

# Transposable Element AI Training Guide

## Key Metrics to Monitor

### 1. Mean Fitness (Target: 0.7-0.9)

Mean fitness: 0.8250 ← GOOD (High baseline fitness)

Mean fitness: 0.7899 ← OK (Slight drop is normal early on)

- **Good:** Steady or increasing fitness, especially after stress events
- **Bad:** Continuous decline or stuck below 0.5
- **What it means:** How well the population recognizes antigens

### 2. Population Stress (0.0-1.0)

Population stress: 0.0000 ← Currently no stress (pre-mutation)

- **Expected Pattern:**
  - 0.0-0.3: Normal evolution
  - 0.3-0.7: Mild stress, increased transposition
  - 0.7-1.0: High stress, transposition cascade

### 3. Population Size

Population after selection: 175 → 305 → 531 ← GOOD (Healthy growth)

- **Good:** Controlled growth (100-5000)
- **Bad:** Explosion beyond max\_population or crash to <50
- **What it means:** Diversity and adaptation capacity

## Transposition Events to Watch

Good Patterns 

## 1. Diverse Transposition Types

-  Gene jumped ← Exploration
-  Gene duplicated ← Creating variants
-  Gene inverted ← Trying opposites

You want to see all types, not just one.

## 2. Position-Based Specialization

- V genes: 0.0-0.3 (early position)
- D genes: 0.3-0.6 (middle position)
- J genes: 0.6-1.0 (late position)

Genes should mostly stay in their zones but occasionally explore.

## 3. Stress-Responsive Bursts

- Generation 50: 2-3 transpositions ← Normal
-  VIRUS MUTATED!
- Generation 51: 15-20 transpositions ← Good response!

## Warning Signs

### 1. Excessive Deletions

-  Gene deleted (silenced)
-  Gene deleted (silenced)
-  Gene deleted (silenced) ← Too many!

Some deletion is normal, but >30% indicates problems.

### 2. No Transpositions During Stress

-  VIRUS MUTATED!
- Generation X: 0 transpositions ← BAD! System not responding

### 3. Position Chaos

- V gene at position 0.95 ← V genes shouldn't be this late
- J gene at position 0.05 ← J genes shouldn't be this early

## Training Phases to Expect

### Phase 1: Initial Adaptation (Gen 1-50)

- **Fitness:** May drop slightly as system explores
- **Transpositions:** Low rate (1-5 per generation)
- **Population:** Steady growth
- **What's happening:** Building baseline diversity

## Phase 2: First Viral Mutation (Gen 50)

 VIRUS MUTATED TO Alpha Variant! Sites: [5]

- **Fitness:** Should drop sharply (0.8 → 0.4-0.5)
- **Stress:** Spikes to 0.8-1.0
- **Transpositions:** Burst of activity (10-30 events)
- **Good sign:** Fitness recovers within 10-20 generations

## Phase 3: Subsequent Mutations

Each mutation should show:

1. Initial fitness drop (smaller each time)
2. Stress spike (may be lower)
3. Transposition burst (more targeted)
4. Faster recovery (5-10 generations)

## Phase 4: Escape Variant (Gen 250+)

 VIRUS MUTATED TO Hypothetical Escape Variant! Sites: [1, 3, 5, 7, 9, 12, 15, 17, 18]

- **Success Indicator:** System maintains >0.6 fitness despite massive mutation

## Key Performance Indicators

### Adaptation Speed

python

Good: Fitness recovery in <20 generations

Bad: Fitness still low after 50 generations

### Transposition Efficiency

python

Good: Stress → Burst → Quick stabilization

Bad: Constant high transposition (thrashing)

## Diversity Maintenance

python

Good: 50-200 unique gene configurations

Bad: <10 variants (convergence) or >1000 (chaos)

## Module Families

Look for emergence of gene families:

V12 → V12-copy1 → V12-copy1-mut3 ← Gene family tree

## Optimization Targets

1. **Primary Goal:** Maintain fitness >0.7 across all viral variants
2. **Secondary Goal:** Minimize time to recover from mutations
3. **Tertiary Goal:** Develop reusable gene modules

## Live Monitoring Commands

Add these to your code for better insights:

python

*# After each generation*

```
print(f"Active genes per cell: {np.mean([len([g for g in cell.genes if g.is_active]) for cell in center.population.values()]):.1f}")
print(f"Unique gene families: {len(set([g.gene_type + str(g.variant_id) for cell in center.population.values() for g in cell.ge
```

## Red Flags

1. **Fitness Collapse:** Never recovers above 0.5
2. **Population Crash:** <50 cells remaining
3. **Transposition Freeze:** No events during high stress
4. **Memory Explosion:** >20 genes per cell
5. **Premature Convergence:** All cells identical

## Success Metrics Summary

Metric	Bad	OK	Good	Excellent
Fitness Recovery Time	>50 gen	20-50 gen	10-20 gen	<10 gen
Population Size	<50 or >5000	50-100	100-1000	500-2000
Transposition Rate (normal)	0 or >10	5-10	1-5	2-4
Transposition Rate (stress)	<5 or >50	5-10	10-30	15-25
Gene Diversity	<10 types	10-30	30-100	50-150
Stress Response Time	>5 gen	2-5 gen	1-2 gen	Immediate

## What Your Current Output Shows

Looking at your output:

- Generation 1-4:
- ✔ Good fitness (0.79-0.82)
  - ✔ Healthy population growth
  - ✔ Diverse transposition types
  - ⚠ No stress yet (waiting for virus)

You're in the **baseline building phase** - this looks perfectly normal! The system is:

- Exploring the solution space (jumps)
- Pruning ineffective genes (deletions)
- Starting to create variants (duplication)

**Next milestone:** Watch for the first viral mutation around generation 50. You should see stress spike and a transposition cascade.