OGF Mission 23 – DNA Analysis Toolkit

Companion artifact to Mission 23: Advanced DNA – Building Genetic Networks & Solving Mysteries

---

How to use this toolkit. Five tools, five problems. Use them in order or jump to the one that matches where you are stuck. Every tool is designed to work with AncestryDNA as your primary platform; the chromosome mapping guide requires a second platform (GEDmatch or MyHeritage – both free for upload). The contact-message template (Part 5) is the one you reach for the moment a cluster points at a person you need to email.

Part 1 – Shared-Match Clustering Template

For researchers who want to sort their match list into family-branch groups.

What this is

A structured spreadsheet template for the four-color clustering method described in Mission 23. The output is a color-coded map of your match list where each color corresponds to one branch of your family tree.

How to use it

Setup (do this once)

Column	Label	What goes here
A	Match Name	As shown in AncestryDNA
B	Shared cM	Copy from AncestryDNA match card
C	Predicted Relationship	From AncestryDNA (e.g., "2nd cousin")
D	Branch Color	Assigned below – blank until you run the clustering pass
E	Branch Label	Plain text: "Maternal grandmother's side," etc.
F	Notes	Any tree connection you can identify

Clustering pass (do this once per branch, starting with the match you know best)

1. Find your closest known match on Branch 1 (e.g., a confirmed second cousin on your maternal grandmother's side). Write her name in Row 2, Column A. Enter her shared cM. Assign **Color 1** (yellow, say) to her Branch Color cell.

2. Click her name in AncestryDNA. Click "Shared Matches." Copy every shared-match name that appears into your spreadsheet. Assign **Color 1** to each. These people are all likely on your maternal grandmother's side.

3. Return to your match list. Find your closest known match on Branch 2 (maternal grandfather's side, if possible). Assign **Color 2**. Repeat the shared-matches pull.

4. Repeat for Branches 3 and 4.

Reading the result

- **Matches in one color only**: almost certainly on that branch.
- **Matches in two colors**: may have ancestry on both branches (endogamy), or may be a more distant match whose DNA is too small to cluster cleanly.
- **Uncolored matches below 30 cM**: set aside. Too distant to sort reliably without additional paper research.
- **Uncolored matches above 50 cM**: these are your brick walls. A match this close who shares with nobody you can identify is worth a direct message.

Sample template structure

```
Name           | cM   | Predicted | Color   | Branch Label              |
Notes
---------------|------|-----------|---------|---------------------------|----
---
Jane Doe       | 312  | 1st-2nd   | Yellow  | Maternal grandmother side |
Confirmed at reunion
Robert Smith   | 178  | 2nd       | Yellow  | Maternal grandmother side |
Shares with Jane
Unknown User   | 89   | 3rd       | Blue    | Maternal grandfather side |
Shares with Tom C.
Mary Johnson   |  44  | 3rd-4th   |         | (uncolored -- too distant)|
```

When the clusters don't form cleanly

If your match list shows heavy overlap between colors (many matches appearing in three or four groups), you may be researching an endogamous community. See the endogamy note in Part 3 and in Mission 23 Section 5.

Part 2 – Chromosome Mapping Visual Guide

For researchers ready to go beyond clustering to segment-level analysis.

Before you start

Chromosome mapping requires a platform with a chromosome browser. AncestryDNA does not offer one. Your options:

Platform	Cost	What you get

GEDmatch	Free (upload your AncestryDNA raw file)	Chromosome browser, segment data, one-to-many comparison
MyHeritage	Free upload; paid for chromosome browser	Chromosome browser with segment triangulation
FamilyTreeDNA	Transfer free; paid tools vary	Full chromosome browser, bucket feature
23andMe	Only if you tested there	Built-in chromosome browser

Recommended path for most researchers: Upload to GEDmatch (free). The chromosome browser at gedmatch.com gives you the segment data you need for this guide.

Understanding chromosome segments

Your DNA is organized into 23 pairs of chromosomes (plus mitochondrial DNA, which is a separate research track). On each autosomal chromosome (pairs 1-22), you inherited one copy from your mother and one from your father. Each copy was itself assembled from your grandparents' DNA via recombination.

A **segment** is a stretch of DNA on a specific chromosome that you share with a match. When a tool reports "Chromosome 5, positions 45,000,000 to 67,000,000, 28 cM," it means you share a 28-centimorgan stretch of DNA on chromosome 5 with that match.

The core insight: **if two matches share the same segment with you – same chromosome, overlapping positions – they almost certainly share an ancestor with each other, not just with you.** This is triangulation.

The chromosome mapping workflow

Step 1: Export your match segment data from GEDmatch

1. Log in at gedmatch.com.
2. Go to "Tier 1 Tools" > "One-to-one comparison" for your closest matches. Or use "People who match one or both of 2 kits" to triangulate.
3. For each match, note: Chromosome number, Start position, End position, cM length, SNPs.

Step 2: Build a segment map in a spreadsheet

Column	Label
A	Match Name
B	Chromosome #
C	Start Position
D	End Position
E	cM
F	Likely Branch (from clustering, Part 1)
G	Ancestor Assigned (fill in as you identify)

Step 3: Triangulate

For any two matches whose segments overlap on the same chromosome: - Compare their trees. - If their trees share a surname or couple you recognize, you have likely found the ancestor who passed that segment to you. - Enter that ancestor in Column G for both rows.

Step 4: Paint the chromosome (visual tracking)

Once several segments are assigned to ancestors, you can "paint" a visual chromosome map. Tools that do this automatically:

- **DNA Painter** (dnapainter.com) – free tier supports chromosome painting; paste segment data directly; highly recommended for researchers working 10+ matches
- **Spreadsheet painting** – use conditional formatting in Excel or Google Sheets to color each row by the ancestor assigned in Column G

What triangulation looks like

You have a 45-cM match on Chromosome 7. You also have a 38-cM match on Chromosome 7 whose segment overlaps the first match's segment. You compare both trees. Both trees include the surname "Petersen" in Sanpete County, Utah. You have just placed a segment on Chromosome 7: it came from your Petersen great-great-grandparents.

Granny's note: this is the slow work, dear. One segment at a time. But each segment is a brick placed in the wall of what you know, and after two dozen segments the wall starts to show a shape.

Endogamy adjustment for chromosome mapping

In endogamous communities, identical-by-descent (IBD) segments from a single ancestor can arrive via multiple pathways. Two matches who appear to triangulate may actually have inherited the same segment through different ancestors who themselves shared common ancestry. The practical adjustment: require **three-way triangulation** (you + match A + match B all sharing the segment) before assigning an ancestor, and cross-check against paper records before writing the connection into your tree.

Part 3 – Leeds Method Reference Card

One-page working reference. Print and prop next to your laptop.

The Leeds Method – Quick Reference

What it does: Sorts your 2nd-through-4th cousin matches (approx. 30-90 cM) into four color groups, each corresponding to one of your four great-grandparent lines.

When to use it: When clustering (Part 1) gives you too many colors and you need a more structured approach. The Leeds method is especially effective when you have 50+ matches in the 30-90 cM range.

You will need: A spreadsheet (Excel or Google Sheets). AncestryDNA login. About 2-3 hours for the first pass.

Step-by-step

1. Build your match list

Export your AncestryDNA matches in the 30-90 cM range into a spreadsheet. Columns needed: Name | cM | Any note you already have. Sort by cM descending (highest first).

2. Color the first uncolored row

Start at the top. The highest uncolored match gets **Color 1**. Go to that match's Shared Matches in AncestryDNA. Every shared match you find in your spreadsheet gets Color 1.

3. Move to the next uncolored row

The next row with no color assigned gets **Color 2**. Repeat the shared-match coloring.

4. Continue until all matches are colored

Most match lists produce 4 color groups. Occasionally 3 (if one great-grandparent line is underrepresented in the database) or 5-6 (if one line had an endogamous subpopulation).

5. Label the groups

Once you can identify at least one match in each group with a confirmed family connection, label the group. Example: Color 1 = "Anderson/Jensen line (maternal grandma's side)."

Interpreting your results

Pattern	Meaning
4 clean color groups	Normal – one group per great-grandparent line
3 groups	One line underrepresented – common for lines from small/isolated communities or countries with low DNA testing rates
5+ groups	Possible endogamy, or one great-grandparent line has both known and unknown branches
Matches in multiple groups	Either too distant (< 30 cM) or endogamy is present

---

Endogamy note

If you are researching an LDS pioneer-stock, Ashkenazi Jewish, Acadian, Mennonite/Amish, or other closely intermarried community: your color groups will overlap significantly. This is normal. The Leeds method still works; you will just have more matches in multiple groups. The adjustment: focus on your **highest matches** within each group and use paper records to anchor the branch assignment before trusting the color alone.

The Leeds method vs. plain clustering

	Leeds Method	Plain Clustering (Part 1)
Best for	30-90 cM band; structured four-quadrant output	Flexible; works at any cM range
Output	Four quadrants mapped to great-grandparent lines	Color groups of variable number
Time	2-3 hours first pass	1-2 hours first pass
Requires	Only shared-match data from AncestryDNA	Only shared-match data from AncestryDNA
Works with endogamy	Yes, with adjustment	Yes, with adjustment

---

Worked example — one researcher's Leeds pass *(page 2)*

The steps above tell you what to do. This is what it actually looks like when you sit down and do it. Follow along — the names are invented, but the pattern is exactly what you will see on your own screen.

Meet **Carol.** Carol knows three of her four grandparents well. The fourth — her father's father, who left the family when her dad was small — is a blank. She has tested at AncestryDNA, she has 47 matches in the 30–90 cM range, and she has two hours on a Saturday morning. Here is her Leeds pass.

She exports and sorts. Carol pulls her 47 matches into a spreadsheet, three columns: Name, cM, Notes. She sorts highest-to-lowest. Her top match in the band is **"R. Whitfield," 88 cM.**

Row 1 — Color 1 (yellow). Carol clicks R. Whitfield in AncestryDNA, opens *Shared Matches*, and finds nine names that are also in her spreadsheet. She paints all nine — plus R. Whitfield — **yellow.** She recognizes two of the nine: both are confirmed cousins on her **maternal grandmother's** side. So yellow is already labeled in her head: *Mom's mom.*

Next uncolored row — Color 2 (blue). The highest match with no color yet is **"J. Okafor," 71 cM.** Shared Matches gives Carol six names. She paints them **blue.** One of the six is a second cousin she met at a funeral — **maternal grandfather's** side. Blue is *Mom's dad.*

Next uncolored row — Color 3 (green). "the_hansen_tree," 64 cM. Shared Matches returns eight names, painted **green.** Carol recognizes a surname from her father's mother's family — Hansen. Green is *Dad's mom.*

Next uncolored row — Color 4 (pink). "S. Delgado," 58 cM. Shared Matches returns five names. None of them appear in yellow, blue, or green. None of them appear in any tree Carol recognizes. She paints them **pink** and labels the group, for now, *unknown — probably Dad's dad.*

She reads the result. Four clean color groups, which is exactly what a Leeds pass is supposed to produce — one per great-grandparent line, give or take. Three of her four groups map to grandparents she already knows. **The pink group is the one that matters.** Five matches, sharing 30–58 cM with Carol, sharing with each other, and connected to *none* of the family she can name. That is not noise. That is her missing grandfather's family, sitting right there in pink.

What she does next. Carol does not write anything into her tree yet — pink is a lead, not a conclusion. She opens the trees attached to her two strongest pink matches (S. Delgado, 58 cM, and the next pink match down, 41 cM). Both trees contain a couple surnamed **Reyes** in Cameron County, Texas, in the right generation. That convergence — two independent pink matches, same surname, same place, right years — is her next thread to pull. She moves to Part 5 of this toolkit and drafts a message to S. Delgado.

> **GRANNY SAYS:** Notice what Carol did *not* do, dear. She did not decide her grandfather was a Reyes on the strength of a color. Pink told her *where to look* — it pointed at a household, not a name. The Reyes connection is a hypothesis now, and a hypothesis is a lovely thing to have when an hour ago you had a blank. But a hypothesis earns its place in the tree only after the paper agrees. Color first, conversation second, courthouse third.

The one-paragraph version, for when you forget: Sort your 30–90 cM matches highest first. Color the top uncolored match and everyone it shares with. Repeat down the list. You will land on three or four groups. The groups you recognize confirm what you knew; the group you *don't* recognize is the answer to whatever question made you test in the first place.

Credit

Dana Leeds developed and published this method. Cite her when you teach it to someone else.

> Leeds, Dana. "The Leeds Method for Autosomal DNA." danaleeds.com.

Part 4 – Unknown-Parentage Worksheet

For researchers working to identify one or both biological parents using DNA evidence.

> **A note before you begin.** Unknown-parentage research sometimes surfaces information that is unexpected, sensitive, or difficult for family members – including the researcher. This worksheet is a tool for organizing evidence; the human dimensions of the search are real and deserve care. If you reach a point where the evidence is pointing at a conclusion that will affect living people, Granny recommends slowing down and deciding deliberately how to proceed. The DNA Detectives and ISOGG communities (links in Further Reading of the main post) have experienced searchers who have been where you are.

Phase 1 – Establish what you know

Before DNA can do its work, document the baseline.

Question	Your answer
Which parent is unknown?	Mother / Father / Both
What is known about the unknown parent?	(Name fragments, location, time period, any identifying information)
What is the parentage event?	(NPE – non-paternity event; adoption; donor conception; foundling; other)
Who has tested?	(List names and platforms)
Closest known match and their confirmed relationship	

Phase 2 – Categorize your closest matches

List your 10 clo