Supplementary Tables

Table S1 List of populations taken under this study

Sl. No.	Population/Ethnicity/Tribe	State	Language Family
1	Bengali (Hindu)	Assam	Indo-Aryan
2	Bengali (Muslim)	Assam	Indo-Aryan
3	Dimasa	Assam/Nagaland	Tibeto-Burman
4	Karbi	Assam (hill areas)	Sino-Tibetan
5	Khasi	Meghalaya	Austro-Asiatic
6	Rongmei Naga	Nagaland	Tibeto-Burman
7	Tea_tribe	Assam/Nagaland	Indo-Aryan
8	Bengali (Christian)	Assam	Indo-Aryan
9	Karamkar/Goala/Tanti	Tripura	Indo-Aryan
10	Hmar	Mizoram	Kuki-Chin
11	Ahom	Assam	Indo-Aryan
12	Manipuri/Vishnupriya (Hindu)	Manipur	Tibeto-Burman

Table S2 List of primers used in the targeted re-sequencing experiment

Sl. No.	Chr.	GENE /SNP	Position (Start)	Position (Stop)	Ion_AmpliSeq_Fwd_Primer	Ion_AmpliSeq_Rev_Primer	PRIMER POOL
1	chr17	TP53(6)	7577397	7577615	TGTGATGAGAGGTGGATGGGTA	CCTCATCTTGGGCCTGTGTTAT	
2	chr17	TP53(1)	7572852	7573081	ATTGCAAGCAAGGGTTCAAAGAC	GCCTTAGGCCCTTCAAAGCATT	
3	chr17	TP53(8)	7578208	7578421	TCATAGGGCACCACCACACTAT	TACAAGCAGTCACAGCACATGA	
4	chr17	TP53(13)	7579287	7579488	CCAGGCATTGAAGTCTCATGGA	GTCCAGATGAAGCTCCCAGAATG	
5	chr11	HRAS(1)	533779	533936	GGTTCACCTGTACTGGTGGAT	TGTCCTCCTGCAGGATTCCTA	
6	chr2	CASP8(6)	202139542	202139748	AAAAAGACCTTATGTTGTCCTATGTGCT	CTTGTCAGTTTATCTGGCAGCAAAAA	
7	chr2	CASP8(14)	202151112	202151310	AGAAAGAACTACAGAACTATCTCTGAGCA	ACAAAACAAACAAAATAGCACCATCAATCA	
8	chr17	TP53(4)	7576782	7577010	CCCAATTGCAGGTAAAACAGTCA	GAGCACTAAGCGAGGTAAGCAA	
9	chr12	rs2051526	11907033	11907251	CAGAAGTCATCCTGTGACCGAT	CCTGTTTGTGATGAGTGTTTGATACTTTTT	
10	chr10	rs7084834	106403145	106403329	CATAAGCAATACCAAAGTGCAGATTTCTT	GACTTGGAAGTCTTCAATCACCCA	
11	chr16	rs243842	55527323	55527521	AAATGCTTCCCAGAATGACCTGA	GCTTTAAATGAGCTTGTGTAGTTCAGTTTG	
12	chr20	rs6139737	5453656	5453881	TAGAAGTATTTGGCACATTTGGGATGA	CCCAGACACTTCTCCTGGAAGA	
13	chr12	rs5744750	133253966	133254194	ACTTACAGGTCGTTCAACAAGGTC	GAGTACGATGTTCCCTACCACAT	
14	chr5	rs1457115	110567577	110567698	GTAAATTATAGCTAGACTTGTCCACATGCA	GCGAGAACATCCTTTTCTTTTCAATTACTT	
15	chr17	rs7501817	32229452	32229673	AGTACACATCTTCCCATTTAGCTAATGC	TGTCTTGCATTGACTGTTGGAGT	
16	chr10	rs6602275	7903632	7903851	ATTGTTTCTGATAAGAAGTCAGACGTCA	GTCTCAGAAGGAGATGATAGGGAAAAA	
17	chr18	rs2270954	51057185	51057411	CTGAAGCAGTTCAGAAGGAATAAGC	CCATGGCATGGTACAGGCTTAA	
18	chr4	rs1435218	54401830	54401968	GATACAGATGTGAATAAGATGTGGAAGGA	TGTGCCAACAGATGAATGAGAAGT	
19	chr22	rs2156921	23568939	23569149	GTGGCTTGCAAAGCTGAAGATATTC	CACCAATGGCAAATGCTCGATC	POOL 1
20	chr2	rs13026208	155252575	155252763	CTAAGAGAAAATCTGAAGTGTAAGCCCTT	TCATGGTCCTTGAAATATGAGAAGTTGTTT	
21	chr15	rs963425	35227517	35227709	GGGCATGAAATATTTTCAGAGTAAAAGCT	CTTTCCGTTTTCCTGTACACTACAAATAA	
22	chr15	rs10519967	35177396	35177618	CACTTGTCAACCCTCTGTCAGT	GGAATCCAGGAATCCACATTTTTATTTAGC	
23	chr11	rs1670661	21230560	21230778	CTAGATAGTAACCATTGTGAAGACCAGAT	AGGAACAAGATGGTTCCTAGAATGTAAA	
24	chr6	rs1395122	100865684	100865900	AGGTTATCAACAACCCATGCAGT	CTAGAAATGTTCTATGAAATGCACATGCA	
25	chr21	rs392686	46023556	46023780	GGCAACAGGTGGAACTTCAGAA	CATTTTGCTGCATTCAAGATTGTCAG	
26	chr13	rs2148229	93483469	93483664	AATCAAATTCCTGTTACCCTGTATTAGGTG	GGAGTTTACAATCTGAAAATGGGAGTAATG	
27	chr13	rs6491676	102907313	102907530	AAACAGCTAGGTGAAGTCTCTAAATCAC	CAGAATAAGTGGCCAAAGGAGTAATAGAAT	
28	chr10	rs10885701	117147511	117147697	ATTGTGTGATTAAGTGTTGTTACCACTTTG	CACACCACTCTTAGTACAAAACAGTAGATA	
29	chr18	rs11874922	50886063	50886278	GCCATGAAAAGGAAAGATGATTAATCTGTT	CACACAGATACAGGAAATGATATTATGCAG	
30	chr13	rs188014	50969823	50970043	TCATTGAAAGGCAGGTGATGATAGG	GTGCCCAGTGCATAATAATTTCTTAACATT	
31	chr8	rs12679182	121810050	121810247	CACCTCAGCCCTATAAAGAAAGCA	CGTGCTTTGTAGCAAACATAAGACATTTTT	
32	chr1	rs587168	23371457	23371681	GTTGTTACAGGGTATTGGGTACCT	ACTGGAACAAAAACATCCTTCAAGAGA	
33	chr6	rs1335024	102471825	102472002	AAGACTGTAGGTCTTTAATGCGCA	GTCAGTCATTCAAAATCATTACCAACACTT	
34	chr7	rs2237306	24757087	24757311	CAGCCTCGTGCCTTATACATAAGA	AGGATTTCAGTTGTTTCCATAACCTGT	
35	chr13	rs2314232	21087152	21087381	GTGCTAATACCCTGACTTGATCGT	CCCTGCTACCTCCTGACTGTAA	
36	chr12	rs7295342	23879459	23879657	AATAAAAAGTGGTTCATTTTGGACCAACAT	CTGACAGTCATTTTTCCCATGACTTTTC	
37	chr8	rs12056907	28239981	28240201	CTCACACAATCATAGACTTTACACTTTCCT	GGGTGGAAGACTTACAGTCTCTTTG	
38	chr6	rs1883930	52400788	52401007	CTGCCTAATGCAGAGTTAGCATTAGT	CCTAACCTTCTTGAAGGTTGTGACTTTATT	
39	chr11	rs2270994	48157746	48157977	ACAAGCACGTATGACAAAGCGA	GCAGAACCATTGCTGAGACCAA	

40	chr11	rs7396030	1083263	1083458	TGGCCAAGTGCTCCTGTTAC	GTGAGCCTGATGCCTCTCA	
41	chr14	rs10148795	33845140	33845364	TTCACCTAAGTTGCTGGGCTTT	CTTAGGATTAACATCCAAAGGCAGTACA	
42	chr20	rs7273624	62385387	62385571	CCCATGACAAGGATCTCCCAAA	GCCAGGAAGAGCTTGGGATTTC	
43	chr3	rs9849237	2675108	2675322	TCTGCTCCTGTACATGTTTAATAAACAGAG	TGGAGAGGATATTTTGGATTTAGCTGTTTT	
44	chr6	rs9485296	148871285	148871432	CCTCTTTTAGTCTAGTGTCTGGTTTTCTAG	TTCAATCAAGGGCTTGTGGTCA	
45	chr9	rs290974	93590331	93590541	AATCCACATTTACTGTGTCCTTTTTGC	GATCCTTTCCTACTCAGCTGTGAAG	
46	chr11	rs4753436	92898234	92898450	GTTTACAACAGCAATAATTGTAGCTGTGA	GCTCCCAGTTTATTTAATGCCTAACATTTG	
47	chr7	rs3757769	127540900	127541125	GCTGTCTTTGTACTGCTGCATCT	CCTTATCTGCTCCACAGAGCAA	
48	chr9	rs10961490	14370077	14370299	TGAACCATTCGCCATCTCCAAA	TGACCTTATGGCTCCATAGACCATTAATAT	
49	chr11	rs1784522	132578730	132578939	CTTTACACATACATTATGGCTACCAGTGT	AGAGAAGAGAAATGACAACTGGTGAAA	
50	chr3	rs11130760	60196411	60196627	TGTGCTGTTAATATGACAAATGCAATCG	TATCTTTGCTAGTGTGTCTTTTGGATAGAC	
51	chr20	rs6021247	50108880	50109107	CCCAGCTTCTAGAAAAGACACAGA	CCTTATCTTGGCTGGTGGTGTT	
52	chr2	rs2124437	33682601	33682823	ACAGCTTTGAATGAATGCATTAAGATTGTT	GAACAGGAAACCAGGATTCTCGT	
53	chr4	rs9307326	109882303	109882498	CACCTATCTTTTCTTAACAGGAATCAGAGT	TACTGTATCTCATTCCATATGCTAGTGTGT	
54	chr12	rs10491990	11939422	11939628	GTAGATTTAGTCCTCATAGTCTCACTGGAT	TGCTGGAAGAGTCTTTAGAAATCACC	
55	chr9	NOTCH1(2)	139396461	139396569	GCGACTCACCCGGAACTT	TTCCTAGGGAGCTCGCTCA	
56	chr9	NOTCH1(17)	139413188	139413405	GTTGACACACACGCAGTTGTAG	GCAGAGGAAGGAGTGATTTAGGC	
57	chr9	NOTCH1(14)	139412703	139412918	CAGTTGGAGCCCTCGTTACA	CTGAAAGGGCCTTTCTGGTCA	
58	chr9	NOTCH1(7)	139400069	139400296	CCGAAGCTGTAGTCCAGGAT	CGTGTGAGAATGACGCTCGTA	
59	chr9	NOTCH1(4)	139399336	139399533	TCACGCTTGAAGACCACGTT	CCTGTACGACCAGTACTGCA	
60	chr4	FAT1(6)	187530339	187530551	CGTAAATCAACCTTTTTCACTTACCGT	AAACAACAACAAAAAGAAACAGCTAAGCTA	POOL 1
61	chr9	NOTCH1(9)	139408932	139409122	GCTCCCAGGGTTTAGGACTGAT	TGTGTAACATCAACATCGATGAGTGT	
62	chr9	NOTCH1(11)	139412135	139412278	CGGAAGCAACCCACAGATGTT	CGAGATCGACGTCAACGAGT	
63	chr4	FAT1(1)	187517665	187517888	ACCTCACTGTTTATCTCAAGTGACAC	TCCAAGTGACTCAAGAAACAATCTGG	
64	chr17	rs2333990	78784783	78785008	TCTGGGAGCTGTTTGAAAGGTT	TGGGTATCTATGAAAAATTCACAGCACA	
65	chr7	rs826802	147034911	147035113	GCAGGGTGGCTCCAATTAGTAT	TGTGAGAACTAACTAACACAGTAGCCTA	
66	chr12	rs1821777	66606104	66606333	GAAGGAGGAGGCTTTCGAGTAG	TGAGCTCCTTGAGGACAGAACTAT	
67	chr4	FAT1(22)	187557638	187557867	CCAGCAGGTCACAAACACTTCA	CATAGCCACCGACAAGGATGAG	
68	chr4	FAT1(29)	187629849	187630071	GGTGTGTTGGGAGGAGCAAATT	GGTGACCTTCTCCAGCAGTTTA	
69	chr4	FAT1(25)	187627765	187627969	CAACGCCAGAGCCATCTCTAAT	GTATAATCTCACTGTGAGGGCCAAA	
70	chr2	CASP8(8)	202149665	202149891	GACCACAGTAACATGGACTGCTT	GGGATATATCTCGTTTGAGGTGATGA	
71	chr4	FAT1(11)	187538078	187538283	GTGATCATAAAACACATGTGAATTACACCA	ATGAATGGAAGGTATATGTGAAGAAACCTC	
72	chr4	FAT1(4)	187522402	187522543	TATGGTTTGAAAGTCTTCCACTAAGATGC	CACTGACTGGAAACAGCTACGT	
73	chr2	CASP8(1)	202131152	202131381	GCCAGCAAATGGTACTTTTCTTCC	AGCTCCTTCAGGAAGGACAGAT	
74	chr4	FAT1(31)	187630655	187630876	TCAATGTGTAGTGATCCTTCACTTCTCTA	GAACAGACTCCTCTGCAGTTTACA	
75	chr4	FAT1(17)	187541475	187541697	CACCTCAGTGTCCACTTTAACAAC	GGAATCCAATCAATGAGCCTTTGTTTTAT	
76	chr4	FAT1(9)	187534221	187534433	GTTTTGAAACAGAGGCTCCTAGGAT	GCAAGCTGTTTATCATCTTCTCGTCA	
77	chr4	FAT1(13)	187539774	187540004	CCAATGAGGCTCTCCTTTGTAGT	GATGCTGGAGGAAAAGTTGCTT	
78	chr2	CASP8(4)	202136137	202136357	AAACCATGCCATTAACTGGCTTTATG	CCAAGTCTCAGGACTAAATCTCAGTAAGA	
79	chr4	FAT1(20)	187542249	187542466	TGTTTCTTCATAGTCCAGACTTAGTACTGT	GGTTCTAGTTCACTTGCAGGATGA	
80	chr4	FAT1(15)	187540792	187541014	CCCAAACATCTGGTATGAGATTCCT	TGGAGTTATCAATGTCATAGCTCCTCT	
81	chr4	FAT1(27)	187628070	187628292	TGAGTTTATCCACATCGAAGTTTCCTT	GCATGAGCACTCCTTAAAGATTGAGG	

82	chr17	TP53(7)	7578086	7578296	GGGTTATAGGGAGGTCAAATAAGCA	GGCCTCTGATTCCTCACTGATTG	
83	chr17	TP53(5)	7576971	7577196	AAGGTGATAAAAGTGAATCTGAGGCAT	AATGGGACAGGTAGGACCTGAT	
84	chr9	NOTCH1(1)	139393576	139393702	GGCTCACCCAGGTCATCTAC	CCTGTTGCCCAGATCCTGATC	
85	chr9	NOTCH1(3)	139397546	139397753	CACACCCGTGGGTAGCAA	CGTCTACCTGGAGATTGACAACC	
86	chr17	TP53(3)	7573885	7574110	TCCTATGGCTTTCCAACCTAGGA	GTTGCTTTTGTACCGTCATAAAGTCA	
87	chr11	HRAS(2)	534218	534348	CGCCAGGCTCACCTCTATA	CAGGAGACCCTGTAGGAGGAC	
88	chr17	TP53(14)	7579403	7579594	GGTTTTCTGGGAAGGGACAGAAG	CTGACTGCTCTTTTCACCCATCT	
89	chr17	TP53(11)	7578384	7578551	CTGCTCACCATCGCTATCTGAG	TGTCTCCTTCCTCTTCCTACAGTAC	
90	chr2	CASP8(7)	202141468	202141690	GTGGTGCAATGGAAAGCAAGTC	ACTTTTCTCTTTTGGGTTCTACTGTACC	
91	chr9	NOTCH1(15)	139412943	139413166	CCAGAAAGGCCCTTTCAGGTTA	CTACAACTGCGTGTGTGTCAAC	
92	chr9	NOTCH1(8)	139401645	139401848	CCTTGGCCTCACACAGGAAAA	GGTGAACTGCTCTGAGGAGAT	
93	chr21	rs2254480	35186652	35186846	CTGCATCTTGACCAGTGGGTAA	CAGAATAACCACCTTTATCCTTTCAGACAT	
94	chr12	rs4883543	133201936	133202069	CTCGGCAAGACCTTCCTGT	CCTGAGCCTGATGGACTGTT	
95	chr13	rs2451078	20098259	20098368	GGGATCTAGAATTTTTCTGGTTAACTGA	ACTGATTTGCTTATTGGTTCACTTGTCATA	
96	chr20	rs595292	56820584	56820801	CCACTTTCCTATCAGAACTGACTTCAG	TTGTTCATTGTCATTAAAGAGAAGATGGGA	
97	chr6	rs1335022	102466294	102466507	CAATTTAGGCTCTGGTAAAAGAATTTCTCC	CAGATAGGTGGATATAGAGAACCACAAATT	
98	chr11	rs11601310	48085061	48085280	GAATGAATGAATGAAAGGTAATTGGCCTT	TAGCTCATGATTTCACAAAGCAAAAGG	
99	chr12	rs7963499	11917278	11917498	CCTCCTGATTTAGACAAAACATTGTGAG	CAAACGTGATCATCCCTGAAGTAAAGA	
100	chr2	rs3827538	111807200	111807425	GCTTGGCCAAATTCCACATCTC	CAGGTTATACCATTTGGAAATGCCTCT	
101	chr4	rs2306058	73942562	73942791	CCTCAACTGCCCATATTCACCA	CCTTTTTCTGTGTATGGGAATGCA	
102	chr7	rs3808081	127567795	127568016	GCCTACCTCACTGCTTCTGAAT	CTGGATTTTAGAGTTTCAGGTTTTCAGC	POOL 2
103	chr13	rs201803	50980074	50980295	GATGTTGGTGTTATAGACCGCCTAT	GCCCAACCTGCATGATCTAATCT	
104	chr7	rs7778413	127357584	127357814	TGCCTGCATCAGGATATTGTTGT	CCCAATGCTGCCAAAATTTCCA	
105	chr12	rs10771021	23875326	23875481	TGAATTGAGATTCTCCCTCAGGGTAA	GTCACAATTTGTGAAATGAATCCAAGGTTA	
106	chr16	rs2667552	78505818	78506007	ACCAAAATATTCCAGATAACGGTTATCAGT	GTTACTTTTTCCTGTTTTAACCTTGAGCTT	
107	chr5	rs7702840	147865772	147865992	GCTCATAGCCGCAGTTCTCATT	CCTGTCTTCTGGTTGCTACACT	
108	chr8	rs10090787	121810434	121810656	GCCTTGATTGTAGATTAGAACCACCT	TTGGTGCAAGTAGCATTATCTTCTTCT	
109	chr3	rs12637722	23605166	23605386	ACCTCTTTGTTTATCTGTATTGCATCCTT	CTTGTCAGAGGTAACAGGACAAAACT	
110	chr12	rs11047417	24555318	24555540	GTGACCTGTCCCTACCACTTTTT	TTTTGCATTCACAATTTTGCATTCTTTCTC	
111	chr6	rs3798502	100859071	100859294	CTGAGATGGAAGATAGGGTCAAATGG	ACCTGAGTGCTGAGGAGGAATTA	
112	chr17	rs4789378	74930263	74930489	CCCAAGACGCTACTGTTCACTG	CCAAGGCTTCATAATGCTGAACAG	
113	chr14	rs1957358	55222361	55222589	CAGCAGGTCACTATGCCATGAT	TCAGAACATGGCTTAGTTTCACAGT	
114	chr4	rs2194861	109889139	109889358	GGTGACCGACAGCACAGATAAA	CAGTATACATTGAACCAATTTCAGGCCTA	
115	chr11	rs1508608	92893698	92893895	GCTGTGGGAAAAGGGAAGGATTA	GAAATAATGTCAAAATTCTTTCCCGGTCTT	
116	chr15	rs1376784	35179531	35179753	CATTCTGGTGCTAGTTCATTTCATCAC	CTTTCTTCCCTAACTGTGTGTTTTCG	
117	chr19	rs10401422	47244460	47244619	GATGAGGGCAGTCCCAGAAATC	CCCTCACATCCTCATTAGGAGGA	
118	chr6	rs1335033	102489304	102489517	GGGTTAGCAGAGGTGGGTAAGA	TGGTTATATCTGCTTATACATGAGGACACT	
119	chr18	rs12604940	50972137	50972329	GAAGGAAAAGGCTGAAACATTTTATCATGT	AGGTCCACCCAATGTATTATATCCTATGTA	
120	chr11	rs948034	120677290	120677491	GGGATGAAAGAGACCCTAGTGATG	GCTTCCAGCATCGTTAGCTTCT	
121	chr10	rs10829607	129030197	129030420	ACCTAGGTAAAGGCTGTCTGCTA	TCTTCCAATACATATGACCACCAAGATG	
122	chr3	rs1911746	62100737	62100953	TCAACTGACTGTTGCAATTATACAGAGG	TGTGTGATTTCTAGTCATCTGCATTAATGT	
123	chr10	rs4748011	12941065	12941285	GGGAGTCAGGTGGCCTAATTAA	GTGAAAAATAGTAATTCACCATGGCTTTCA	
124	chr16	rs708278	54144816	54145035	TCTGTTTCCTTCTCTCTGATCCTCAT	CAAAATTAAATTAAGAGGCCAGCAAGAGTT	

125	chr1	rs12046563	43137165	43137379	CAGTTGTAGCTAGTGGCTACTGTAC	CTTTACATTAGACTGCCACCCAGAT	
126	chr8	rs4512367	68927525	68927737	TCAAATCGTTTTCACGATTTTGCGT	AGTGAGGAAGTAGTAACAATGCATCAAAAT	
127	chr13	rs200220	50960260	50960475	CATAAACACAGGTGCCAACCAA	GGAAGATCCACAGGAGGAAAAGAAATG	
128	chr22	rs139838	38268771	38268960	CAAGTAGCTAGGAACGCAGGTT	AAAAAGGGCAGACATTTTCACAAAA	
129	chr20	rs8708	10618466	10618680	CCAGGTCTTTGAGAACTCCAGAT	ATGAACATACAAATGCATCTTTGATGTGTT	
130	chr6	rs3799627	166967255	166967480	CTTTAAACAGACTTGGCTAATTTCTGCA	CCTGTAGGTGAGAGCTCCTTCT	
131	chr11	rs12418890	11893108	11893319	GAATTCTGGTTGGTGGGAACCTA	CAGCAATCTCCCTGAGTTAAAGGA	
132	chr9	rs2377527	78791087	78791288	TTTTTACATTTGGCCATTAATTCATCTGGA	AACAGTGTGGGTGGTATTGGTTT	
133	chr2	rs2353292	158158395	158158616	GGTATCCAAAGGGAAGAAGCTAAGC	AGGAAGCTGAAATAAAACAAAAACACTGG	
134	chr9	NOTCH1(13)	139412296	139412507	GCACTCGTTGACGTCGATCT	GATGAGAAGTCGATTCTGGCTTCA	
135	chr9	NOTCH1(10)	139409608	139409836	TGGTCCCGCTGCAATCTC	TTCAGGACCCAACTGCGAGAT	
136	chr9	NOTCH1(5)	139399683	139399896	CCTGAGCAGAGCCTTAGAACTG	GTGACTGCTCCCTCAACTTCAA	
137	chr4	FAT1(21)	187549772	187549995	TCTACTGCTCTGCAGAGTGTAGA	TTGCCACAGTATCGTTTCAATATACATGA	
138	chr13	rs7323929	97200668	97200886	TCTGTAAGAGCTCTGCTTACGTAAAC	CACAAGCCAATTCTTTATTTGATGACAGAT	
139	chr2	CASP8(11)	202149875	202150097	ACTGATTCAGAGGAGCAACCCTA	GGCTCTGAATATGAGTAGATAGTAGTGTGA	POOL 2
140	chr4	FAT1(26)	187627911	187628135	GGGTGCAGGTTCTCATTCACAT	AAGGAACCGTCATCATGTGGTT	
141	chr2	CASP8(5)	202137280	202137499	GCAGGTCCTTGGTTGGAGAAAT	CCGGCTGTCAGGTTGTTTCTAC	
142	chr4	FAT1(16)	187541252	187541480	GAACGATAACTTCCGCTGAAAAGG	GGTGTTTGTCAACCTTCCCTACT	
143	chr4	FAT1(7)	187532673	187532891	TTCAGTTCCAACAAGAATGTCCTCA	CGGAAGATTTTATACTCACTGATTGACTCT	
144	chr4	FAT1(24)	187584493	187584697	TGTATTGAAAAGAATCCTTGTGGATTTCCA	CCATTATTGGCTAACAGTCTTTGCAAC	
145	chr4	FAT1(12)	187538969	187539192	TCATCCTCACTCACAGTCCCTT	GGCCAAGTTATGTATAGCCTGGAT	
146	chr4	FAT1(5)	187522499	187522723	GAATATGTTCTGAGCCTCATGGTCA	GCTGGGCTTTTCATTAAGCAAATAGA	
147	chr4	FAT1(14)	187540709	187540873	GCTCGTAATCCAGGGTTCTGAGTA	GTGTTTGCTCAGCAGTCTTATGC	
148	chr4	FAT1(3)	187521161	187521390	GATACATCCACCGACTCTTCGATG	GCCCTGGAAGTGAATGGAAACT	
149	chr4	FAT1(18)	187541803	187542026	CGCAGAGTAGACATCCTGGGTA	GAGTAAAAGTCATCACAGTAAATGCTACAG	
150	chr4	FAT1(30)	187630490	187630686	GACTCTTGCGATACTGGTCCTT	TTCTCGGAGACTTTTGCTTTCTAAGAATAA	
151	chr2	CASP8(2)	202131298	202131521	AGTTCCTGAGCCTGGACTACAT	CCAGTCCCACAATGGGAGTTTC	
152	chr4	FAT1(10)	187535375	187535590	TGTGTCTGGATTTAGTTTGAATTTTTCTGC	TAGTCATTTCCCATGTTTTCTGTGGT	
153	chr4	FAT1(28)	187629250	187629472	CCACGCCTAGATCTCTGGGAAT	GGTACGTGACATACAGTATCGCA	
154	chr10	rs7085532	114859344	114859564	TCTTAGGGAAAAGTACAATGTAGGGCTA	GTCACTGCTCCCACTTCTCAAT	

Table S3 List of 115 significant SNPs associated with oral cancer obtained from "Discovery Set" samples

Sl. No.	CHR	SNP	Position	GENE	Type	P-value
1	1	rs587168	23371608	*AOF2 [KDM1A]*	*SNV*	< 0.001
2	1	rs12046563	43137280	*PPIH*	*SNV*	< 0.001
3	2	rs2124437	33682737	*RASGRP3*	*SNV*	< 0.001
4	2	rs3827538	111807303	*ACOXL*	*SNV*	< 0.001
5	2	rs13026208	155252653	*GALNT13*	*SNV*	< 0.001
6	2	rs2353292	158158514	*GALNT5*	*SNV*	< 0.001
7	3	rs9849237	2675189	*CNTN4*	*SNV*	< 0.001
8	3	rs12637722	23605325	*UBE2E2*	*SNV*	< 0.001
9	3	rs11130760	60196537	*FHIT*	*SNV*	< 0.001
10	3	rs1911746	62100840	*PTPRG*	*SNV*	< 0.001
11	4	rs1435218	54401899	*LNX1*	*SNV*	< 0.001
12	4	rs2306058	73942678	*ANKRD17 PROTEIN*	*SNV*	< 0.001
13	4	rs9307326	109882400	*COL25A1*	*SNV*	< 0.001
14	4	rs2194861	109889251	*COL25A1*	*SNV*	< 0.001
15	4	rs76442451	187522722	*FAT1*	*SNV*	< 0.001
16	4	rs2249916	187534363	*FAT1*	*SNV*	< 0.001
17	4	rs2249917	187534375	*FAT1*	*SNV*	< 0.001
18	4	rs1280100	187538133	*FAT1*	*SNV*	< 0.001
19	4	rs114279527	187628054	*FAT1*	*SNV*	< 0.001
20	4	rs555992573	187628276	*FAT1*	*SNV*	< 0.001
21	4	rs41278611	187630597	*FAT1*	*SNV*	< 0.001
22	5	rs1457115	110567598	*CAMK4*	*SNV*	< 0.001
23	5	rs7702840	147865882	*HTR4*	*SNV*	< 0.001
24	6	rs1883930	52400889	*TRAM2*	*SNV*	< 0.001
25	6	rs3798502	100859189	*SIM1*	*SNV*	< 0.001
26	6	rs1395122	100865806	*SIM1*	*SNV*	< 0.001
27	6	rs1335022	102466397	*GRIK2*	*SNV*	< 0.001
28	6	rs1335024	102471903	*GRIK2*	*SNV*	< 0.001
29	6	rs1335033	102489410	*GRIK2*	*SNV*	< 0.001
30	6	rs9485296	148871389	*SASH1*	*SNV*	< 0.001
31	6	rs3799627	166967389	*RPS6KA2*	*SNV*	< 0.001
32	7	rs2237306	24757202	*ICERE-1* or *DFNA5*	*SNV*	< 0.001
33	7	rs145303927	24757243	*DFNA5*	*SNV*	< 0.001
34	7	rs7778413	127357716	*SND1*	*SNV*	< 0.001
35	7	rs3757769	127541062	*SND1*	*SNV*	< 0.001
36	7	rs3808081	127567900	*SND1*	*SNV*	< 0.001
37	7	rs826802	147035067	*CNTNAP2*	*SNV*	< 0.001
38	8	rs12056907	28240092	*ZNF395*	*SNV*	< 0.001
39	8	rs4512367	68927592	*DEPDC2 [PREX2]*	*SNV*	< 0.001
40	8	rs12679182	121810147	*SNTB1*	*SNV*	< 0.001
41	8	rs10090787	121810549	*SNTB1*	*SNV*	< 0.001
42	9	rs10961490	14370172	*PTPRD / NFIB*	*SNV*	< 0.001
43	9	rs2377527	78791189	*PCSK5, DKFZp667I0310*	*SNV*	< 0.001
44	9	rs182361	93590359	*SYK*	*SNV*	< 0.001
45	9	rs290974	93590403	*SYK*	*SNV*	< 0.001
46	9	rs169724	93590499	*SYK*	*SNV*	< 0.001
47	10	rs6602275	7903733	*TAF3*	*SNV*	< 0.001
48	10	rs7084834	106403209	*SORCS*	*SNV*	< 0.001
49	10	rs7085532	114859463	*TCF4*	*SNV*	< 0.001
50	10	rs10885701	117147589	*ATRNL1*	*SNV*	< 0.001
51	10	rs10829607	129030343	*DOCK1*	*SNV*	< 0.001
52	11	rs7396030	1083364	*MUC2*	*SNV*	< 0.001
53	11	rs530072530	11893126	*USP47*	*SNV*	< 0.001
54	11	rs12418890	11893217	*USP47*	*SNV*	< 0.001
55	11	rs1670661	21230670	*NELL1, IDH3GL*	*SNV*	< 0.001
56	11	rs11601310	48085189	*PTPRJ*	*SNV*	< 0.001

57	11	rs2270994	48157869	*PTPRJ*	*SNV*	< 0.001
58	11	rs1508608	92893825	*SLC36A4*	*SNV*	< 0.001
59	11	rs4753436	92898342	*SLC36A4*	*SNV*	< 0.001
60	11	rs948034	120677398	*GRIK4*	*SNV*	< 0.001
61	11	rs1784522	132578834	*OPCML*	*SNV*	< 0.001
62	12	rs2051526	11907128	*ETV6*	*SNV*	< 0.001
63	12	rs7963499	11917397	*ETV6*	*SNV*	< 0.001
64	12	rs10491990	11939528	*ETV6*	*SNV*	< 0.001
65	12	rs10771021	23875372	*SOX-5*	*SNV*	< 0.001
66	12	rs7295342	23879554	*SOX-5*	*SNV*	< 0.001
67	12	rs11047417	24555441	*SOX-5*	*SNV*	< 0.001
68	12	rs1821777	66606228	*IRAK3*	*SNV*	< 0.001
69	12	rs4883543	133202004	*POLE*	*SNV*	< 0.001
70	12	rs5744750	133254083	*POLE*	*SNV*	< 0.001
71	13	rs2451078	20098289	*TPTE2*	*SNV*	< 0.001
72	13	rs4769700	20098320	*TPTE2*	*SNV*	< 0.001
73	13	rs2314232	21087335	*CRYL1*	*SNV*	< 0.001
74	13	rs200220	50960378	*XTP6 alias DLEU1*	*SNV*	< 0.001
75	13	rs188014	50969923	*XTP6 alias DLEU1*	*SNV*	< 0.001
76	13	rs201803	50980202	*XTP6 alias DLEU1*	*SNV*	< 0.001
77	13	rs2148229	93483546	*GPC5*	*SNV*	< 0.001
78	13	rs7323929	97200777	*HS6ST3*	*SNV*	< 0.001
79	13	rs112894494	102907383	*FGF14*	*SNV*	< 0.001
80	13	rs6491676	102907408	*FGF14*	*SNV*	< 0.001
81	14	rs10148795	33845322	*NPAS3*	*SNV*	< 0.001
82	14	rs1957358	55222475	*SAMD4A*	*SNV*	< 0.001
83	15	rs10519967	35177507	*AQR*	*SNV*	< 0.001
84	15	rs1376784	35179643	*AQR*	*SNV*	< 0.001
85	15	rs963425	35227613	*AQR*	*SNV*	< 0.001
86	15	rs2060810	35254596	*AQR*	*SNV*	< 0.001
87	16	rs708278	54144925	*FTO*	*SNV*	< 0.001
88	16	rs243842	55527422	*MMP2*	*SNV*	< 0.001
89	16	rs17638536	78505895	*WWOX*	*SNV*	< 0.001
90	16	rs2667552	78505906	*WWOX*	*SNV*	< 0.001
91	16	rs16947620	78505948	*WWOX*	*SNV*	< 0.001
92	16	rs16947623	78505992	*WWOX*	*SNV*	< 0.001
93	17	rs1625895	7578115	*TP53*	*SNV*	< 0.001
94	17	rs1042522	7579472	*TP53*	*SNV*	< 0.001
95	17	rs7501817	32229537	*ACCN1 alias ASIC2*	*SNV*	< 0.001
96	17	rs4789378	74930366	*MGAT5B*	*SNV*	< 0.001
97	**17**	**rs71910526**	**78784797**	***RPTOR***	***INDEL***	< 0.001
98	17	rs2333990	78784925	*RPTOR*	*SNV*	< 0.001
99	17	rs9894955	78784990	*RPTOR*	*SNV*	< 0.001
100	18	rs11874922	50886168	*DCC*	*SNV*	< 0.001
101	18	rs12604940	50972215	*DCC*	*SNV*	< 0.001
102	18	rs2270954	51057298	*DCC*	*SNV*	< 0.001
103	19	rs10401422	47244530	*STRN4*	*SNV*	< 0.001
104	20	rs6139737	5453856	*LOC149837 / LOC643406*	*SNV*	< 0.001
105	20	rs8708	10618574	*JAG1*	*SNV*	< 0.001
106	20	rs6021247	50108980	*NFATC2*	*SNV*	< 0.001
107	20	rs595292	56820702	*PPP4RL1 / [PP4R1]*	*SNV*	< 0.001
108	20	rs13041528	62385446	*ZBTB46*	*SNV*	< 0.001
109	20	rs7273624	62385479	*ZBTB46*	*SNV*	< 0.001
110	20	rs12481497	62385541	*ZBTB46*	*SNV*	< 0.001
111	20	rs4299387	62385558	*ZBTB46*	*SNV*	< 0.001
112	21	rs2254480	35186743	*ITSN1*	*SNV*	< 0.001
113	21	rs392686	46023668	*ORF REGION / TSPEAR*	*SNV*	< 0.001
114	22	rs2156921	23569063	*BCR*	*SNV*	< 0.001
115	22	rs139838	38268922	*EIF3S6IP [EIF3L]*	*SNV*	< 0.001

Table S4 List of SNPs distributed in four pools for Sequenom experiment

Sl. No.	SNPs in Pool 1	SNPs in Pool 2	SNPs in Pool 3	SNPs in Pool 4
1	rs10090787	rs10491990	rs10771021	rs10961490
2	rs10148795	rs10829607	rs11047417	rs12056907
3	rs10401422	rs10885701	rs112894494	rs12679182
4	rs1042522	rs12418890	rs114279527	rs1376784
5	rs10519967	rs12604940	rs12046563	rs169724
6	rs11130760	rs1280100	rs12481497	rs182361
7	rs11601310	rs1335024	rs13026208	rs188014
8	rs11874922	rs139838	rs1335033	rs1911746
9	rs12637722	rs1435218	rs1395122	rs1957358
10	rs13041528	rs1457115	rs16947620	rs243842
11	rs1335022	rs1625895	rs1883930	rs3757769
12	rs145303927	rs1670661	rs200220	rs392686
13	rs1508608	rs16947623	rs2060810	rs41278611
14	rs1784522	rs17638536	rs2148229	rs4769700
15	rs2194861	rs1821777	rs2156921	rs6139737
16	rs2249917	rs201803	rs2254480	rs7084834
17	rs2306058	rs2051526	rs2270954	rs7085532
18	rs2451078	rs2124437	rs2270994	rs71910526
19	rs2667552	rs2237306	rs2333990	rs7273624
20	rs290974	rs2249916	rs2353292	rs7963499
21	rs3808081	rs2314232	rs3798502	
22	rs3827538	rs2377527	rs5744750	
23	rs4512367	rs3799627	rs587168	
24	rs4753436	rs4299387	rs6602275	
25	rs4883543	rs4789378	rs7396030	
26	rs555992573	rs530072530	rs7778413	
27	rs595292	rs6021247	rs826802	
28	rs6491676	rs7295342	rs9485296	
29	rs708278	rs76442451	rs963425	
30	rs7323929	rs7702840	rs9894955	
31	rs7501817	rs9307326		
32	rs8708	rs948034		
33		rs9849237		

Table S5 Overall summary of the strongly associated SNPs obtained for **Case1** and **Case2** after stratified analysis

Sl. No.	CHR	SNP ID	BP	Genetic Region	MA	MAF_ Case	MAF_ Control	AA	CHISQ	P-value	OR	L95	U95
Tobacco related Oral Cancer/OSCC													
1	12	rs1821777	66606228	*IRAK3*	T	0.07	0.222	C	7.777	0.005	0.23	0.08	0.71
2	**9**	***rs182361***	93590359	***SYK***	A	0.33	0.13	C	7.74	0.005	**3.46**	1.42	8.43
3	**7**	rs2237306#	24757202	***DFNA5***	A	0.15	0.306	C	6.831	0.009	**0.33**	0.14	0.78
4	13	rs4769700	20098320	*TPTE2*	A	0.58	0.361	G	6.582	0.01	2.26	1.19	4.29
5	6	rs1883930	52400889	*TRAM2*	T	0.42	0.217	C	6.352	0.012	2.46	1.21	5.01
6	**9**	rs290974	93590403	***SYK***	A	0.33	0.135	G	5.85	0.016	**3.05**	1.23	7.53
7	17	rs2333990	78784925	*RPTOR*	A	0.02	0.151	G	5.575	0.018	0.12	0.02	0.93
8	16	rs16947623	78505992	*WWOX*	C	0	0.083	G	5.217	0.022	0	NA	NA
9	9	rs10961490	14370172	*NFIB*	C	0.05	0.176	T	4.91	0.027	0.25	0.07	0.91
10	16	rs16947620	78505948	*WWOX*	T	0	0.083	C	4.816	0.028	0	NA	NA
11	**9**	rs169724	93590499	***SYK***	C	0.38	0.208	T	4.104	0.043	**2.55**	1.03	6.32
12	**11**	***rs1670661***	21230670	***NELL1***	C	0.18	0.083	T	3.916	0.049	**2.71**	1.01	7.29
Non-tobacco related Oral Cancer/OSCC													
13	21	rs392686	46023668	*TSPEAR*	T	0.09	0.389	G	17.4	< 0.001	0.16	0.06	0.41
14	11	rs530072530	11893126	*USP47*	G	0.19	0.019	C	15.35	< 0.001	12.2	2.64	56.67
15	11	***rs1670661***	21230670	***NELL1***	C	0.25	0.083	T	8.986	0.003	**3.67**	1.51	8.9
16	4	rs555992573	187628276	*FAT1*	G	0.14	0.028	T	7.674	0.006	5.62	1.46	21.61
17	11	rs4753436	92898342	*SLC36A4*	G	0.39	0.202	A	7.095	0.008	2.53	1.27	5.07
18	20	rs6021247	50108980	*NFATC2*	G	0.58	0.38	A	6.389	0.012	2.24	1.19	4.21
19	11	rs1508608	92893825	*SLC36A4*	A	0.36	0.189	G	6.153	0.013	2.41	1.19	4.88
20	4	rs2306058	73942678	*ANKRD17*	T	0.42	0.24	C	6.106	0.014	2.31	1.18	4.51
21	2	rs13026208	155252653	*GALNT13*	C	0.05	0.17	T	5.57	0.018	0.24	0.07	0.85
22	8	rs12056907	28240092	*ZNF395*	A	0.34	0.185	C	5.475	0.019	2.31	1.14	4.68
23	13	rs6491676	102907408	*FGF14*	A	0.25	0.113	G	5.428	0.02	2.61	1.14	5.96
24	16	rs16947620	78505948	*WWOX*	T	0	0.066	C	4.408	0.036	0	0	nan
25	20	rs595292	56820702	*PPP4RL1*	G	0.25	0.125	T	4.334	0.037	2.33	1.04	5.25
26	6	rs1335024	102471903	*GRIK2*	G	0.42	0.269	A	4.304	0.038	1.99	1.03	3.82
27	6	rs1335022	102466397	*GRIK2*	T	0.42	0.269	C	4.192	0.041	1.98	1.03	3.83
28	9	***rs182361***	93590359	***SYK***	A	0.25	0.13	C	4.044	0.044	**2.24**	1.01	4.97
29	20	rs8708	10618574	*JAG1*	T	0.56	0.404	C	4.009	0.045	1.9	1.01	3.57

The protective SNP (OR<1.0). **Case1** Cancer patients **WITH** prolonged tobacco chewing habit; **Case2** Cancer patients **WITHOUT** prolonged tobacco chewing habit; **BP** Base pair (as per hg19 assembly); **MA** Minor Allele; **AA** Alternate (major) Allele; **MAF** Minor Allele Frequency; **CHISQ** Chi-square Value; **OR** Odds Ratio; **L95** Lower 95% CI; **U95** Upper 95% CI; P-value<0.05 considered as statistically significant; The highlighted SNP IDs are the final SNPs that came after validation & replication in **Sequenom iPLEX MassARRAY** platform. The ***bold & italicized*** SNP IDs were found to be common in both the types of case categories.