

**Assessment of genetic homogeneity of somatic embryo- derived plants and seed-derived plants of a robusta coffee cultivar using molecular markers and functional genes sequencing**

**Manoj Kumar Mishra\*, Pavankumar Jingade, Arun Kumar C. Huded, Bychappa Muniswamy**

Tissue Culture & Biotechnology Division, Unit of Central Coffee Research Institute, Coffee Board, Manasagangothri, Mysore-570 006. India

**Journal Name:** Plant Cell, Tissue and Organ culture

**Corresponding author:** Manoj Kumar Mishra

**Email:** [manojmishra.m@gmail.com](mailto:manojmishra.m@gmail.com)

**Online resource 1:** Similarity coefficient of somatic embryo-derived plants and CxR mother plant generated using the combined SRAP-SCoT data.

	MP	SE-1	SE-2	SE-3	SE-4	SE-5	SE-6	SE-7	SE-8	SE-9	SE-10	SE-11	SE-12
MP	1.000												
SE-1	0.980	1.000											
SE-2	0.976	0.996	1.000										
SE-3	0.980	1.000	0.996	1.000									
SE-4	0.959	0.980	0.976	0.980	1.000								
SE-5	0.929	0.934	0.929	0.934	0.946	1.000							
SE-6	0.922	0.918	0.914	0.918	0.913	0.942	1.000						
SE-7	0.976	0.988	0.984	0.988	0.976	0.929	0.922	1.000					
SE-8	0.976	0.996	0.992	0.996	0.984	0.938	0.922	0.992	1.000				
SE-9	0.964	0.984	0.980	0.984	0.980	0.942	0.909	0.980	0.988	1.000			
SE-10	0.880	0.885	0.889	0.885	0.879	0.863	0.890	0.880	0.889	0.893	1.000		
SE-11	0.946	0.942	0.938	0.942	0.937	0.923	0.942	0.954	0.946	0.942	0.899	1.000	
SE-12	0.976	0.988	0.984	0.988	0.975	0.929	0.922	0.992	0.992	0.980	0.879	0.946	1.000

**Online resource 2:** Similarity coefficient of seed-derived plants and CxRmother plant generated using the combined SRAP-SCoT markers data.

	<b>MP</b>	<b>SD-1</b>	<b>SD-2</b>	<b>SD-3</b>	<b>SD-4</b>	<b>SD-5</b>	<b>SD-6</b>	<b>SD-7</b>	<b>SD-8</b>
<b>MP</b>	1.000								
<b>SD-1</b>	0.967	1.000							
<b>SD-2</b>	0.964	0.980	1.000						
<b>SD-3</b>	0.954	0.979	0.967	1.000					
<b>SD-4</b>	0.951	0.975	0.972	0.962	1.000				
<b>SD-5</b>	0.967	0.959	0.955	0.970	0.950	1.000			
<b>SD-6</b>	0.955	0.946	0.959	0.966	0.954	0.979	1.000		
<b>SD-7</b>	0.955	0.946	0.951	0.949	0.946	0.954	0.958	1.000	
<b>SD-8</b>	0.951	0.943	0.955	0.946	0.943	0.959	0.954	0.988	1.000

**Online resource 3:** The sequence size of the *matK*, partial NAC25, Zinc Finger Protein and UDP-glycosyltransferase genes isolated from seed-derived plants and somatic embryo-derived plant samples derived from the same CxR mother plant.

Sr. No.	CxR samples	Sequence Size (bp)			
		<i>matK</i>	PartialNAC25	Zinc Finger Protein	UDP-glycosyltransferase
1	<b>MP</b>	881	568	423	495
2	<b>SD-2</b>	881	568	422	495
3	<b>SD-5</b>	881	567	424	495
4	<b>SD-8</b>	879	568	421	495
5	<b>SE-1</b>	882	568	419	495
6	<b>SE-10</b>	881	566	424	496
7	<b>SE-11</b>	881	568	420	496
<b>Average size</b>		<b>880.9</b>	<b>567.6</b>	<b>421.9</b>	<b>495.3</b>

**Note:** MP-Mother Plant; SD-Seed-derived plantsand SE-Somatic embryo-derived plants

**Online resource 4:** The sequence size and nucleotide base contents of the *matK*, partial NAC25, Zinc Finger Protein and UDP-glycosyltransferase genes obtained from seed-derived plants and somatic embryo-derived along with CxR mother plant.

Sr. no.	Gene	Avg. Size (bp)	Nucleotide Content (%)	CxR Samples							Avg.
				MP	SD-2	SD-5	SD-8	SE-1	SE-10	SE-11	
1	<i>matK</i>	881	A	36.78	36.78	36.78	36.97	36.73	36.78	36.78	36.80
			C	15.32	15.32	15.32	15.36	15.42	15.32	15.32	15.34
			G	17.93	17.93	17.93	17.75	17.91	17.93	17.93	17.90
			T	29.97	29.97	29.97	29.92	29.93	29.97	29.97	29.96
			G+C	33.26	33.26	33.26	33.11	33.33	33.26	33.26	33.25
2	Partial NAC25	568	A	28.35	28.17	29.10	28.35	28.35	28.27	27.99	28.37
			C	17.43	17.25	17.46	17.43	17.43	17.31	17.25	17.37
			G	24.30	24.30	23.81	24.30	24.30	24.38	24.30	24.24
			T	29.93	30.28	29.63	29.93	29.93	30.04	30.46	30.03
			G+C	41.73	41.55	41.27	41.73	41.73	41.70	41.55	41.61
3	Zinc Finger Protein	422	A	25.53	25.12	25.00	25.18	25.78	26.18	25.71	25.50
			C	28.13	27.73	27.83	27.79	27.45	28.07	27.14	27.73
			G	17.02	17.06	16.75	17.34	18.38	17.45	16.67	17.24
			T	29.31	30.09	30.42	29.69	28.40	28.30	30.48	29.53
			G+C	45.15	44.79	44.58	45.13	45.82	45.52	43.81	44.97
4	UDP-glycosyltransferase	495	A	26.06	26.67	26.87	26.46	26.26	26.41	26.21	26.42
			C	23.84	23.43	23.43	23.43	23.64	23.59	23.19	23.51
			G	25.25	24.85	24.65	24.85	24.85	25.20	25.00	24.95
			T	24.85	25.05	25.05	25.25	25.25	24.80	25.00	25.04
			G+C	49.09	48.28	48.08	48.28	48.48	48.79	48.19	48.46

**Note:** MP-Mother Plant; SD-Seed-derived plantsand SE-Somatic embryo-derived plants

**Online resource 5:** Parsimony Informative sites (PIC) and Singleton variable sites of four genes.

Gene	Parameter	Sites	Positions	Total	Grand Total
<i>matK</i>	PIC	Two variants	-	-	-
		Three variants	-	-	
	Singleton Variable sites	Two variants	865, 878 and 879	3	3
		Three variants	-	-	
Partial NAC25	PIC	Two variants	5 and 556	2	2
		Three variants	-	-	
	Singleton Variable sites	Two variants	6, 204, 526, 556 and 565	5	5
		Three variants	-	-	
Zinc Finger Protein	PIC	Two variants	82, 92, 114, 166, 200, 251, 257, 265, 271, 272, 278, 280, 287, 296, 306, 311, 316, 319, 320, 321, 325, 365, 370, 377, 385 and 419	26	28
		Three variants	209 and 275	2	
	Singleton Variable sites	Two variants	2, 23, 83, 102, 162, 184, 202, 240, 248, 269, 277, 279, 294, 299, 301, 303, 308, 309, 312, 317, 318, 326, 331, 334, 340, 341, 342, 350, 375, 383, 407, 415, 422, 425, 426, 427 and 428	37	38
		Three variants	307	1	
UDP-glycosyltransferase	PIC	Two variants	95, 96, 189, 195, 353 and 433	6	7
		Three variants	363	1	
	Singleton Variable sites	Two variants	3, 4, 5, 56, 82, 99, 170, 233, 350 and 404	10	10
		Three variants	-	-	

**Online resource 6:** Maximum likelihood based substitution matrix of the *matK*, partial NAC25, Zinc Finger Protein and UDP-glycosyltransferase genes estimated using Tamura-Nei model.

<i>matK gene</i>					NAC25				Zinc Finger Protein				UDP-glycosyltransferase						
	A	T	C	G		A	T	C	G		A	T	C	G		A	T	C	G
<b>A</b>	-	9.29	4.78	<b>0</b>	<b>A</b>	-	10.19	5.94	<b>8.15</b>	<b>A</b>	-	4.75	4.42	<b>14.33</b>	<b>A</b>	-	6.68	6.25	<b>7.94</b>
<b>T</b>	11.45	-	<b>12.88</b>	5.52	<b>T</b>	9.69	-	<b>5.19</b>	8.31	<b>T</b>	4.03	-	<b>15.82</b>	2.77	<b>T</b>	7.03	-	<b>14.72</b>	6.62
<b>C</b>	11.45	<b>25.06</b>	-	5.52	<b>C</b>	9.69	<b>8.90</b>	-	8.31	<b>C</b>	4.03	<b>17.00</b>	-	2.77	<b>C</b>	7.03	<b>15.73</b>	-	6.62
<b>G</b>	<b>0</b>	9.29	4.78	-	<b>G</b>	<b>9.51</b>	10.19	5.94	-	<b>G</b>	<b>20.90</b>	4.75	4.42	-	<b>G</b>	<b>8.43</b>	6.68	6.25	-

**Note:** Highlighted percent is **Transition** while **Non-highlighted** percent is **Transversion**

**Online resource 7:** The number of synonymous and non-synonymous SNP from the four genes isolated from CxR samples.

Sr. No.	Gene	SD-2		SD-5		SD-8		SE-1		SE-10		SE-11	
		S	NS	S	NS	S	NS	S	NS	S	NS	S	NS
1	<i>matK</i>	-	-	-	-	-	-	-	4	-	-	-	-
2	Partial NAC25	-	2	1	7	-	2	-	-	-	2	1	2
3	Zinc Finger Protein	6	9	8	11	2	12	3	34	9	24	3	23
4	UDP-glycosyltransferase	2	4	-	13	-	5	1	4	1	3	1	7