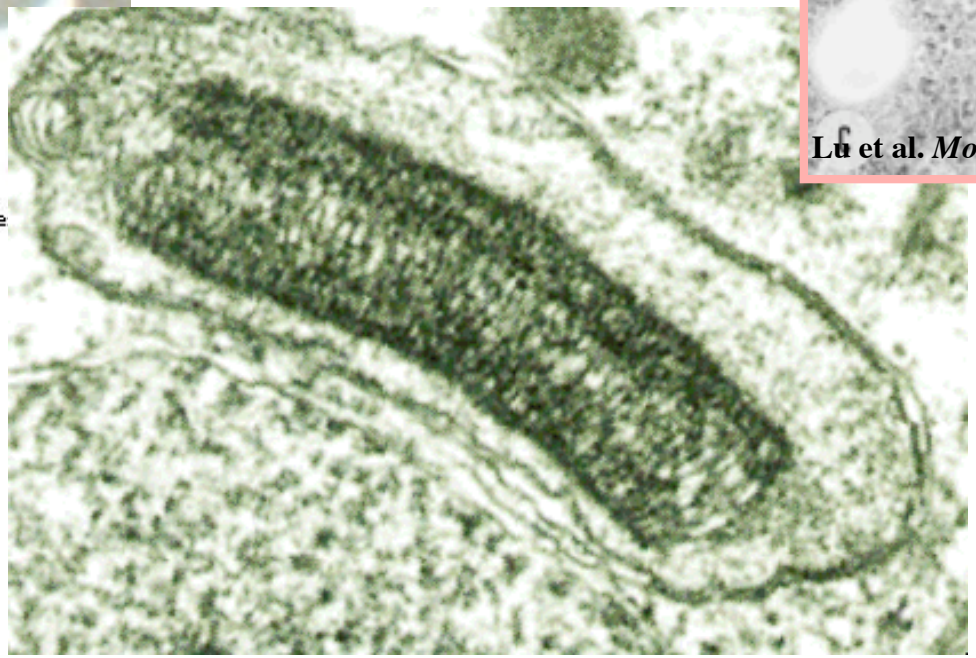
- **Flagellum**
- **Paraflagellar rod**
- **Subpellicular microtubules**
- Surface membrane is densely packed with a protein called VSG (variant surface glycoprotein)
- **Kinetoplast**: Region of the mitochondrion containing highly packed DNA (20% of total)
- **Glycosomes**: peroxisomes, contain all the glycolyitc enzymes

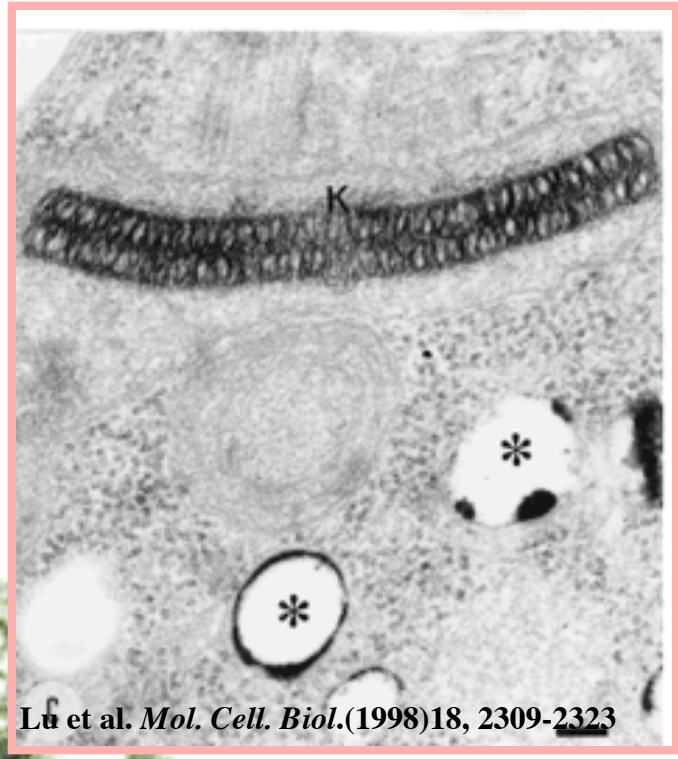
The Kinetoplast

DNA with a peculiar network structure
Disk shape structure near the flagellar basal body



Mark F. Wiser, Tulane University
www.tulane.edu/~wiser/protozoology/note





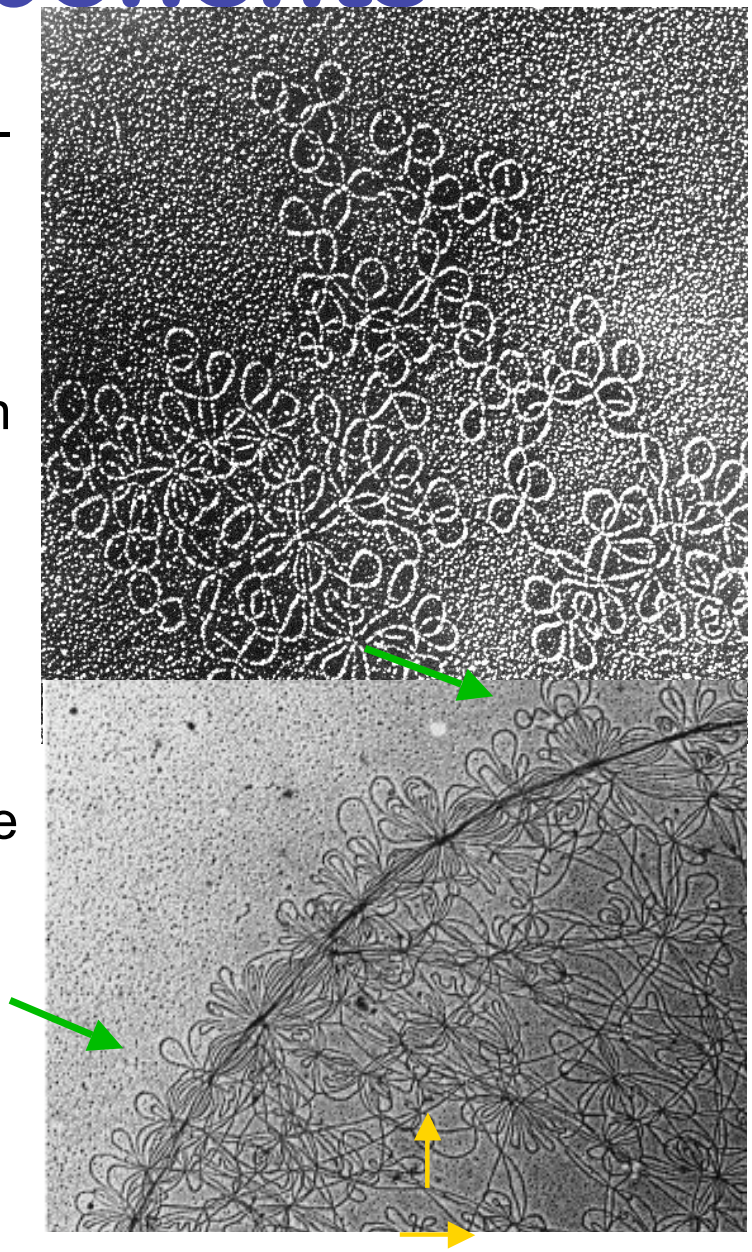
Lu et al. *Mol. Cell. Biol.*(1998)18, 2309-2323

KDNA components

- The network contains 5-20 x10^3 catenated minicircles (0.8-2.5 kb) and 20-50 catenated maxicircles (23-36 kb)

- **Maxicircles**: 20-40 kb. Functional homologue of mitochondrial DNA: Region encoding ribosomal RNAs, subunits of respiratory complexes and a variable region non-transcribed. Maxicircles transcripts undergo RNA editing.

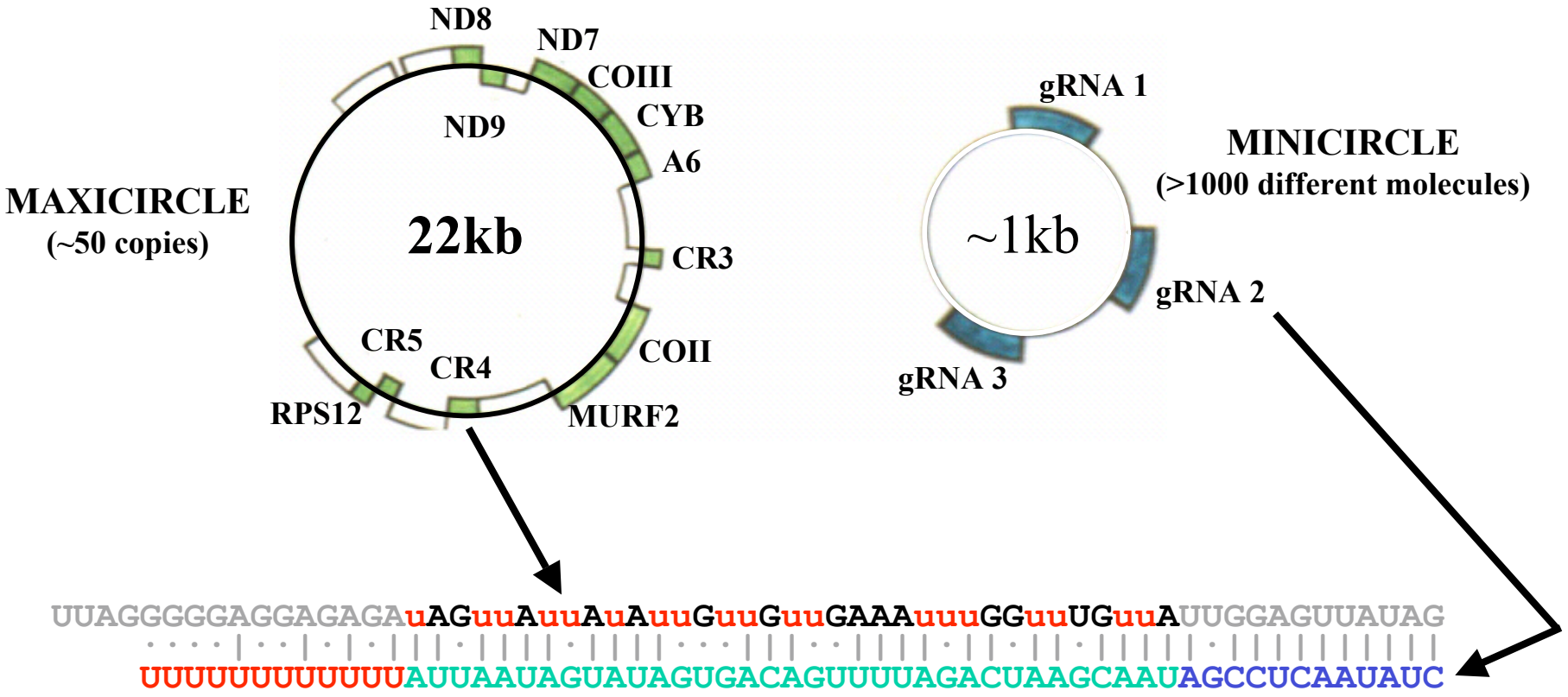
- **Minicircles**: 0.5-2.5 kb. Encode the guide RNAs. All minicircles share the UMS



RNA EDITING IN KINETOPLASTIDS

- The precursors of messenger RNAs (pre-mRNAs) have their coding information remodeled by the site-specific insertion and deletion of uridylate (U) residues.
- Pre-mRNAs are encoded in the maxicircle DNA
- This process creates initiation and termination codons, corrects frameshifts and even builds entire open-reading frames from nonsense sequences.
- The edited transcripts are translated into components of the oxidative phosphorylation: subunits of complex I (NADH-UQ reductase), III (Cyt bc1), IV (cyt oxidase) and V (ATP synthase)
- Minicircles encode guide RNAs (gRNAs) that specify the editing.

Pre-edited mRNAs on Maxicircle; guide RNAs on Minicircle



Courtesy of K. Stuart

Issue #2

- Need multiple types of mitochondrial genome sequence types
  - Maxi-circle (Exists)
    - mRNA features
  - Mini-circle
    - gRNA features

The extent of editing varies between mRNAs from the same organism

- Occurs only with mitochondrial genes encoded in maxicircles (12)
- **No editing**
  - 5 genes produce RNAs that do not appear to be edited: cox1, MURF1, ND1, ND4, ND5
- **Limited editing**:
  - 5'end of CYb, MURF2, cox2, ND7 and cox3 *(L. t. and C. f.),* ND8 *(C.f.)*
- **Extensive editing (pan edited cryptogenes)**:
  - *T. brucei*: ND7, cox3, MURF4, CR1-6, G1-6
  - *L. tarentolae*; 5'portion of ATP6 RNA, entire G6 transcript, ND8, ND9, G3, G4 and ND3

Some genes are very heavily edited!

COXIII
Cytochrome oxidase III

From *Trypanosoma brucei*

Lower case Us were inserted by editing. The deleted Ts (found in the DNA) are indicated in upper case.

Edited *T. brucei* ND7 mRNA

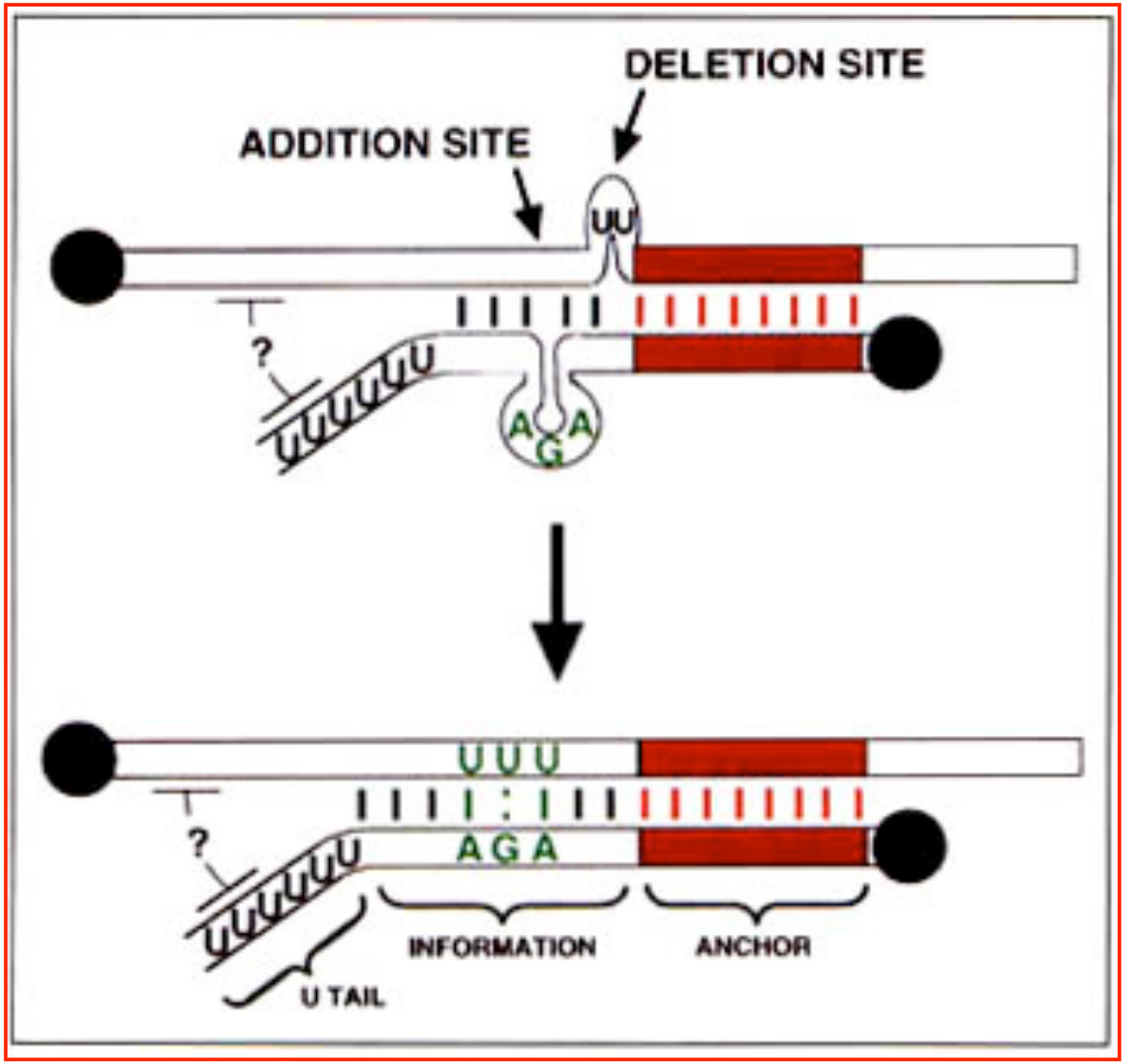
UGAUACAAAAAAACAUGACUACAUGAUAAGUAuCAuuuuAuGuuAuuuuu
GGuAGuuuuuuuuACAuuuGuAuCGuuuuACAuuuG*GUCCACAGCAuCCCG*
CAGCACAuGGuGuuuuAuGuuGuuuAuuGuAuuuuuGuGGuGA*AuuuAu
uGuuuA**UAUUGAuUGuAuuAuA***GGuuAUUUGCAUCGUGGUACAGA
AAAGUUAUGUGAAUAUAAAAGUGUAGAACAAUGUCUUCCGuAUUUC
GACAGGUUAGAuuAuGuuA*GuGuuuGuuGuAAuGAGCAuuuGuuGuCuuu
A***UGuuuuGAGuAuAuGuuGCGAuGuuGuuuGuCGuuACGuuGuGCAuuuA
uuAAuuGuA****GAAuuuAC***CCGuAGuuuuAAuGGuuuGuuGuGuAuAuC
AuGuAuGGuuuuGG*AuuuAGGuuGuuuGuCUCCGuuG*UUAuGAuCAuuuG
AGGAA***CG*UGACAAAuuGAuGACAuuuuuuGAuuuAuG**UUGuGGuuG
uCGuAuGCAuuuGGCUUUCAuGGuuuuAuuA*GGuAUUCUUGAUGAuuuuG
uuuuuGGuuuuGuuGAuuuuuuGuuGuuGuuGA***UAAuAuCAuGuuuGuuuGu
uAuGGAuuGuuAuGAuuuGuuAuuuGuGGGuAAUCGuuuAuuuUAuuuGCGuu
uGC***GuGGuuuGuCAuuuuuuGAuuuAuAuGAuuuA**GuuuuuA**A**UAG
uuuAAGuGGuGuuuuGuCuCGuuCGuuAGGuAuGGuGuGAGAuuGUCGuuu
AuuuAGuuGuuA****UGA****GuUGuAuuuuAuGuuuuGuuAuGAuuAuuGuu
uuuGuuuuAuAGGuGAuGCAuuuGA*UCGuuuAuuuuuACGuuuGuuuGAUAu
GCGuAuGAGuuuGuuGAuuuGuAAGCAAuGuuuuuuuGuuGGuuuuuuuGuuu
uuG*****GuuuuGuuuGuuuGuuuG**AuuAuuuAuAuuGuGAuAuuACCAuuG*
***AGACCAuuAuuAuGuuAuuuuAuAGuuuGuGGuGuuGuuGuuuGCCGGGu
AuA*UCAuuuGC*UUGUGuuGAACACCCCAAAGGuGA***GuAuuGuuuGu
uAuuA****UGuuuuuGuGuuGGuuuAuGuuCUCGuuuACGuuuGCGuuGuGC
GGAuuuuuuGCA*UAUUUGuuuAuuGGAuGuuuGuuuGCGuGGuuuuuuAuuG
CAuGAuuuAGuuGC***C*GuuuuAGGuAAuAuuGAuGuuGuuuuuGGAuCCG
UAGAUCGuuA*GuuuuAuAuGuG**A******GGUUAUUGuAGGAUUGUUU
AAAAUUGAAUAAAAA-poly(A)

Issue #3

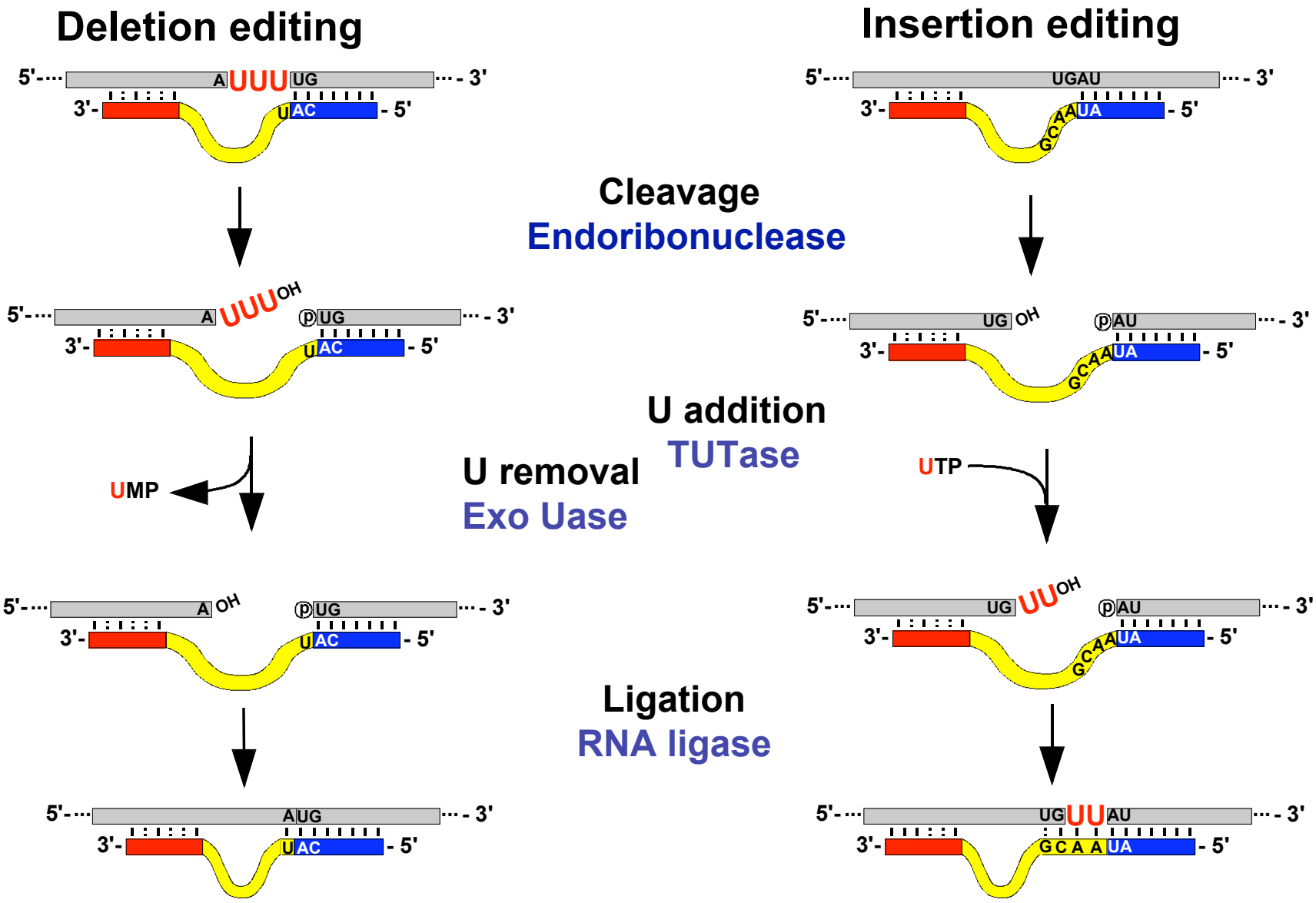
- Do we need to indicate the type and style of editing?
  - Types include :
    - U - Deletion (Exists)
    - With reference to DNA would be T deletion
    - U - Insertion (Exists)…would be a T insertion
  - Styles include:
    - Single/few edits
    - 5' editing
    - Pan editing

Guide RNA (gRNA)

- Complementary to edited RNA sections if G·U is allowed
- Small RNAs (40-70 nt)
- Guide because they are not conventional templates
- Structural elements: anchor, informational part and Oligo(U)tail



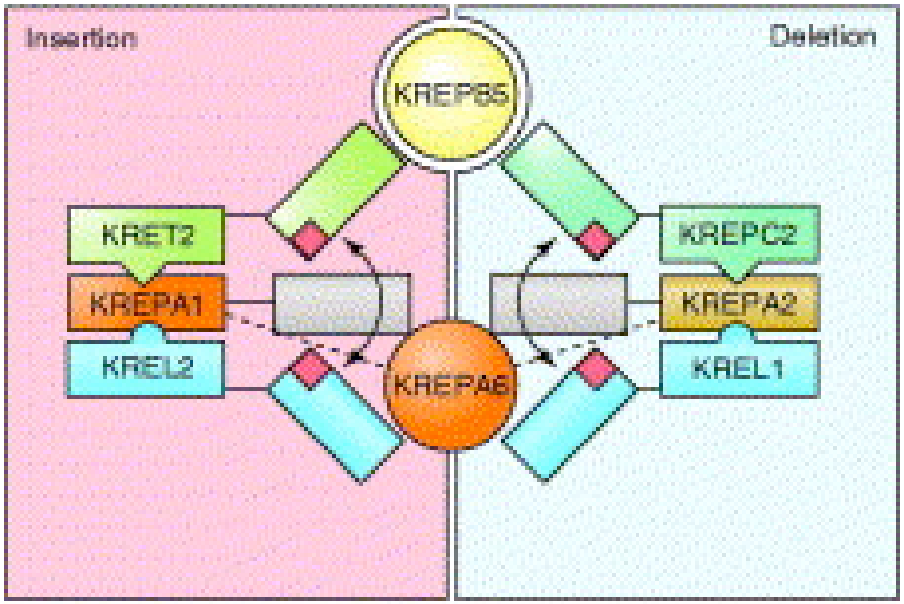
RNA editing mechanism



Editing is catalyzed by a multiprotein complex

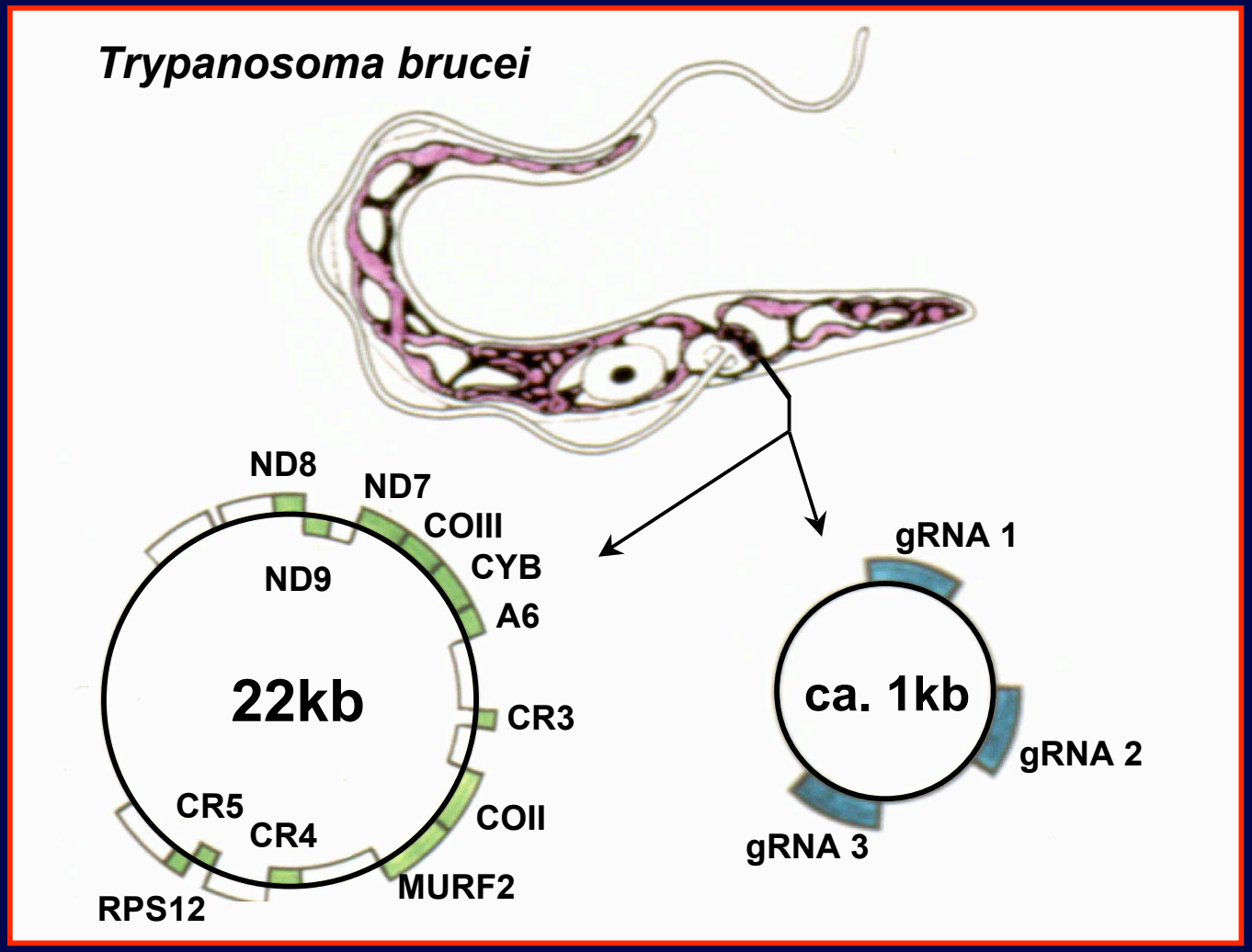
- Editing is catalyzed by a multiprotein complexes that have not been fully defined yet. A complex has been purified from glycerol gradients that contains the four key enzyme activities: **20S editosome**

- **Endonuclease:** cleavage in vitro occurs at an unpaired nucleotide immediately upstream of the gRNA-mRNA anchor duplex.
- **Exonuclease:** exoUase removes non-base-paired U nucleotides after cleavage of deletion editing sites
- **TUTase:In insertion editing, Us are added to the 3' end of the 5' pre-mRNA fragment by a terminal uridyl transferase as specified by the gRNA.**
- **RNA ligase:** the natural editing ligase substrates are nicked dsRNAs that are completely base-paired after the correct addition or removal of U nucleotides
- **Helicase:** each gRNA must be displaced from the sequence that it creates to enable binding by the subsequent gRNA and also from the mRNA completely before translation
- **Other 20S editosome proteins**



TIBS

The Kinetoplast and Its DNA



MAXICIRCLE
(~50 copies)

MINICIRCLE
(>1000 different molecules)

Courtesy of K. Stuart

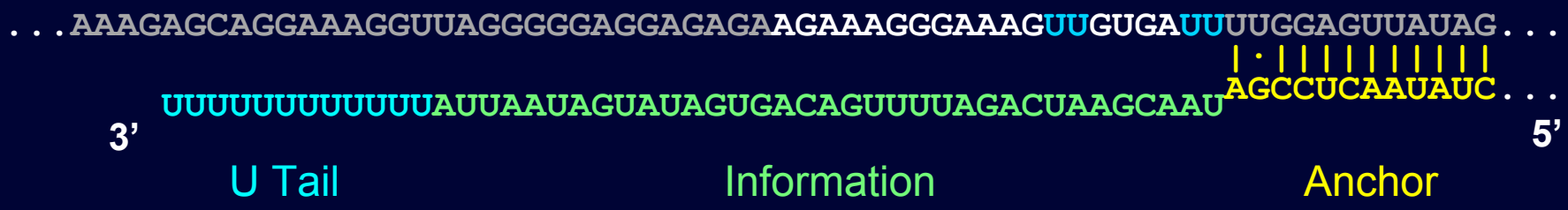
Sequence of *T. brucei* ATPase Subunit 6 (A6) mRNA

AAAAAUAAGUAUUUUGAUAUUAUUAAAGUAAAuAuGuuuuuAuuuuuuuuuuuGuGAuuuA
UUUUGGuuGCGuuuGuuAuuAuGuAuGuAuuAuuGuGuAuGAuCuAGGuuAuGuuuuAuu
GuGuAuuuuAAuUGuuuAAuGuuGAuuuuuGAuuuuuuAuuAuuuuGuuuG*UUUGAuuu
GuAuuuGuuuGuuGGuuuGuG***UUUGuuuuuAuuGuuGuGGuuuAuGuuGuuuAAuuu
AuAuAGuuuAAUUUUGuAuuA*UUGuAuuACuUAUUUG***AAuuuG*UAuuUGuuGuuu
uGuAuuGuuuuuuuAuuGuAuAuuGCAuuuuuAuuuuuGuuuuGuuuuuuAuGuGAuuuu
uuuuuGuuuAAuAAuuuGuUAGuuGGuGAuA****GuuuuAuGGAuGuuuuuuuuAUUC*
*GuuuuuuGuuGuGuuuuuuAGAGuGuuuuuCuuuGuuGuGuCGuuGuuuGuCGACGuuu
uuGCGuuuGUUUUGuAAuuuAuuAuCAuCCCAuUUUUUAuuGuuGAuGuuuuuuGAuuuu
uuuUAuuuuAuuuuuGuuuuuuuuuuuuuAuGGuGuuuuuuGuuAuuGAuuuAuuuuAuuu
AuuuuuGuGuuuuGuuuuuGuuuAuuAuuuuAUGuGuuuuuAuAuUUGuuGGAuuuAuUU
GCC***GCCAuAuuAC****AGuuAuuuAuuuuuuGuAAuAuGAuuuuGCAGuuGAuAAu
GG**AuuuuuuGuuGuuuuuGuuGuuuGuuuAGuuuuGuAuuuGAuuuuuGAuAGuuAuu
AuAuuGuuGuuGAAAuuuG**GuuUGuuA**UUGGAGUUAUAGAAUAAGAUCAAAUAAGU
UAAUAAUA

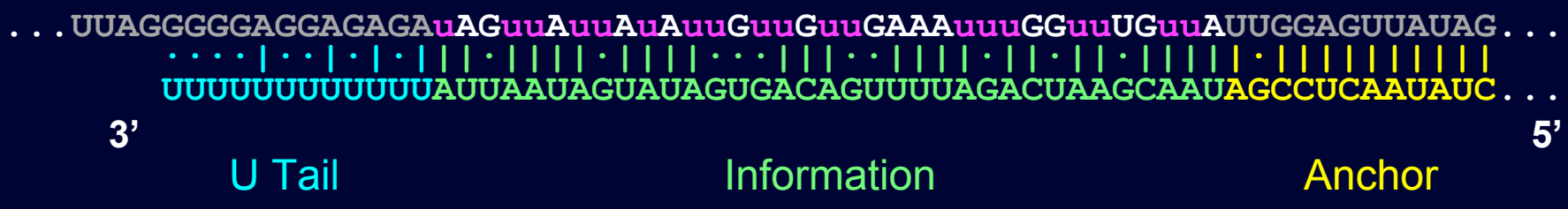
Courtesy of K. Stuart

Structure of a Guide RNA (gRNA)

A6 Pre-mRNA



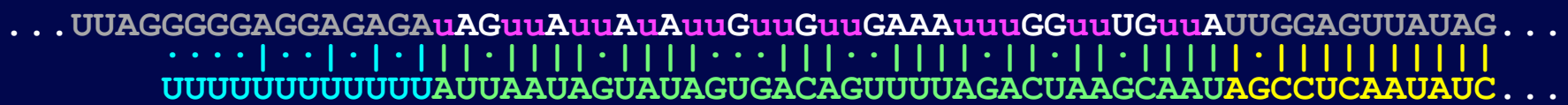
gA6[14] gRNA



Courtesy of K. Stuart

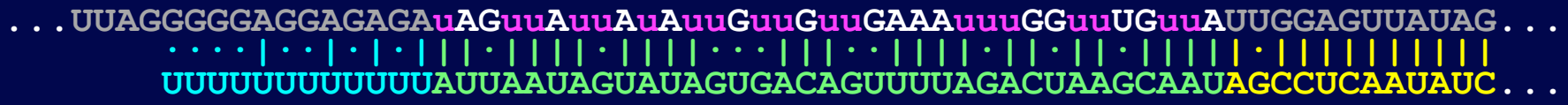
Two Models for Editing Directed by One gRNA

1. The Zipper



U Tail Information Anchor

2. Dynamic Realignment

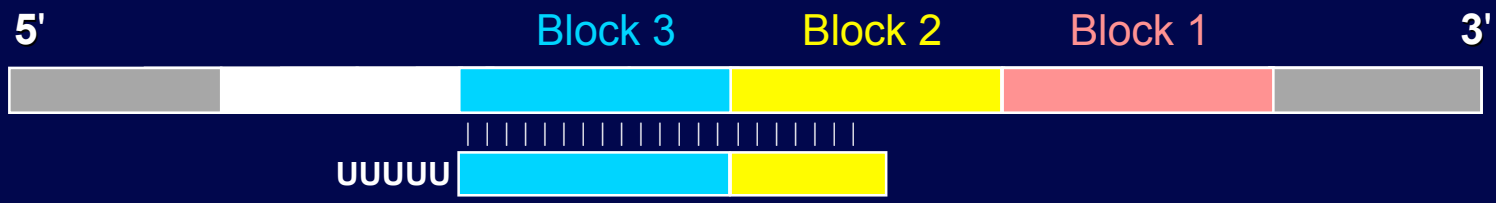


U Tail Information Anchor

Courtesy of K. Stuart

Editing Proceeds in the 3′ → 5′ Direction

Editing of a downstream region is required for binding of the subsequent gRNA's anchor region.



Courtesy of K. Stuart

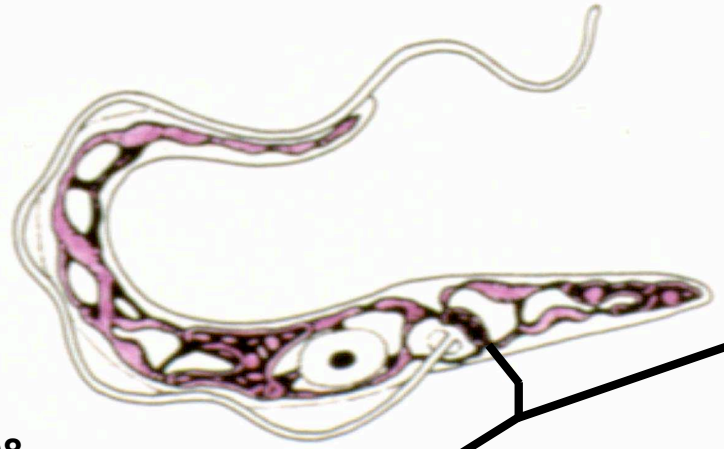
edited *T. brucei* ATPase subunit 6 RNA

edited

```
                                    M  F  L  F  F  F  C  D
L  F  W  L  R  L  L  L  C  M  Y  Y  C  V  W  S  R  L  C  F
  I  V  Y  F  N  C  L  M  L  I  F  D  F  L  L  F  C  L  F
  D  L  Y  L  F  V  G  L  C     L  F  L  L  L  W  F  M  L
 F  N  L  Y  S  L  I  L  Y  Y  C  I  T  Y  L     N  L  Y
  L  L  F  C  I  V  F  L  L  Y  I  A  F  L  F  L  F  C  F
 L  C  D  F  F  L  F  N  N  L  L  V  G  D     S  F  M  D
 V  F  F  I     R  F  L  L  C  F  L  E  C  F  S  L  L  C  R
  C  L  S  T  F  L  R  L  F  C  N  L  L  S  S  H  F  L  L
 L  M  F  F  D  F  F  Y  F  I  F  V  F  F  F  W  C  F  L
L  L  I  Y  F  I  Y  F  C  V  L  F  L  F  I  I  L  C  V  F
  I  F  V  G  F  I  C     R  H  I  T     V  I  Y  F  L ter
```

Bhat *et al.* (1990)

RNA editing in kinetoplastid mitochondria



gRNA 1

~1kb

gRNA 2

gRNA 3

MINICIRCLE

(>1000 different molecules)

1000s of gRNAs

ND8

ND7

COIII

CYB

A6

CR3

COII

MURF2

RPS12

CR4

CR5

ND9

22kb

MAXICIRCLE

(~50 copies)

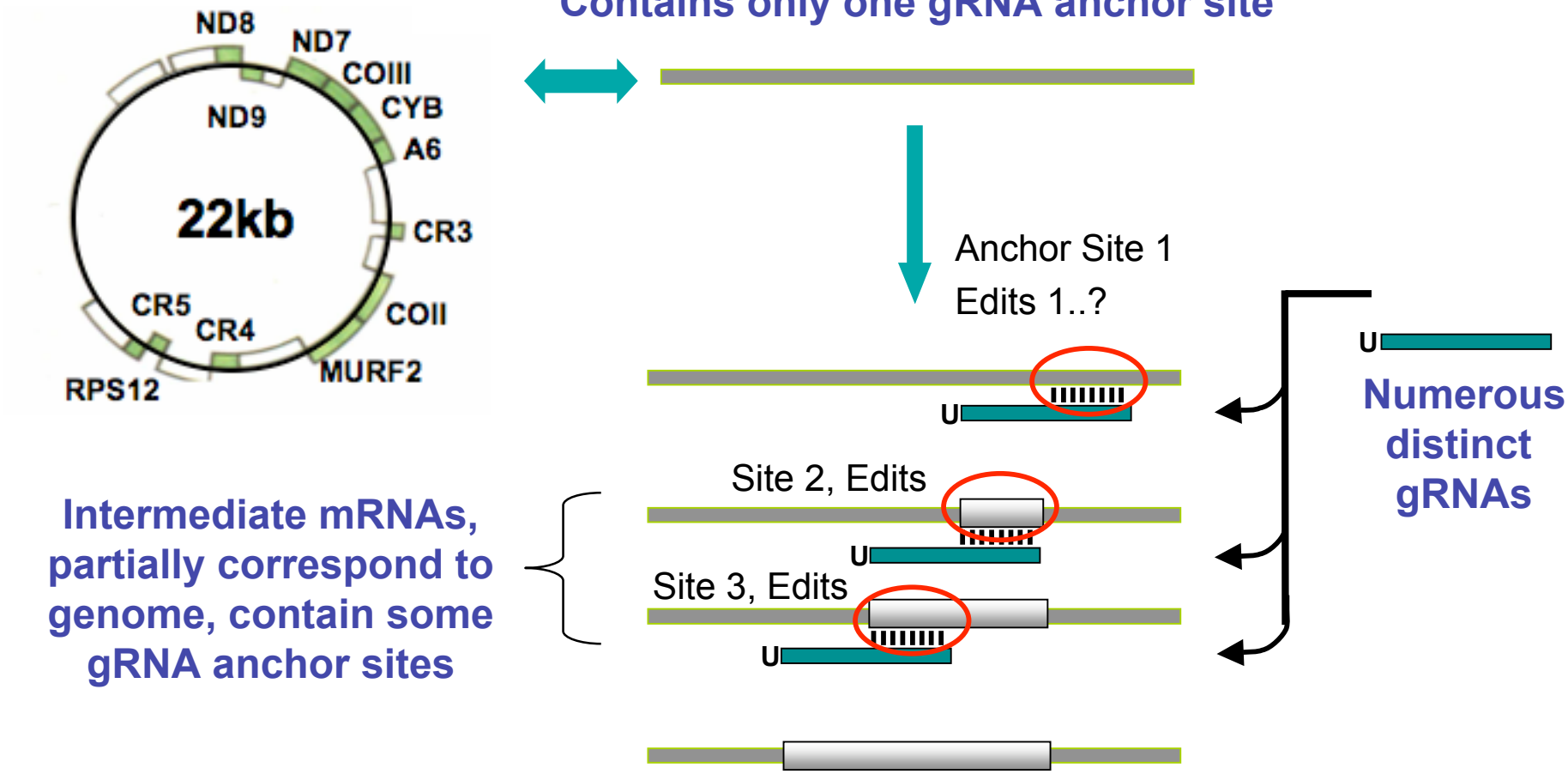
12 pre-edited RNAs

U

U

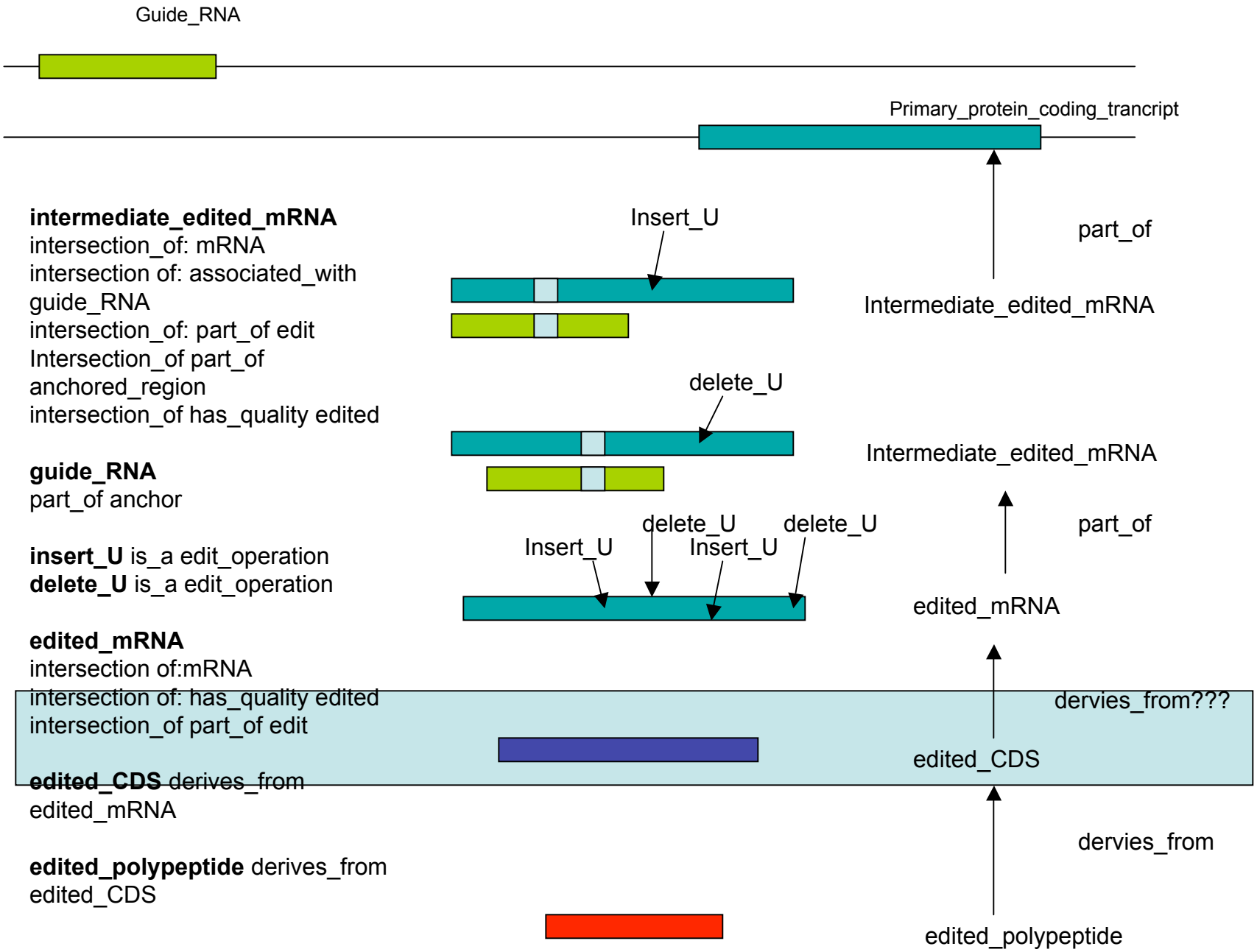
U

Courtesy of K. Stuart


Pre-mRNA - nonsense ORF
Corresponds to genome sequence
Contains only one gRNA anchor site
ND8
ND7
COIII
CYB
ND9
A6
22kb
CR3
CR5
CR4
COII
RPS12
MURF2
Anchor Site 1
Edits 1..?
U
Numerous distinct gRNAs
Intermediate mRNAs, partially correspond to genome, contain some gRNA anchor sites
Site 2, Edits
Site 3, Edits
Mature mRNA - sensible ORF
Does not correspond to genome sequence
Contains all gRNA anchor sites

Issue #4

- Do we need to create RNA intermediates?
  - Currently, we don't have mRNA splice intermediates
  - A gRNA only has an anchor site on an edited RNA (except first gRNA), It cannot be mapped to the Maxi-circle DNA directly
- Do we need to label/capture the 3'→5' direction of processing? Do other systems do it differently?


Guide_RNA
Primary_protein_coding_trancript
intermediate_edited_mRNA
intersection_of: mRNA
intersection of: associated_with guide_RNA
intersection_of: part_of edit
Intersection_of part_of anchored_region
intersection_of has_quality edited
guide_RNA
part_of anchor
insert_U is_a edit_operation
delete_U is_a edit_operation
edited_mRNA
intersection of:mRNA
intersection of: has_quality edited
intersection_of part_of edit
edited_CDS derives_from edited_mRNA
edited_polypeptide derives_from edited_CDS
Insert_U
part_of
Intermediate_edited_mRNA
delete_U
Intermediate_edited_mRNA
delete_U
delete_U
Insert_U
Insert_U
part_of
edited_mRNA
dervies_from???
edited_CDS
dervies_from
edited_polypeptide