



US 20250263795A1

(19) **United States**

(12) **Patent Application Publication**  
Kennedy et al.

(10) **Pub. No.: US 2025/0263795 A1**  
(43) **Pub. Date:** **Aug. 21, 2025**

(54) **METHODS FOR CLASSIFICATION OF TISSUE SAMPLES AS POSITIVE OR NEGATIVE FOR CANCER**

(71) Applicant: **Veracyte, Inc.**, South San Francisco, CA (US)

(72) Inventors: **Giulia C. Kennedy**, San Francisco, CA (US); **Bonnie H. Anderson**, Half Moon Bay, CA (US); **Darya I. Chudova**, San Jose, CA (US); **Eric T. Wang**, Milpitas, CA (US); **Hui Wang**, San Bruno, CA (US); **Moraima Pagan**, San Francisco, CA (US); **Nusrat Rabbee**, South San Francisco, CA (US); **Jonathan I. Wilde**, Burlingame, CA (US)

(21) Appl. No.: **19/197,756**

(22) Filed: **May 2, 2025**

#### Related U.S. Application Data

(63) Continuation-in-part of application No. 18/823,253, filed on Sep. 3, 2024, now Pat. No. 12,297,503, which is a continuation of application No. 17/157,876, filed on Jan. 25, 2021, now Pat. No. 12,110,554, which is a continuation of application No. 14/153,219, filed on

Jan. 13, 2014, now Pat. No. 10,934,587, which is a continuation of application No. 13/318,751, filed on May 10, 2012, filed as application No. PCT/US2010/034140 on May 7, 2010, now Pat. No. 8,669,057.

(60) Provisional application No. 61/176,471, filed on May 7, 2009.

#### Publication Classification

(51) **Int. Cl.**  
*CI2Q 1/6883* (2018.01)  
*CI2Q 1/6886* (2018.01)  
(52) **U.S. Cl.**  
CPC ..... *CI2Q 1/6883* (2013.01); *CI2Q 1/6886* (2013.01); *CI2Q 2600/112* (2013.01); *CI2Q 2600/156* (2013.01); *CI2Q 2600/158* (2013.01)

#### ABSTRACT

The present invention relates to compositions, kits, and methods for molecular profiling and cancer diagnostics, including but not limited to genomic DNA markers associated with cancer. In particular, the present invention provides molecular profiles associated with thyroid cancer, methods of determining molecular profiles, and methods of analyzing results to provide a diagnosis.

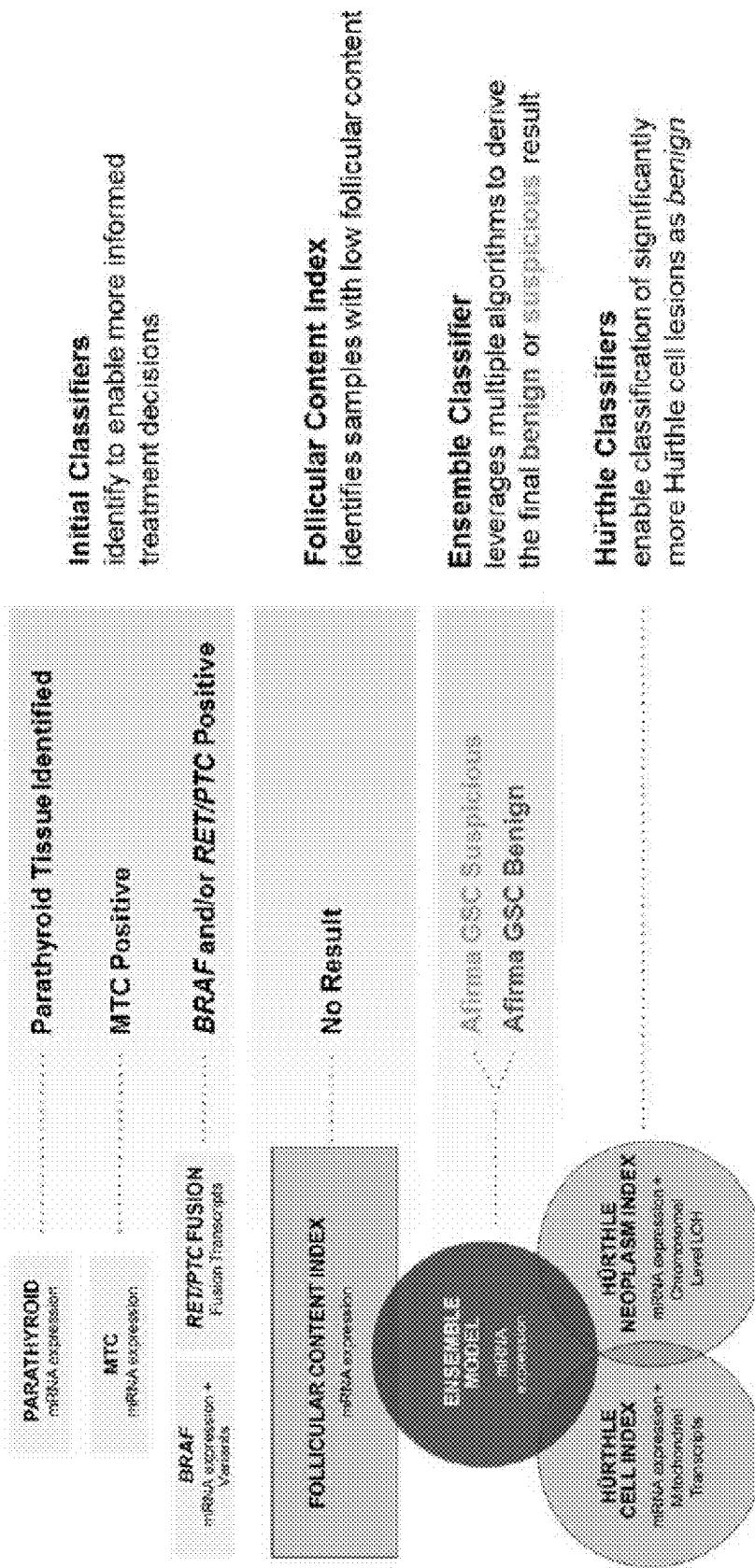


Fig. 1

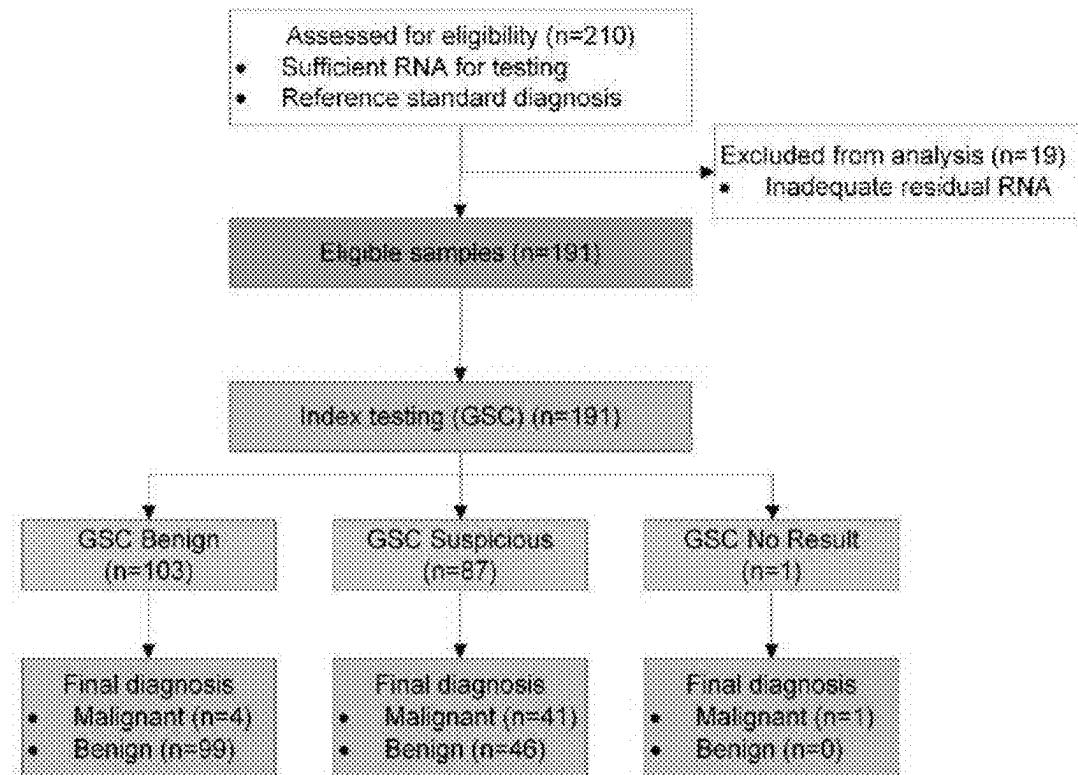


Fig. 2

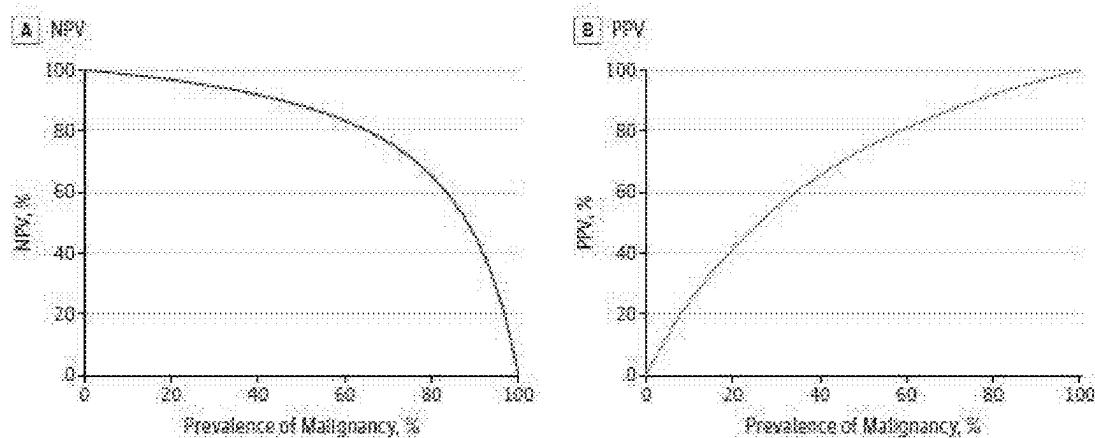


Fig. 3

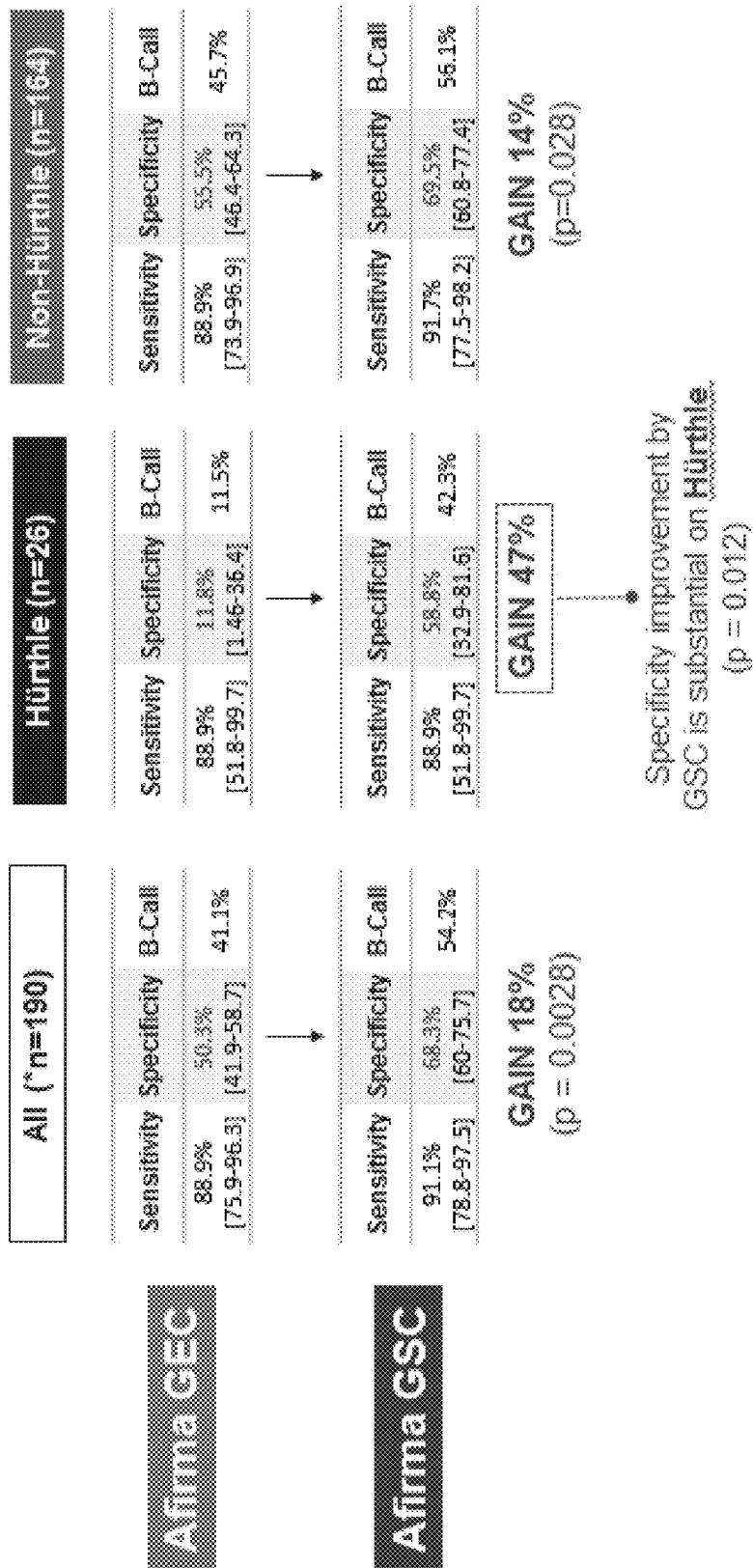


Fig. 4

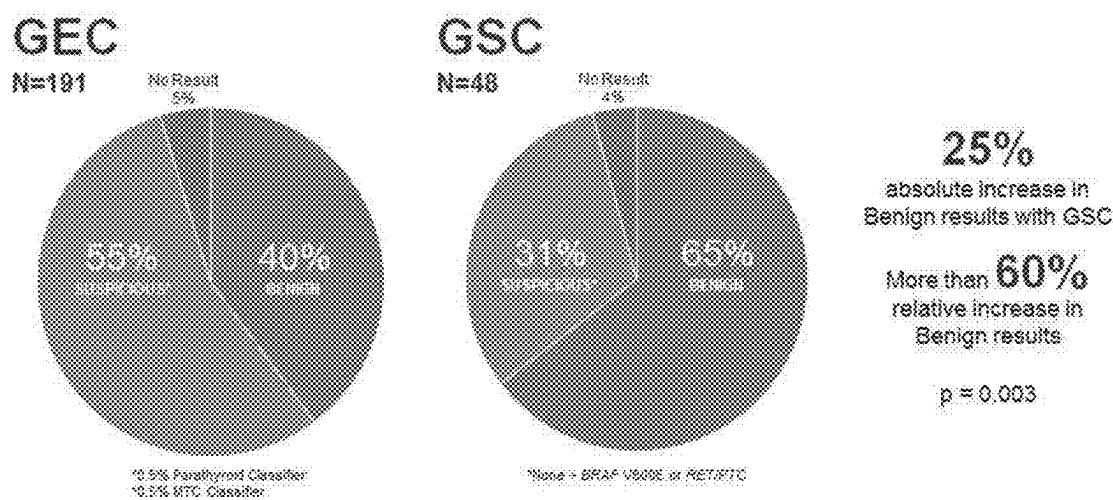


Fig. 5

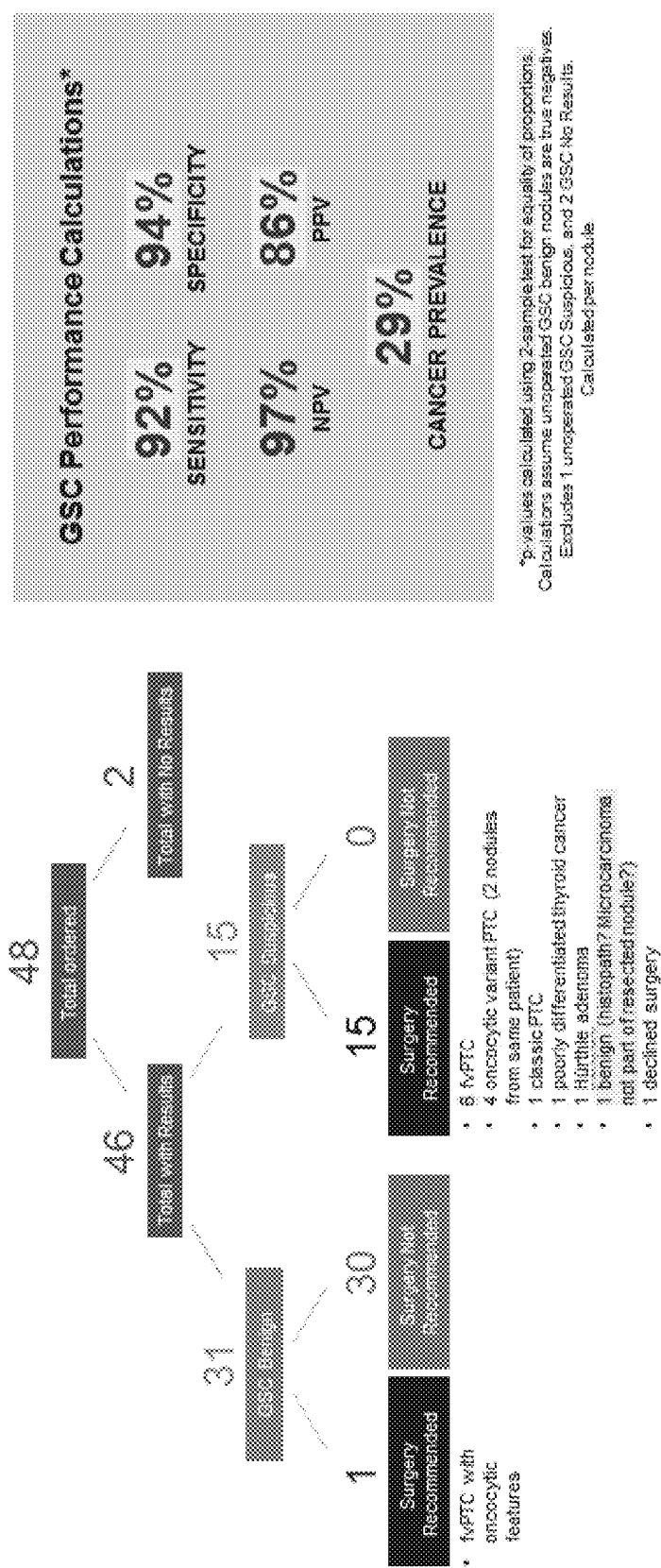


Fig. 6

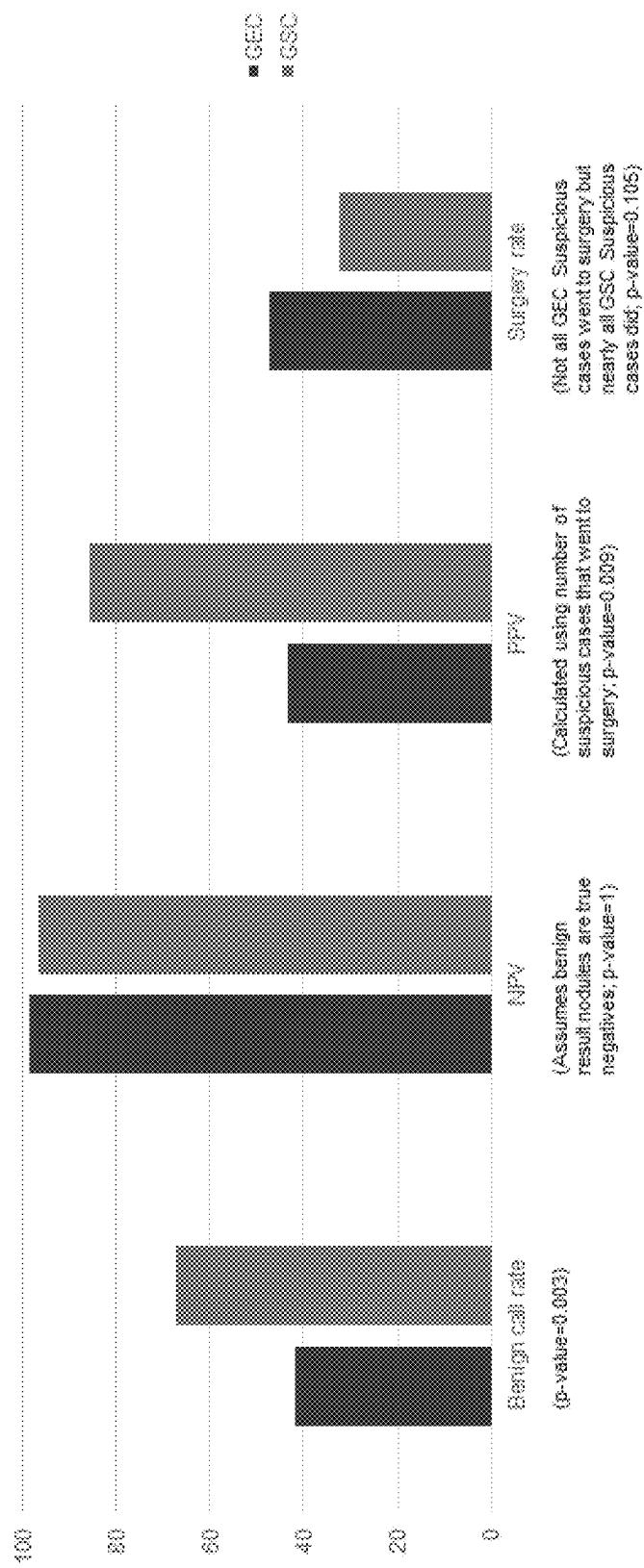


Fig. 7

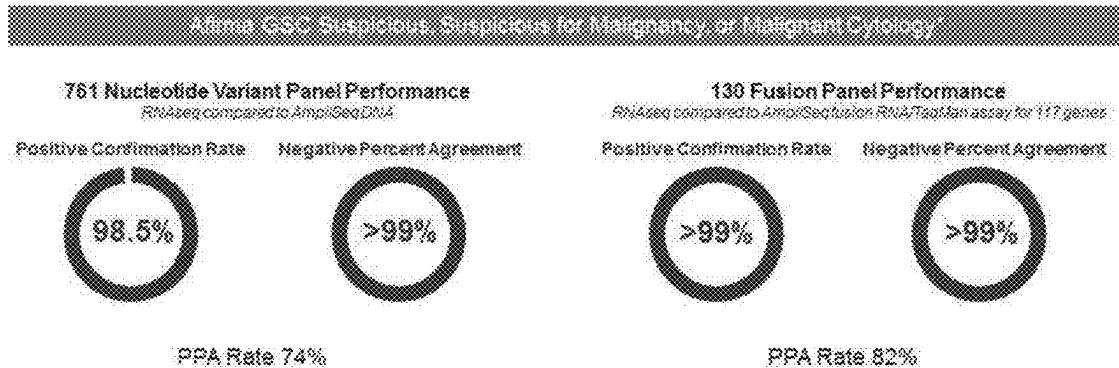


Fig. 8

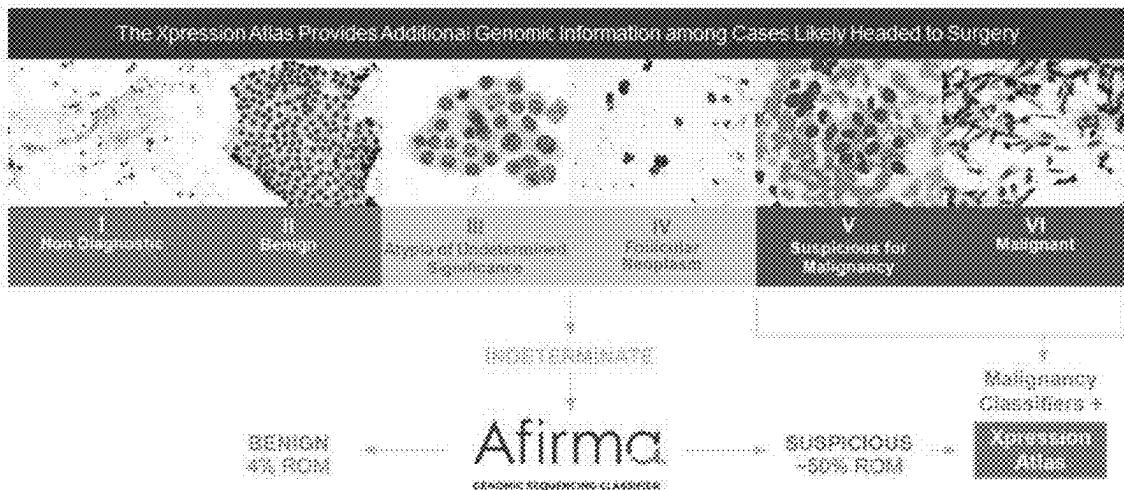


Fig. 9

Variant Detected	Result
<b>Sample Case 1: BRAF V600E</b>	
ROM	Highly associated with malignant nodules
Node type	Conventional PTC & infiltrative NFTC
Prognostics	Likely lymph node metastases*
Anticipated action	Variable extent of thyroid and LN surgery
<b>Sample Case 2: NRAS Q61R</b>	
ROM	Associated with benign and malignant nodules
Node type	MPTC, FTC, NFTC, RA*
Prognostics	Likely no lymph node metastases*
Anticipated action	Hemithyroidectomy alone
<b>Sample Case 3: None</b>	
ROM	Unchanged from GSC "Suspicious"
Node type	Variable
Prognostics	Likely no lymph node metastases*
Anticipated action	Hemithyroidectomy alone

- XA with variants detected may increase ROM of a GSC Result

**BUT**

- XA with NO variants detected does NOT reduce the ROM of a GSC "Suspicious" result

Fig. 10

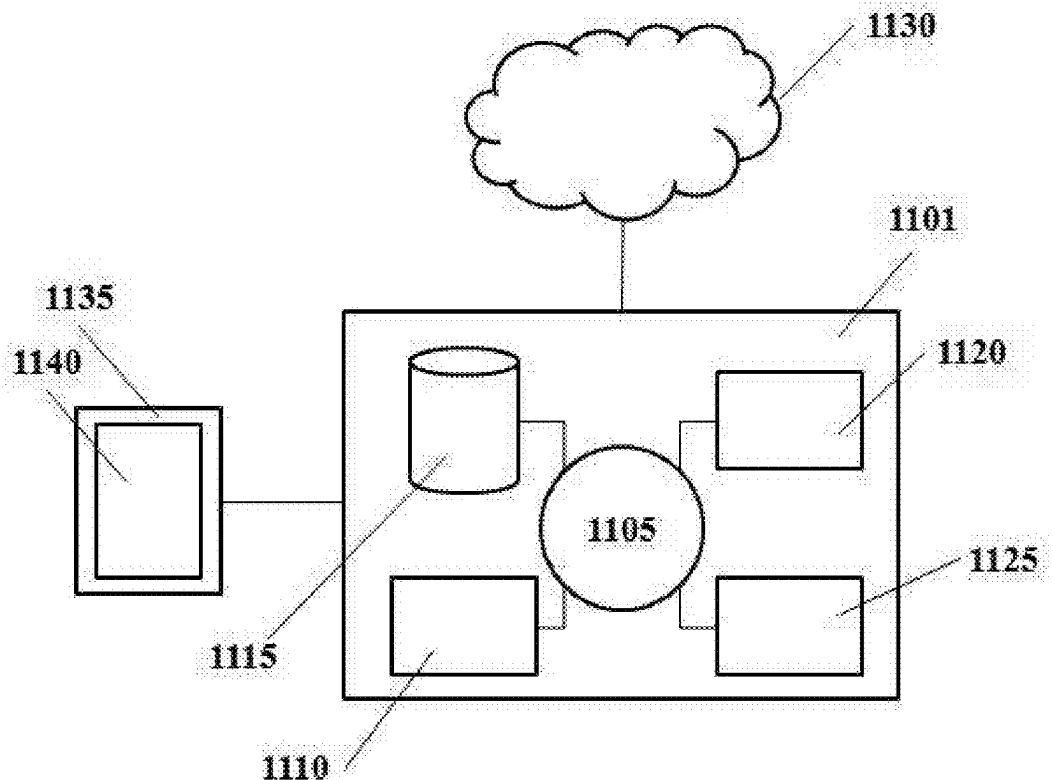


Fig. 11

TCID	GENE_na29
2337524	2337524
2357193	2357193
2383807	2383807
2406391	2406391
2472880	2472880
2489140	2489140
2494151	2494151
2506693	2506693
2641232	2641232
2649710	2649710
2652217	2652217
2832019	2832019
2908499	2908499
2908502	2908502
3092561	3092561
3107603	3107603
3139557	3139557
3153400	3153400
3184925	3184925
3253452	3253452
3321980	3321980
3398455	3398455
3511158	3511158
3619197	3619197
3749600	3749600
3777770	3777770
3881010	3881010
3893716	3893716
3895675	3895675
3958399	3958399
3289445	A1CF
3443348	A2M
2558150	AAK1
3697015	AARS
3070183	AASS
3719362	AATF
3218528	ABCA1
2756309	ABCA10
3768880	ABCA10
2756309	ABCA11P
3768880	ABCA11P
3768969	ABCA5
3768791	ABCA6
3815416	ABCA7
3768627	ABCA8
3768627	ABCA9
3768703	ABCA9
3060117	ABCB1
3060182	ABCB1
2459924	ABCB10
3060117	ABCB4
2599993	ABCB6
4012949	ABCB7
3690470	ABCC11
3914786	ABCC13
3260447	ABCC2
3726691	ABCC3
3446919	ABCC9
3450861	ABCD2
2777276	ABCG2
2550755	ABCG5
2474265	ABHD1
3535307	ABHD12B
3607447	ABHD2
2523689	ABI2
2686458	ABI3BP
3307939	ABLIM1
3739962	ABR
3368940	ABTB2
3754469	ACACA
3430959	ACACB
3432030	ACAD10
2695648	ACAD11
2641341	ACAD9
2597347	ACADL
3708306	ACADVL
2712040	ACAP2
3723317	ACBD4
3279058	ACBD7
3328069	ACCSL
3730601	ACE
4000605	ACE2
3164312	ACER2
3528994	ACIN1
3528895	ACIN1
3757433	ACLY
3013952	ACN9
3166477	ACO1
2394626	ACOT7
4002741	ACOT9
3771215	ACOX1
2759857	ACOX3
2645387	ACPL2
3981164	ACRC
3603199	ACSBG1
3726406	ACSF2
2796553	ACSL1
4017810	ACSL4
3651294	ACSM5
3901696	ACSS1
3424218	ACSS3
2459837	ACTA1
3299504	ACTA2
3569814	ACTN1
2386943	ACTN2
3861503	ACTN4
3832643	ACTN4
3304406	ACTR1A
2501697	ACTR3
2678901	ACTR8
2582562	ACVR1
3415148	ACVR1B
2509557	ACVR2A
3415109	ACVRL1
2623441	ACY1
3572263	ACYP1
3702382	ADAD2
3311832	ADAM12
2539821	ADAM17
3011492	ADAM22
3927446	ADAMTS1
3641633	ADAMTS17
3700158	ADAMTS18
2889916	ADAMTS2
3927480	ADAMTS5
2859601	ADAMTS6
2680046	ADAMTS9
2977471	ADAT2
2800711	ADCY2
3558168	ADCY4
3453252	ADCY6
3153716	ADCY8
3775906	ADCYAP1
2558736	ADD2
3263555	ADD3
2779271	ADH1A
2779231	ADH1A
2779199	ADH1B
2779271	ADH1B
2779231	ADH1B
2779199	ADH1B
2779271	ADH1C
2779231	ADH1C
2779124	ADH4
2387711	ADH5
2779095	ADH5
2779163	ADH6
3252170	ADK
3320123	ADM
3248824	ADO
2375596	ADORA1
3940001	ADORA2A
3711869	ADORA2B
2427981	ADORA3
3265047	ADRB1
4049835	ADRB3
3940631	ADRBK2
2464353	ADSS
3607232	AEN
2834957	AFAP1L1
3307851	AFAP1L2
2399743	AFARP1
2734784	AFF1
3554592	AFF1
2875555	AFF4
2533670	AGAP1
3031880	AGAP3
3606682	AGBL1
3372420	AGBL2
2949830	AGER
2530599	AGFG1
3083936	AGPAT5
3095815	AGPAT6
2734047	AGPAT9
3039791	AGR2
3039830	AGR3
3695726	AGRP

Fig. 12

TCID	GENE_na29
3212728	AGTPBP1
2647015	AGTR1
2889486	AGXT2L2
3023384	AHCYL2
2975385	AHI1
3375735	AHNAK
3581221	AHNAK2
2991233	AHR
3545466	AHSA1
2555277	AHSA2
3522398	AIDA
2381876	AIDA
2457573	AIDA
4042837	AIDA
3191877	AIF1L
2439554	AIM2
3226138	AK1
2340315	AK3L1
3728097	AKAP1
3749086	AKAP10
3487220	AKAP11
2931569	AKAP12
3606304	AKAP13
3184408	AKAP2
4008170	AKAP4
3540068	AKAP5
2925724	AKAP7
3853299	AKAP8
3012381	AKAP9
2969201	AKD2
3073981	AKR1B1
3274758	AKR1C1
3274758	AKR1C2
3233049	AKR1C3
3232944	AKR1CL2
2399743	AKR7A3
2399743	AKR7L
4009849	ALAS2
3838522	ALDH16A1
3209726	ALDH1A1
3611625	ALDH1A3
3169331	ALDH1B1
3714068	ALDH3A2
3337329	ALDH3B1
3379091	ALDH3B1
3337329	ALDH3B2
3379091	ALDH3B2
3571727	ALDH6A1
2873785	ALDH7A1
2975257	ALDH8A1
3750767	ALDOC
3990566	ALG11
3490504	ALG11
3965393	ALG12
2424148	ALG14
2708407	ALG3
2339682	ALG6
3383164	ALG8
3391149	ALG9
TCID	GENE_na29
2546409	ALK
3389878	ALKBH8
2488785	ALMS1
2488785	ALMS1P
3709417	ALOX15B
3244622	ALOX5
2532314	ALPI
2532294	ALPI
3605884	ALPK3
2532314	ALPP
2532294	ALPP
2532272	ALPP
2532294	ALPPL2
2532272	ALPPL2
2594773	ALS2CR12
2852742	AMACR
3371544	AMBRA1
3692856	AMFR
3393670	AMICA1
3452478	AMIGO2
4017961	AMMECR1
3410322	AMN1
4018454	AMOT
2696309	AMOTL2
3046739	AMPH
2571075	ANAPC1
2788143	ANAPC10
3527597	ANG
3572782	ANGEL1
3148463	ANGPT1
3122489	ANGPT2
2445982	ANGPTL1
3819474	ANGPTL4
3388517	ANGPTL5
3850020	ANGPTL6
2740067	ANK2
3290875	ANK3
3598199	ANKDD1A
3727787	ANKFN1
3741997	ANKFY1
2849469	ANKH
3349535	ANKK1
3525679	ANKRD10
3778252	ANKRD12
3431376	ANKRD13A
3716151	ANKRD13B
2773023	ANKRD17
3205834	ANKRD18A
3205834	ANKRD18B
2565559	ANKRD23
3282117	ANKRD26
2356273	ANKRD35
2564816	ANKRD36
2565935	ANKRD36
2564816	ANKRD36B
2565935	ANKRD36B
2754673	ANKRD37
3146661	ANKRD46
3876142	ANKRD5
TCID	GENE_na29
2857488	ANKRD55
2916825	ANKRD6
2904329	ANKS1A
3467351	ANKS1B
3678279	ANKS3
3217361	ANKS6
3441542	ANO2
3428333	ANO4
3323748	ANO5
2487082	ANTXR1
2774971	ANTXR2
3286975	ANUBL1
3174816	ANXA1
3627248	ANXA2
3167110	ANXA2
3627248	ANXA2P1
2790109	ANXA2P1
3167110	ANXA2P2
3249171	ANXA2P3
2732844	ANXA3
2487412	ANXA4
2784027	ANXA5
2881747	ANXA6
2358591	ANXA9
3046062	AOAH
3722195	AOC3
2522247	AOX1
3754677	AP1GBP1
3016177	AP1S1
2601287	AP1S3
3718682	AP2B1
2655476	AP2M1
2863730	AP3B1
3636216	AP3B2
3295032	AP3M1
3096007	AP3M2
2455418	AP3S1
3638665	AP3S2
2428855	AP4B1
3593770	AP4E1
2766893	APBB2
2878368	APBB3
2824198	APC
2622196	APEH
2434341	APH1A
3597521	APH1B
3369249	APIP
2486851	APLF
3356115	APLP2
2361584	APOA1BP
3945572	APOBEC3C
3945651	APOBEC3F
3945651	APOBEC3G
3835891	APOC1
4054204	APOD
3944404	APOL1
3959350	APOL3
2902531	APOM
4002809	APOO

Fig. 12 Continued

TCID	GENE_na29	TCID	GENE_na29	TCID	GENE_na29
3983105	APOOL	2531779	ARMC9	2695295	ATP2C1
2833924	APOOL	3984945	ARMCX3	3671727	ATP2C2
2625606	APPL1	4015838	ARMCX6	2351817	ATP5F1
3469319	APPL2	3321150	ARNTL	2714200	ATP5I
3203311	APTX	3409127	ARNTL2	3916576	ATP5J
3802396	AQP4	3471198	ARPC3	3535125	ATP5S
3595594	AQPF9	3188993	ARPC5L	3996381	ATP6AP1
3617757	AQR	2616596	ARPP-21	3105749	ATP6V0D2
3381241	ARAP1	2866704	ARRDC3	2841284	ATP6V0E1
2765590	ARAP2	3768474	ARSG	2636589	ATP6V1A
2731542	AREG	2782694	ARSJ	3088544	ATP6V1B2
3453370	ARF3	2731928	ART3	2469529	ATP6V1C2
3962587	ARFGAP3	3445723	ART4	2949038	ATP6V1G2
3139035	ARFGEF1	3359881	ART5	3982423	ATP7A
3888055	ARFGEF2	3261923	AS3MT	2767378	ATP8A1
2747893	ARFIP1	3395464	ASAM	3482274	ATP8A2
2925841	ARG1	3153428	ASAP1	3809826	ATP8B1
3524618	ARGLU1	2468811	ASAP2	3712835	ATPAF2
2746693	ARHGAP10	2553262	ASB3	3617920	ATPB4D
3283920	ARHGAP12	3191074	ASB6	2698844	ATR
2508611	ARHGAP15	2966636	ASCC3	3874313	ATRN
2973694	ARHGAP18	3429008	ASCL1	4013224	ATRX
3302187	ARHGAP19	3743371	ASGR1	3948754	ATXN10
3390641	ARHGAP20	2437417	ASH1L	3471588	ATXN2
2734421	ARHGAP24	3094447	ASH2L	3576889	ATXN3
2833286	ARHGAP26	3966597	ASMT	2627390	ATXN7
3777263	ARHGAP28	3137530	ASPH	2456849	AURKA
2423829	ARHGAP29	3214845	ASPN	2456849	AURKAPS1
3531479	ARHGAP5	2695295	ASTE1	3006572	AUTS2
3999568	ARHGAP6	3881874	ASXL1	2996033	AVL9
3948259	ARHGAP8	2544925	ASXL2	3459722	AVPR1A
3445786	ARHGDIB	3069399	ASZ1	3063589	AZGP1
3642747	ARHGDIG	3299255	ATAD1	2667181	AZI2
3352503	ARHGEF12	3151534	ATAD2	3147591	AZIN1
2437801	ARHGEF2	3716893	ATAD5	2386418	B3GALNT2
2677723	ARHGEF3	3310413	ATE1	2449104	B3GALT2
2505833	ARHGEF4	2588066	ATF2	3375935	B3GAT3
3029646	ARHGEF5	2379132	ATF3	3824596	B3GNT3
3029646	ARHGEF5L	2949622	ATF6B	3475511	B3GNT4
3501661	ARHGEF7	3647827	ATF7IP2	3418394	B4GALNT1
3537884	ARID4A	2818212	ATG10	3400236	B4GALNT3
2461786	ARID4B	3769779	ATG12	2440664	B4GALT3
3600744	ARIH1	3578278	ATG2B	3908963	B4GALT5
2621827	ARIH2	2339511	ATG4C	3803120	B4GALT6
2842530	ARL10	2526759	ATIC	3110217	BAALC
2632453	ARL13B	2548776	ATL2	3393257	BACE1
3767169	ARL17P1	3347658	ATM	3921933	BACE2
3334783	ARL2	3390143	ATM	3981120	BAG1
3304475	ARL3	3670668	ATMIN	2911372	BAG2
2931391	ARL4A	2726072	ATP10D	3267314	BAG3
2990464	ARL4A	2654855	ATP11B	3094494	BAG4
2724853	ARL4A	2711225	ATP13A4	3924929	BAGE
3683037	ARL6IP1	2711205	ATP13A4	3924929	BAGE2
2628682	ARL6IP5	2353477	ATP1A1	3924929	BAGE3
2451139	ARL8A	2645764	ATP1B3	3924929	BAGE4
2608765	ARL8B	3431483	ATP2A2	3924929	BAGE5
3017080	ARMC10	3464983	ATP2B1	3589947	BAHD1
3238702	ARMC3	2375706	ATP2B4	2912416	BAI3
2644461	ARMC8	2642325	ATP2C1	3737697	BAIAP2

Fig. 12 Continued

TCID	GENE_na29
3240452	BAMBI
2737596	BANK1
2949038	BAT1
2367154	BAT2D1
2949256	BAT5
3838067	BAX
3560711	BAZ1A
3056044	BAZ1B
3457947	BAZ2A
2583014	BAZ2B
3324453	BBOX1
3600960	BBS4
2996321	BBS9
2634965	BBX
3835777	BCAM
3018535	BCAP29
4026669	BCAP31
3729569	BCAS3
3311775	BCCIP
3269662	BCCIP
3454147	BCDN3D
2420808	BCL10
2554975	BCL11A
3811339	BCL2
3635198	BCL2A1
3902489	BCL2L1
2500275	BCL2L11
3936256	BCL2L13
3405207	BCL2L14
3386737	BCL7B
2356818	BCL9
3348940	BCO2
2814527	BDP1
3758157	BECN1
2411799	BEND5
4016308	BEX1
4016308	BEX2
3649245	BFAR
3899111	BFSP1
2608725	BHLHE40
3448088	BHLHE41
2817251	BHMT
2817251	BHMT2
3951927	BID
3346548	BIRC3
3772187	BIRC5
2476219	BIRC6
3893250	BIRC7
3499585	BIVM
3608298	BLM
3751830	BLMH
3301713	BLNK
3303392	BLOC1S2
2366490	BLZF1
3619229	BMF
3238491	BMI1
3089215	BMP1
2774817	BMP2K
2958172	BMP5

TCID	GENE_na29
2893895	BMP6
2331558	BMP8A
2331511	BMP8A
2331511	BMP8B
3256074	BMPR1A
2523213	BMPR2
3938817	BMS1
2841528	BNIP1
3314040	BNIP3
2761285	BOD1L
2536625	BOK
3683050	BOLA2
2593838	BOLL
3732448	BPTF
2877257	BRD8
2798915	BRD9
2442858	BRP44
3842059	BRSK1
3267678	BRWD2
4013730	BRWD3
3333595	BSCL2
3854417	BST2
3636470	BTBD1
3363645	BTBD10
3430462	BTBD11
3554592	BTBD6
3577277	BTBD7
2952497	BTBD9
2773545	BTC
2553771	BTF3
2375664	BTG2
3926080	BTG3
2899372	BTN3A1
2899298	BTN3A1
2899372	BTN3A2
2899298	BTN3A2
2899372	BTN3A3
2899298	BTN3A3
2844859	BTNL8
3261165	BTRC
3014742	BUD31
3764289	BZRAP1
2690012	BZW1
2522439	BZW1
2690012	BZW1L1
2522439	BZW1L1
3286776	C10orf10
3279089	C10orf111
3307795	C10orf118
3309755	C10orf119
3259087	C10orf129
3259253	C10orf131
3231846	C10orf139
3280787	C10orf140
3275132	C10orf18
3244539	C10orf25
3261886	C10orf26
3260099	C10orf28
3261923	C10orf32

TCID	GENE_na29
3309124	C10orf46
3294959	C10orf55
3254337	C10orf57
3254488	C10orf58
3241601	C10orf68
3288518	C10orf72
3305198	C10orf79
3314720	C10orf92
3314720	C10orf93
3279410	C10orf97
3334847	C11orf2
3334446	C11orf20
3358112	C11orf35
3356417	C11orf44
3325052	C11orf46
3375999	C11orf48
3329537	C11orf49
3380996	C11orf51
3396144	C11orf61
3353441	C11orf63
3390143	C11orf65
2708855	C11orf72
3343252	C11orf73
3327166	C11orf74
3386737	C11orf75
3336486	C11orf80
3368707	C11orf91
3465188	C12orf12
3430389	C12orf23
3424379	C12orf26
3421824	C12orf28
3464622	C12orf29
3471819	C12orf30
3443434	C12orf33
3410384	C12orf35
3407849	C12orf39
3441215	C12orf4
3411234	C12orf40
3453177	C12orf41
3473331	C12orf49
3401756	C12orf5
3442249	C12orf53
3403077	C12orf57
3427282	C12orf63
3410322	C12orf72
3429857	C12orf75
3471073	C12orf76
3486956	C13orf15
3480681	C13orf3
3484117	C13orf33
3485740	C13orf36
3563372	C14orf104
3528994	C14orf119
3550328	C14orf129
3550234	C14orf132
3563687	C14orf138
3573933	C14orf145
3573994	C14orf145
3566949	C14orf149

Fig. 12 Continued

TCID	GENE_na29
3543619	C14orf169
3569778	C14orf181
3571727	C14orf45
3543979	C14orf45
3540155	C14orf50
3554523	C14orf79
3554868	C14orf80
3569285	C14orf83
3633191	C15orf17
3623472	C15orf33
3629761	C15orf44
3641597	C15orf51
3619479	C15orf57
3643114	C16orf14
3649714	C16orf45
3701433	C16orf46
3695819	C16orf48
3655708	C16orf53
3644191	C16orf73
3693511	C16orf80
3678542	C16orf89
3774975	C17orf101
3744300	C17orf44
3712098	C17orf45
3710277	C17orf48
3708201	C17orf49
3724591	C17orf57
3744254	C17orf59
3766621	C17orf72
3769969	C17orf80
3735752	C17orf86
3742627	C17orf87
3740664	C17orf91
3803500	C18orf34
3791341	C18orf49
3793588	C18orf55
3795850	C18orf56
3798829	C18orf58
3828162	C19orf2
3817651	C19orf30
3832280	C19orf33
3868330	C19orf41
3822347	C19orf53
3862785	C19orf54
3864597	C19orf61
3839400	C19orf63
3850020	C19orf66
3837664	C19orf68
2557549	C1D
4019967	C1GALT1C1
2406735	C1orf102
2443305	C1orf114
2381249	C1orf115
2453065	C1orf116
2369609	C1orf125
2407128	C1orf149
2431886	C1orf152
2391005	C1orf159
2320392	C1orf187
TCID	GENE_na29
2436576	C1orf189
2371547	C1orf21
2406412	C1orf216
2369325	C1orf220
2364016	C1orf226
2371738	C1orf26
2436576	C1orf43
2369325	C1orf49
2358136	C1orf51
2381876	C1orf58
2402111	C1orf63
2389718	C1orf71
2359736	C1orf77
2336963	C1orf83
2438093	C1orf85
3236786	C1QL3
2852742	C1QTNF3
2719440	C1QTNF7
3403168	C1S
2902844	C2
3876084	C20orf103
3890109	C20orf108
3893072	C20orf11
3902764	C20orf112
3899495	C20orf12
3904797	C20orf132
3746881	C20orf191
3891723	C20orf197
3895679	C20orf27
3901665	C20orf3
3896174	C20orf30
3877221	C20orf7
3900091	C20orf74
3918104	C21orf63
3926138	C21orf91
3958045	C22orf30
3965102	C22orf34
3955357	C22orf36
3952703	C22orf39
3954764	C22orf43
3963676	C22orf9
3933331	C2CD2
3381702	C2CD3
2525682	C2orf21
2506335	C2orf27A
2506335	C2orf27B
2496628	C2orf29
2482230	C2orf30
2473831	C2orf39
2498274	C2orf40
2477372	C2orf56
2594313	C2orf60
2566383	C2orf64
2560317	C2orf65
2597273	C2orf67
2525852	C2orf67
2559494	C2orf7
2482440	C2orf73
2556017	C2orf86
TCID	GENE_na29
2520069	C2orf88
3848039	C3
2662581	C3orf10
2627080	C3orf14
2688882	C3orf17
2612012	C3orf20
2663083	C3orf31
2641577	C3orf37
2619521	C3orf41
2610417	C3orf42
2700780	C3orf44
2627368	C3orf49
2636062	C3orf52
2674168	C3orf62
2677653	C3orf63
2681114	C3orf64
2902958	C4A
2902958	C4B
2758076	C4orf10
2734352	C4orf12
2770427	C4orf14
2791419	C4orf18
2765865	C4orf19
2717757	C4orf23
2743029	C4orf29
2739714	C4orf32
2766492	C4orf34
2753994	C4orf41
2730303	C4orf7
2715440	C4orf8
2875929	C5orf15
2805176	C5orf22
2829542	C5orf24
2855578	C5orf28
2831519	C5orf32
2853388	C5orf33
2855614	C5orf34
2882834	C5orf4
2853642	C5orf42
2858793	C5orf43
3724591	C5orf45
2844479	C5orf45
2828564	C5orf56
2854915	C6
2951057	C6orf1
2941972	C6orf105
2951221	C6orf106
2927967	C6orf115
2986084	C6orf120
2901841	C6orf134
2956217	C6orf138
2909772	C6orf141
2960774	C6orf147
2960399	C6orf155
2916307	C6orf165
2966193	C6orf168
2973232	C6orf174
2984275	C6orf176
2969350	C6orf186

Fig. 12 Continued

TCID	GENE_na29
2974671	C6orf192
2939593	C6orf201
2902633	C6orf26
2985342	C6orf54
2945677	C6orf62
2930753	C6orf72
2931700	C6orf97
3060051	C7orf23
3041409	C7orf30
3005956	C7orf42
3048134	C7orf44
3022465	C7orf54
3060450	C7orf62
3011911	C7orf63
2995491	C7orf67
3037304	C7orf70
3194969	C8G
3085933	C8orf12
3124344	C8orf16
3107234	C8orf39
3095313	C8orf4
3096271	C8orf40
3130823	C8orf41
3121198	C8orf42
3101802	C8orf44
3101765	C8orf44
3086181	C8orf49
3151809	C8orf54
3119236	C8orf55
3089569	C8orf58
3118651	C8orf60
3099390	C8orf71
3086809	C8orf79
3086774	C8orf79
3158812	C8orf82
3140478	C8orf84
3180717	C9orf102
3226709	C9orf114
2549092	C9orf126
3225224	C9orf126
3179872	C9orf129
3167684	C9orf131
3195083	C9orf142
3162529	C9orf150
3216931	C9orf156
3223903	C9orf31
3213847	C9orf33
3185618	C9orf43
3037304	C9orf51
3173974	C9orf61
3197231	C9orf68
3193018	C9orf77
3192580	C9orf9
3192912	C9orf96
3181240	C9orf97
3867264	CA11
3628498	CA12
3105506	CA13
3105600	CA2
TCID	GENE_na29
3729419	CA4
3969855	CA5B
3969855	CA5BP
3137120	CA8
3168066	CA9
2531522	CAB39
3513752	CAB39L
3939707	CABIN1
3781531	CABLES1
3400730	CACNA1C
2624385	CACNA1D
3726618	CACNA1G
3643580	CACNA1H
3058991	CACNA2D1
2675315	CACNA2D2
3755580	CACNB1
3237396	CACNB2
2581349	CACNB4
3959787	CACNG2
3732092	CACNG4
3841184	CACNG6
2368198	CACYBP
3392332	CADM1
3864519	CADM4
2679406	CADPS
3364127	CALCA
3321592	CALCB
3456353	CALCOCO1
2591367	CALCRL
3025545	CALD1
3304767	CALHM2
2551924	CALM2
3599280	CALML4
3822049	CALR
3235516	CAMK1D
2782545	CAMK2D
3294854	CAMK2G
2400177	CAMK2N1
2823880	CAMK4
3229529	CAMSAP1
2374345	CAMSAP1L1
3420713	CAND1
3772775	CANT1
3861503	CAPN12
3590853	CAPN3
4018194	CAPN6
2612278	CAPN7
3831168	CAPNS1
3326183	CAPRIN1
2649182	CAPRIN1
3389450	CARD16
3389353	CARD17
3866958	CARD8
3820865	CARM1
2814693	CARTPT
3447798	CASC1
3266583	CASC2
3590014	CASC5
3013178	CASD1
TCID	GENE_na29
3644541	CASKIN1
3389450	CASP1
3389353	CASP1
3389257	CASP12
2796484	CASP3
3389273	CASP4
2781693	CASP6
2916952	CASP8AP2
2429556	CASQ2
2638824	CASR
2868131	CAST
2821194	CAST
3326400	CAT
3378043	CATSPER1
2825514	CATSPER2
3020302	CAV1
3020273	CAV2
3293998	CBARA1
3704567	CBFA2T3
3665116	CBFB
3352070	CBL
3812864	CBLN2
3910724	CBLN4
3919834	CBR3
3933923	CBS
2993639	CBX3
3456630	CBX5
3259367	CC2D2B
2619480	CCBP2
3686587	CCDC101
3251490	CCDC109A
2739160	CCDC109B
3941643	CCDC117
3987607	CCDC121
2545999	CCDC121
3858794	CCDC123
2860614	CCDC125
2992963	CCDC126
2670903	CCDC13
3822287	CCDC130
2692573	CCDC14
2560149	CCDC142
3765167	CCDC144A
3748400	CCDC144A
3009838	CCDC146
2582701	CCDC148
2855285	CCDC152
2423264	CCDC18
3367036	CCDC34
2622026	CCDC36
3731228	CCDC45
2641449	CCDC48
2657981	CCDC50
3468225	CCDC53
2343170	CCDC55
3434193	CCDC64
3413643	CCDC65
3344685	CCDC67
3808745	CCDC68

Fig. 12 Continued

TCID	GENE_na29
2881860	CCDC69
3241601	CCDC7
2425173	CCDC76
2348792	CCDC76
2688813	CCDC80
3387771	CCDC82
3351733	CCDC84
2553771	CCDC88A
2773407	CCDC90B
3817400	CCDC94
2759393	CCDC96
2840002	CCDC99
2948821	CCHCR1
3718204	CCL13
3753966	CCL14
3753966	CCL15
3753956	CCL16
3204285	CCL19
3204301	CCL21
2855542	CCL28
3754070	CCL3
3754070	CCL3L1
3754070	CCL3L3
3485674	CCNA1
2741768	CCNA2
2966371	CCNC
3380065	CCND1
3338192	CCND1
3401704	CCND2
2953866	CCND3
2838598	CCNG1
3259400	CCNJ
2884578	CCN JL
2702307	CCNL1
4041923	CCNL2
2391532	CCNL2
3453218	CCNT1
2507209	CCNT2
2620736	CCR9
3213530	CCRK
3336351	CCS
3421630	CCT2
3003193	CCT6A
3003193	CCT6P1
3928070	CCT8
2913694	CD109
3316344	CD151
2432714	CD160
3442706	CD163
3835035	CD177
2860178	CD180
3655109	CD19
2362180	CD1A
2353669	CD2
2636125	CD200
3830359	CD22
4035833	CD24
3161082	CD274
3601229	CD276
TCID	GENE_na29
2909404	CD2AP
3770305	CD300A
3770305	CD300C
3770361	CD300LF
2583254	CD302
2453307	CD34
3010503	CD36
3838385	CD37
3393744	CD3D
3836243	CD3EAP
3865378	CD3EAP
3992575	CD40LG
3326635	CD44
2377427	CD46
2687739	CD47
2440354	CD48
3332729	CD5
2326463	CD52
2351572	CD53
2377229	CD55
3368707	CD59
3457160	CD63
3443868	CD69
2881370	CD74
3834502	CD79A
3402315	CD9
2635741	CD96
3822657	CD97
3966836	CD99
3489350	CDADC1
3235461	CDC123
2348896	CDC14A
3503164	CDC16
2857042	CDC20B
3221543	CDC26
3760625	CDC27
2969467	CDC2L6
2921086	CDC40
2324634	CDC42
3119339	CDC42
2459042	CDC42BPA
3580498	CDC42BPB
3335070	CDC42EP2
3770029	CDC42EP4
2434776	CDC42SE1
3720896	CDC6
2346399	CDC7
2372967	CDC73
3402874	CDCA3
2992243	CDCA7L
3040897	CDCA7L
2671728	CDCP1
3694657	CDH11
3671202	CDH13
3695315	CDH16
3811949	CDH19
3802602	CDH2
3251068	CDH23
3666366	CDH3
TCID	GENE_na29
2805078	CDH6
3417146	CDK2
3476097	CDK2AP1
3458783	CDK4
3223425	CDK5RAP2
2813481	CDK7
3482498	CDK8
3563861	CDKL1
2773719	CDKL2
2905169	CDKN1A
3201488	CDKN2B
2871896	CDO1
3396770	CDON
4024373	CDR1
3684782	CDR2
2734270	CDS1
3863669	CEACAM1
3834257	CEACAM21
3834379	CEACAM5
3834341	CEACAM5
3834379	CEACAM6
2321813	CELA2A
2321813	CELA2B
2473991	CENPA
2771654	CENPC1
2780172	CENPE
3505937	CENPJ
2859667	CENPK
2444451	CENPL
3187577	CEP110
2463864	CEP170
2463864	CEP170L
3779817	CEP192
3464622	CEP290
2369843	CEP350
3258444	CEP55
3345593	CEP57
2351632	CEPT1
2518272	CERKL
3692701	CES1
3664982	CES2
3692701	CES4
3665049	CES8
4026263	CETN2
3662417	CETP
2902844	CFB
2373336	CFH
2373406	CFH
2373336	CFHR1
2373406	CFHR1
2373406	CFHR3
2373406	CFHR4
2781736	CFI
2522616	CFLAR
3508696	CG030
3867573	CGB
3867573	CGB2
3867573	CGB5
3867573	CGB7

Fig. 12 Continued

TCID	GENE_na29
3867573	CGB8
3595315	CGNL1
3817501	CHAF1A
3920003	CHAF1B
3483159	CHCHD2
3099089	CHCHD7
2356721	CHD1L
3609138	CHD2
3709244	CHD3
3442054	CHD4
3906160	CHD6
3660858	CHD9
2676854	CHDH
3354799	CHEK1
3479438	CHFR
3549092	CHGA
3875179	CHGB
2451593	CHI3L1
3358538	CHID1
2607568	CHL1
4006841	CHMP5
2994835	CHN2
3915569	CHODL
3468080	CHPT1
3428671	CHPT1
4018080	CHRDL1
3615791	CHRFAM7A
3359897	CHRNA10
3129026	CHRNA2
3913775	CHRNA4
3603436	CHRNA5
3615791	CHRNA7
3708663	CHRNBT1
2360346	CHRNBT2
2532378	CHRND
3742351	CHRNE
2532399	CHRNG
2567242	CHST10
2646125	CHST2
3802416	CHST9
2494447	CIAO1
3629877	CIDEB
3629529	CILP
2737840	CISD2
2675504	CISH
3474104	CIT
4012178	CITED1
2976768	CITED2
3205659	CKAP2
3490655	CKAP2
3371719	CKAP5
3580769	CKB
3664785	CKLF
3621351	CKMT1A
3178583	CKS2
2573641	CLASP1
2668425	CLASP2
3862108	CLC
2345023	CLCA1
TCID	GENE_na29
2345095	CLCA3P
3378758	CLCF1
2751385	CLCN3
2320472	CLCN6
2710599	CLDN1
3497195	CLDN10
3012019	CLDN12
26657808	CLDN16
2932593	CLDN20
3007960	CLDN4
3952762	CLDN5
3743551	CLDN7
3928415	CLDN8
3868987	CLDND2
3443891	CLEC2B
3404436	CLEC2D
2620448	CLEC3B
3443183	CLEC4E
3848525	CLEC4G
3848525	CLEC4GP1
3076868	CLEC5A
3444009	CLEC7A
2787005	CLGN
2949330	CLIC1
4027769	CLIC2
3230594	CLIC3
2955638	CLIC5
2955556	CLIC5
3919278	CLIC6
3860229	CLIP3
2475407	CLIP4
2437329	CLK2
2437329	CLK2P
3577940	CLMN
3494502	CLN5
3629494	CLPX
2406420	CLSPN
2395890	CLSTN1
3168415	CLTA
2842530	CLTB
3729172	CLTC
3129065	CLU
3775808	CLUL1
2945882	CMAH
3407926	CMAS
2668021	CMTM6
2817464	CMYA5
3793760	CNDP2
2382419	CNIH4
3971219	CNKS2
2980516	CNKS3
3821263	CNN1
3815399	CNN2
2424102	CNN3
3260265	CNNM1
3261971	CNNM2
3693673	CNOT1
3421897	CNOT2
3074362	CNOT4
TCID	GENE_na29
2844709	CNOT6
2774565	CNOT6L
2963859	CNR1
3758157	CNTD1
3331730	CNTF
3163728	CNTLN
3411721	CNTN1
3345940	CNTN5
2607757	CNTN6
3721989	CNTNAP1
3029900	CNTNAP2
3531163	COCH
3769969	COG1
3488253	COG3
3067080	COG5
2425756	COL11A1
2961177	COL12A1
3250486	COL13A1
3181642	COL15A1
2404546	COL16A1
3762198	COL1A1
3013054	COL1A2
2889542	COL23A1
2781441	COL25A1
2642261	COL29A1
2519577	COL3A1
3924424	COL6A2
2673345	COL7A1
2633390	COL8A1
2406579	COL8A2
3892974	COL9A3
3113133	COLEC10
3518169	COMM6
2768145	COMM8
3369762	COMM9
2363084	COPA
3855324	COPE
2532064	COPS7B
3761054	COPZ2
3417531	COQ10A
2775965	COQ2
3571667	COQ6
3470549	CORO1C
3751590	CORO6
3711165	COX10
3303109	COX15
2772968	COX18
2451309	COX7C
2700244	CP
3106559	CP
2647109	CPA3
3716411	CPD
2750627	CPE
2719361	CPEB2
2841699	CPEB4
2842255	CPLX2
3461341	CPM
3105904	CPNE3
2695453	CPNE4

Fig. 12 Continued

TCID	GENE_na29
3450655	CPNE8
2524817	CPO
3158516	CPSF1
2468920	CPSP3
2315739	CPSF3L
3421446	CPSP6
3043648	CPVL
3895118	CPXM1
2377427	CR1L
2377283	CR2
3603295	CRABP1
2438458	CRABP2
2660648	CRBN
2994558	CREB5
3677795	CREBBP
2567647	CREG2
2610136	CRELD1
3950452	CRELD2
2816536	CRHBP
3554818	CRIP2
2956563	CRISP3
3720228	CRKRS
3855104	CRLF1
2759038	CRMP1
2398193	CROCC
2398193	CROCCL1
2398193	CROCCL2
3726772	CROP
3011317	CROT
3302572	CRTAC1
2616166	CRTAP
3469865	CRY1
3391149	CRYAB
2685776	CRYBG3
2418451	CRYZ
3929821	CRYZL1
3457614	CS
3444252	CSDA
3444252	CSDAP1
2406783	CSF3R
3126504	CSGALNACT1
3243908	CSGALNACT2
2880932	CSNK1A1
2880932	CSNK1A1L
3960478	CSNK1E
3629103	CSNK1G1
3629012	CSNK1G1
2826550	CSNK1G3
3894228	CSNK2A1
3894228	CSNK2A1P
2902559	CSNK2B
3101893	CSPP1
2669930	CSRNP1
3454662	CSRNP2
2450865	CSRP1
3463056	CSRP2
3901333	CST1
3901333	CST4
3335894	CST6

TCID	GENE_na29
2638869	CSTA
3934245	CSTB
3289631	CSTF2T
3368520	CSTF3
4020137	CT47A1
4020137	CT47A10
4020137	CT47A11
4020137	CT47A2
4020137	CT47A3
4020137	CT47A4
4020137	CT47A5
4020137	CT47A6
4020137	CT47A7
4020137	CT47A8
4020137	CT47A9
4020137	CT47B1
3533499	CTAGE5
3029230	CTAGE6
2417390	CTBP2
2420467	CTBS
3665603	CTCF
2527786	CTDSP1
3458911	CTDSP2
2617276	CTDSP1
3591909	CTDSP2
2974330	CTGF
2341663	CTH
3110317	CTHRC1
2830946	CTNNA1
3292169	CTNNA3
3219621	CTNNAL1
2618940	CTNNB1
4000839	CTPS2
3887117	CTSA
3086206	CTSB
3124537	CTSB
3385769	CTSC
3358950	CTSD
3634811	CTSH
2434609	CTSK
3216671	CTSL2
2434575	CTSS
3338552	CTTN
3069470	CTTNBP2
3848644	CTXN1
3279698	CUBN
3764022	CUEDC1
3502497	CUL4A
4019900	CUL4B
3347549	CUL5
2954355	CUL7
2954355	CUL9
2907754	CUL9
3310675	CUZD1
3389745	CWF19L2
2726542	CWH43
2773434	CXCL1
2773972	CXCL11
3286602	CXCL12

TCID	GENE_na29
2732508	CXCL13
2876608	CXCL14
3863640	CXCL17
2773434	CXCL2
2773434	CXCL3
2773947	CXCL9
2578028	CXCR4
4024420	CXorf18
4003895	CXorf21
4002011	CXorf23
3973891	CXorf27
4006504	CXorf36
3981474	CXorf50B
3986230	CXorf57
3980264	CXorf62
3807809	CXXC1
2831350	CXXC5
3709213	CYB5D1
3962530	CYB5R3
2915491	CYB5R4
3973839	CYBB
2515240	CYBRD1
3613300	CYFIP1
2837266	CYFIP2
3660213	CYLD
3157217	CYP11B1
3157217	CYP11B2
3304522	CYP17A1
3261886	CYP17A1OS
2548699	CYP1B1
2523635	CYP20A1
2903034	CYP21A2
3910429	CYP24A1
3258384	CYP26A1
2559189	CYP26B1
2528093	CYP27A1
3458819	CYP27B1
3862944	CYP2A13
3862944	CYP2A6
3862944	CYP2A7
3862944	CYP2A7P1
3259019	CYP2C19
3259019	CYP2C9
2955761	CYP39A1
3063501	CYP3A4
3015040	CYP3A43
3063406	CYP3A5
3063406	CYP3A5P2
3063501	CYP3A7
3551432	CYP46A1
2335048	CYP4A11
2335048	CYP4A22
3853658	CYP4F11
3823340	CYP4F12
2334986	CYP4X1
3060994	CYP51A1
3138204	CYP7B1
2344888	CYR61
4013460	CYSLTR1

Fig. 12 Continued

TCID	GENE_na29
3489138	CYSLTR2
3772525	CYTH1
3837836	CYTH2
3037251	CYTH3
3940001	CYTSA
3714177	CYTSB
3927392	CYYR1
3926138	D21S2089E
3517251	DACH1
3538087	DACT1
2622121	DAG1
2673830	DALRD3
2848464	DAP
2361036	DAP3
3628832	DAPK2
2577958	DARS
4031834	DAZ1
2664760	DAZ2
4031834	DAZ2
2664760	DAZ3
4031834	DAZ3
2664760	DAZ4
4031834	DAZ4
3815834	DAZAP1
2664760	DAZL
3223157	DBC1
2888800	DBN1
2425212	DBT
2515183	DCAF17
2686023	DCBLD2
3788560	DCC
3367788	DCDC1
3367917	DCDC1
2945440	DCDC2
3367788	DCDC5
3361072	DCHS1
2790486	DCHS2
3509473	DCLK1
3465274	DCN
3440192	DCP1B
2824354	DCP2
3520934	DCT
2795819	DCTD
2559967	DCTN1
3458614	DCTN2
2881554	DCTN4
3092325	DCTN6
3683845	DCUN1D3
2726828	DCUN1D4
4018218	DCX
3824212	DDA1
2420832	DDAH1
3329649	DDB2
3050388	DDC
3153633	DDEF1IT1
2321911	DDI2
3251393	DDIT4
2969350	DDO
2901970	DDR1
TCID	GENE_na29
2364231	DDR2
3347831	DDX10
2944068	DDX18
3250055	DDX21
3577513	DDX24
3354896	DDX25
3974838	DDX3X
3974838	DDX3Y
2829488	DDX46
3766893	DDX5
2476116	DDX50
3754736	DDX52
3203086	DDX58
2450416	DDX59
3393946	DDX6
2792800	DDX60
3358262	DEAF1
3122721	DEFA4
3122721	DEFA8P
3873091	DEFB132
2382360	DEGS1
2944068	DEK
3847906	DENND1C
3118651	DENND3
3629811	DENND4A
3362263	DENND5A
3449760	DENND5B
3435490	DENR
2858592	DEPDC1B
3943101	DEPDC5
3113280	DEPDC6
3151401	DERL1
3742756	DERL2
2528476	DES
3638068	DET1
3980482	DGAT2L6
3064501	DGAT2L7
3039247	DGKB
3487095	DGKH
3074912	DGKI
2413907	DHCR24
2326496	DHDDS
2817837	DHFR
3667811	DHODH
3558118	DHRS1
3719210	DHRS11
3529237	DHRS2
2397025	DHRS3
3567187	DHRS7
3235373	DHTKD1
2763805	DHX15
3311775	DHX32
3269662	DHX32
3884922	DHX35
2549007	DHX57
3722554	DHX8
3494502	DHX9
3475511	DIABLO
3983962	DIAPH2
TCID	GENE_na29
3577870	DICER1
2336891	DIO1
3573870	DIO2
3924674	DIP2A
3414561	DIP2B
3273251	DIP2C
3214227	DIRAS2
2417362	DIRAS3
3598613	DIS3L
2385343	DISC1
2526980	DKFZp434H1419
2731636	DKFZP564O0823
2995320	DKFZP586I1420
3860912	DKFZp761D1918
3348852	DLAT
3125116	DLC1
3489708	DLEU1
3513995	DLEU2
3513995	DLEU2L
3384704	DLG2
3980643	DLG3
3708306	DLG4
3743393	DLG4
3082759	DLGAP2
3883819	DLGAP4
3565663	DLGAP5
3544346	DLST
3544346	DLSTP
2333794	DMAP1
4004044	DMD
3859761	DMKN
3159735	DMRT3
2336809	DMRTB1
3981474	DMRTC1
3981474	DMRTC1B
3060051	DMTF1
2825514	DMXL1
3624145	DMXL2
2992243	DNAH11
3040897	DNAH11
2382781	DNAH14
2849056	DNAH5
2491168	DNAH6
2491089	DNAH6
2593013	DNAH7
3690084	DNAJA2
3603247	DNAJA4
3852783	DNAJB1
2656569	DNAJB11
3629811	DNAJB14
2343289	DNAJB4
3397461	DNAJB6
3961699	DNAJB7
3018866	DNAJB9
3280902	DNAJC1
3292413	DNAJC12
2642791	DNAJC13
3457201	DNAJC14
3619650	DNAJC17

Fig. 12 Continued

TCID	GENE_na29
3590129	DNAJC17
2806256	DNAJC21
3367965	DNAJC24
3184940	DNAJC25
3184940	DNAJC25-GNG10
3497270	DNAJC3
2340350	DNAJC6
2847264	DNAL4
2330723	DNAL11
3644664	DNASE1L2
2678298	DNASE1L3
2602770	DNER
2600155	DNPEP
3259503	DNTT
2423597	DNTTIP2
3269939	DOCK1
2601648	DOCK10
3988435	DOCK11
2622970	DOCK3
3090512	DOCK5
3159330	DOCK8
3522398	DOCK9
3929821	DONSON
3929775	DONSON
3919860	DOPEY2
3394192	DPACT1
2902089	DPCR1
3913483	DPH3B
2584018	DPP4
3032647	DPP6
2688166	DPPA2
2688180	DPPA4
2443120	DPT
3065546	DPY19L2
3459801	DPY19L2
3065546	DPY19L2P1
3459801	DPY19L2P1
3065546	DPY19L2P2
3459801	DPY19L2P3
3107606	DPY19L4
2547386	DPY30
2424524	DPYD
2880292	DPYSL3
3984702	DRP2
3803020	DSC1
3802980	DSC2
3802924	DSC3
3393311	DSCAML1
3812074	DSEL
3783529	DSG2
3904566	DSN1
2893794	DSP
2958325	DST
3878533	DTD1
2378937	DTL
3784208	DTNA
2638962	DTX3L
4046876	DTYMK
TCID	GENE_na29
2636626	DULLARD
2815139	DULLARD
3592214	DUOX1
3622176	DUOX1
3592214	DUOX2
3622176	DUOX2
3622239	DUOXA1
3622239	DUOXA2
2887309	DUSP1
2457261	DUSP10
3719515	DUSP14
3444958	DUSP16
2518729	DUSP19
2891241	DUSP22
2362702	DUSP23
3129731	DUSP4
3263743	DUSP5
3464860	DUSP6
3593147	DUT
2391425	DVL1
3552847	DYNC1H1
3013565	DYNC1I1
4037595	DYNC1I2
2515276	DYNC1I2
2981874	DYNLT1
4004819	DYNLT3
2376894	DYRK3
2488252	DYSF
2596386	DYTN
3625391	DYX1C1
3497195	DZIP1
3521372	DZIP1
2635263	DZIP3
2401448	E2F2
2612371	EAF1
3111695	EBAG9
2883878	EBF1
3128411	EBF2
3873923	EBF4
3294242	ECD
2400518	ECE1
2413032	ECHDC2
2358360	ECM1
3214867	ECM2
2652675	ECT2
4011096	EDA2R
2386828	EDARADD
2608801	EDEM1
2447824	EDEM3
2865390	EDIL3
3891447	EDN3
2562435	EDNRB
3483159	EEF1A1
2960903	EEF1A1
2960903	EEF1AL7
2940920	EEF1E1
3962997	EFCAB6
2554018	EFEMP1
3087438	EFHA2
TCID	GENE_na29
2665472	EFHB
2360633	EFNA3
2360633	EFNA4
2869880	EFNA5
3557408	EFS
3635776	EFTUD1
3759356	EFTUD2
2739308	EGF
3969358	EGFL6
2807195	EGFLAM
3002640	EGFR
2830861	EGR1
3291601	EGR2
2484970	EHBP1
3837431	EHD2
3326461	EHF
4031136	EIF1AY
3325052	EIF2AK2
2563654	EIF2AK3
3544387	EIF2B2
2409904	EIF2B3
2330002	EIF2C4
3903288	EIF2S2
3309215	EIF3A
2584712	EIF3E
2483544	EIF3F
3945056	EIF3L
2655688	EIF4G1
2400373	EIF4G3
3008144	EIF4H
3553607	EIF5
3788270	ELAC1
3848689	ELAVL1
4021433	ELF4
3427098	ELK3
3046197	ELMO1
3907830	ELMO2
2409310	ELOVL1
2941721	ELOVL2
2962113	ELOVL4
2781813	ELOVL6
3784727	ELP2
3091628	ELP3
2419432	ELTD1
2779543	EMCN
3865511	EML2
2478748	EML4
3575371	EML5
3679959	EMP2
3822657	EMR2
3852832	EMR3
2458338	ENAH
2382781	ENAH
2730503	ENAM
3345427	ENDOD1
3226709	ENDOG
2395490	ENO1
2386828	ENO1
2733767	ENOPH1

Fig. 12 Continued

TCID	GENE_na29
3795866	ENOSF1
2925953	ENPP1
2909020	ENPP4
2955673	ENPP5
3259253	ENTPD1
3230733	ENTPD2
3571667	ENTPD5
3260383	ENTPD7
3438772	EP400
3438772	EP400NL
2480383	EPAS1
2327677	EPB41
3883690	EPB41L1
2973995	EPB41L2
3797032	EPB41L3
2870964	EPB41L4A
3219788	EPB41L4B
2503109	EPB41L5
3284073	EPC1
2997907	EPDR1
2632225	EPHA3
2600689	EPHA4
2771342	EPHA5
2965206	EPHA7
2643592	EPHB1
2655845	EPHB3
3028858	EPHB6
2346625	EPHX4
3772187	EPR1
2456746	EPRS
3445908	EPS8
3841949	EPS8L1
3511698	EPST11
3728588	EPX
2868131	ERAP1
3720402	ERBB2
3417249	ERBB3
2597552	ERBB4
3865378	ERCC1
3836217	ERCC2
3288707	ERCC6
4012142	ERCC6L
2858752	ERCC8
3931765	ERG
3570049	ERH
3085065	ERI1
3683845	ERI2
3651509	ERI2
3303255	ERLIN1
3564790	ERO1L
2462329	ERO1LB
3445768	ERP27
3217736	ERP44
2395177	ERRFI1
3053380	ERV3
3396107	ESAM
3800779	ESCO1
2856995	ESM1
2534664	ESPNL

TCID	GENE_na29
2931763	ESR1
3568184	ESR2
3107548	ESRP1
3696226	ESRP2
2455933	ESRRG
3633794	ETFA
3868963	ETFB
3864430	ETHE1
2451870	ETNK2
3397589	ETS1
3039177	ETV1
3758510	ETV4
2709132	ETV5
3405207	ETV6
2423017	EV15
3355733	EWSR1
3941907	EWSR1
2938196	EXOC2
3025005	EXOC4
3771336	EXOC7
2396480	EXOSC10
3485863	EXOSC8
2741768	EXOSC9
3150060	EXT1
3140037	EYA1
3887479	EYA2
2403335	EYA3
3078348	EZH2
3502437	F10
2755154	F11
2888741	F12
2816459	F2R
2863363	F2RL2
2423907	F3
2443370	F5
4027639	F8
2334740	FAAH
3979101	FAAH2
2404418	FABP3
3142381	FABP4
3517694	FABP5
3606304	FABP5
2838116	FABP6
3402571	FADS1
3333226	FADS2
3603932	FAH
3644191	FAHD1
2565753	FAHD2A
2494064	FAHD2A
2565753	FAHD2B
2494064	FAHD2B
2984835	FAM103A1
3769969	FAM104A
3278813	FAM107B
3869097	FAM107B
3331926	FAM111A
3331903	FAM111B
2724094	FAM114A1
3029230	FAM115A

TCID	GENE_na29
3029230	FAM115C
2678116	FAM116A
3948543	FAM118A
3355021	FAM118B
3418534	FAM119B
3179872	FAM120A
4009560	FAM120C
3490073	FAM124A
3189422	FAM125B
3041294	FAM126A
2594627	FAM126B
4022925	FAM127A
3991889	FAM127A
4022925	FAM127B
3991889	FAM127B
4022925	FAM127C
3991889	FAM127C
3225952	FAM129B
3077273	FAM131B
2528275	FAM134A
2849992	FAM134B
2777487	FAM13A
3247977	FAM13C
3290649	FAM13C
3360772	FAM160A2
3265432	FAM160B1
2638886	FAM162A
2369713	FAM163A
3104323	FAM164A
4051521	FAM166A
2576281	FAM168B
2825796	FAM170A
3759105	FAM171A2
2519294	FAM171B
3643347	FAM173A
2821981	FAM174A
2776088	FAM175A
2366941	FAM175A
2560625	FAM176A
3532353	FAM177A1
3260829	FAM178A
2475348	FAM179A
3534128	FAM179B
3549436	FAM181A
3649052	FAM18B
2628482	FAM19A1
3949722	FAM19A5
3768535	FAM20A
3244742	FAM21A
3244742	FAM21B
3244742	FAM21C
3244742	FAM21D
3310675	FAM24B
3764738	FAM33A
3256279	FAM35A
3256279	FAM35B2
3704376	FAM38A
3798778	FAM38B
3798829	FAM38B2

Fig. 12 Continued

TCID	GENE_na29
3921992	FAM3B
3070047	FAM3C
2678468	FAM3D
2658785	FAM43A
2962383	FAM46A
2353988	FAM46C
2731986	FAM47E
3509910	FAM48A
2528476	FAM48A
2541699	FAM49A
3996430	FAM50A
2757278	FAM53A
2830698	FAM53C
2634091	FAM55C
3705491	FAM57A
4026560	FAM58A
3803290	FAM59A
2368840	FAM5B
2448710	FAM5C
3449700	FAM60A
3417371	FAM62A
3082248	FAM62B
2434746	FAM63A
2358591	FAM63A
2423175	FAM69A
4019784	FAM70A
3839400	FAM71E1
2343170	FAM73A
3190893	FAM73B
3171425	FAM75A1
3171425	FAM75A2
3171425	FAM75A3
3171425	FAM75A4
3171425	FAM75A5
3171425	FAM75A6
3171425	FAM75A7
3171425	FAM75C1
2820813	FAM81B
3619595	FAM82A2
3143330	FAM82B
3884892	FAM83D
3946146	FAM83F
2470470	FAM84A
3152558	FAM84B
3335338	FAM89B
3114365	FAM91A1
3114365	FAM91A2
3107151	FAM92A1
3107151	FAM92A2
3107151	FAM92A3
3589212	FAM98B
2610241	FANCD2
3607537	FANCI
3534248	FANCM
2584134	FAP
3321269	FAR1
3409605	FAR2
3497881	FARP1
2893130	FARS2

TCID	GENE_na29
3257031	FAS
3257098	FAS
3079336	FASTK
2586227	FASTKD1
2797393	FAT1
2742581	FAT4
3862167	FBL
3948640	FBLN1
3576749	FBLN5
2500667	FBLN7
3623031	FBN1
2874371	FBN2
3656635	FBXL19
3755655	FBXL20
3518455	FBXL3
2966078	FBXL4
2761734	FBXL5
2802963	FBXL7
2552153	FBXO11
3813198	FBXO15
3129361	FBXO16
2396750	FBXO2
3473480	FBXO21
3602526	FBXO22
3602526	FBXO22OS
3368748	FBXO3
2978026	FBXO30
2531129	FBXO36
2980241	FBXO5
3943414	FBXO7
2886977	FBXW11
3223605	FBXW2
3230530	FBXW5
2789957	FBXW7
3433591	FBXW8
2453036	FCAMR
3848492	FCER2
3544216	FCF1
3862188	FCGBP
2363689	FCGR2A
2363689	FCGR2B
2363689	FCGR2C
2440943	FCGR3A
2440943	FCGR3B
3838624	FCGRT
2815139	FCHO2
3381377	FCHSD2
3229338	FCN1
3229338	FCN2
2363852	FCRLA
3086206	FDFT1
3348189	FDX1
3391149	FDXACB1
3770457	FDXR
3333226	FEN1
3396593	FEZ1
2612100	FGD5
3708663	FGF11
2710895	FGF12

TCID	GENE_na29
3523499	FGF14
2742109	FGF2
3441168	FGF23
3593408	FGF7
3480885	FGF9
2761829	FGFBP1
3132016	FGFR1
3413950	FGFR1OP2
3310041	FGFR2
2338487	FGGY
3057955	FGL2
2747961	FHDC1
3992408	FHL1
2568687	FHL2
3695541	FHOD1
3324447	FIBIN
3377964	FIBP
2727226	FIP1L1
3326826	FJX1
3721452	FKBP10
3043936	FKBP14
2472955	FKBP1B
3334339	FKBP2
3401119	FKBP4
2951567	FKBP5
3824963	FKBP8
3498315	FKSG29
3355733	FLI1
3748188	FLII
3110217	FLJ10489
4040117	FLJ12120
2738244	FLJ20184
3829638	FLJ21369
3358049	FLJ23519
3018011	FLJ23834
2369950	FLJ23867
3937967	FLJ26056
3953456	FLJ26056
2460470	FLJ30430
2617041	FLJ31715
3489708	FLJ31945
3346147	FLJ32810
3415046	FLJ33996
3301609	FLJ34077
2716246	FLJ35424
2926447	FLJ35700
2821761	FLJ35946
2485257	FLJ36848
3299782	FLJ37201
2536965	FLJ38379
2907396	FLJ38717
3903461	FLJ38773
3724591	FLJ39349
2672629	FLJ39534
2564816	FLJ40330
2565935	FLJ40330
3887165	FLJ40606
2536996	FLJ41327
3380647	FLJ42102

Fig. 12 Continued

TCID	GENE_na29
3380065	FLJ42258
3338192	FLJ42258
3636879	FLJ43276
3981361	FLJ44635
3673661	FLJ45121
3549989	FLJ45244
2812591	FLJ46010
3918953	FLJ46020
3097401	FLJ46365
2481379	FLJ46838
3722479	FLJ77644
4027176	FLNA
2625907	FLNB
3546924	FLRT2
3898355	FLRT3
3507282	FLT1
3838556	FLT3LG
2387711	FMN2
2510713	FMNL2
3414104	FMNL3
3454006	FMNL3
2366941	FMO3
2433232	FMO5
2451693	FMOD
3994100	FMR1
2598261	FN1
2526806	FN1
3739147	FN3K
3227159	FNBP1
3489212	FNDC3A
2652410	FNDC3B
2545953	FNDC4
2405250	FNDC5
2350316	FNDC7
3343900	FOLH1
3372896	FOLH1
3343900	FOLH1B
3372896	FOLH1B
3339406	FOLR1
3544525	FOS
3836266	FOSB
2743085	FOSL1
3672609	FOXF1
3440598	FOXM1
3715642	FOXN1
2920475	FOXO3
2920475	FOXO3B
3019793	FOXP2
2342176	FPGT
2342220	FPGT
3869237	FPR1
3839910	FPR2
2732655	FRAS1
3486096	FREM2
3212008	FRMD3
3278401	FRMD4A
3621728	FRMD5
3287995	FRMPD2
3287995	FRMPD2L1
TCID	GENE_na29
3287995	FRMPD2L2
2590715	FRZB
3022422	FSCN3
3183238	FSD1L
3815097	FSTL3
2875685	FSTL4
2791894	FSTL5
2545841	FTH1
2545841	FTHL3P
2419235	FUBP1
3656904	FUS
3837934	FUT2
3540552	FUT8
3830166	FXYD3
3830216	FXYD5
3393479	FXYD6
2854327	FYB
2969886	FYN
2659887	FYTD1
3343452	FZD4
3385509	FZD4
2596763	FZD5
3110272	FZD6
2523045	FZD7
2378068	G0S2
2745547	GAB1
3383227	GAB2
3404636	GABARAPL1
3404636	GABARAPL3
3217242	GABBR2
3916576	GABPA
2358700	GABPB2
2768056	GABRA4
2884845	GABRB2
3614534	GABRB3
4054481	GABRD
3585272	GABRG3
2964092	GABRR1
2341083	GADD45A
3816509	GADD45B
2765935	GAFA3
3575103	GALC
2401581	GALE
3593339	GALK2
3623472	GALK2
3181600	GALNT12
2511045	GALNT13
2384788	GALNT2
2585129	GALNT3
3464912	GALNT4
2511603	GALNT5
2751936	GALNT7
3401920	GALNT8
3479015	GALNT9
3363091	GALNTL4
3590853	GANC
2637112	GAP43
3402625	GAPDH
3146661	GAPDHL7
TCID	GENE_na29
3189110	GAPVD1
3183238	GARNL1
3189714	GARNL3
3929721	GART
3323891	GAS2
3428268	GAS2L3
2444451	GAS5
3781245	GATA6
3825713	GATAD2A
3622386	GATM
2437205	GBA
2437205	GBAP
2684187	GBE1
3261544	GBF1
3228523	GBGT1
2421843	GBP1
2421883	GBP1
2421925	GBP2
2421883	GBP2
2421843	GBP3
2421925	GBP3
2421883	GBP3
2421995	GBP4
2421925	GBP4
2422035	GBP5
2345816	GBP6
2421925	GBP7
2512930	GCA
3945014	GCAT
2636062	GCET2
2688605	GCET2
3565524	GCH1
3048468	GCK
2957560	GCM1
3595441	GCOM1
3701459	GCSH
3174510	GDA
3103607	GDAP1
2430370	GDAP2
3855285	GDF1
3287789	GDF10
3824993	GDF15
3996404	GDI1
3275132	GDI2
3729014	GDPD1
3382319	GDPD5
3144934	GEM
2477980	GEMIN6
2649640	GFM1
2862716	GFM2
2558045	GFPT1
3308241	GFRA1
3127156	GFRA2
3685183	GGA2
3770663	GGA3
3044129	GGCT
3719161	GGNBP2
3939914	GGT1
3939009	GGT1

Fig. 12 Continued

TCID	GENE_na29
3939914	GGT2
3939009	GGT2
3939914	GGT3P
3939009	GGT3P
3939914	GGT8P
3223967	GGTA1
3939914	GGTLC2
3939009	GGTLC2
3939914	GGTLC3
3939009	GGTLC3
3031556	GIMAP2
3031533	GIMAP4
3031573	GIMAP5
3079103	GIMAP6
3031517	GIMAP7
3031466	GIMAP8
3880827	GINS1
3703112	GINS2
3471005	GIT2
3656032	GIYD1
3656032	GIYD2
2923661	GJA1
2407755	GJA9
4054427	GJB4
3504213	GJB6
4015763	GLA
2668205	GLB1
2989537	GLCC1
3197955	GLDC
3593931	GLDN
3698919	GLG1
3047660	GLI3
3422855	GLIPR1
3422826	GLIPR1L2
3197140	GLISS3
3197014	GLISS3
2749191	GLRB
3271018	GLRX3
2520291	GLS
3437500	GLT1D1
3824471	GLT25D1
2676319	GLT8D1
3468888	GLT8D2
3451246	GLT8D3
3978169	GLTSCR2
2447066	GLUL
2487478	GMCL1
2487478	GMCL1L
3914021	GMEB2
3861948	GMFG
2898597	GMNN
2528620	GMPPA
3210737	GNA14
3010439	GNA11
3779207	GNAL
3661940	GNAO1
3210808	GNAQ
2391840	GNB1
3402874	GNB3
TCID	GENE_na29
3184940	GNG10
2417272	GNG12
3535628	GNG2
3333595	GNG3
3845944	GNG7
2948379	GNL1
2858752	GNL3L
2858752	GNL3LP
2907513	GNMT
2385197	GNPAT
2879028	GNPDA1
3564872	GNPNAT1
3468103	GNPTAB
3460127	GNS
3225224	GOLGA1
2617041	GOLGA4
3617458	GOLGA8A
3617574	GOLGA8A
3617458	GOLGA8B
3617574	GOLGA8B
3212848	GOLM1
3148871	GOLSYN
2451931	GOLT1A
2437645	GON4L
2971378	GOPC
2669888	GORASP1
4049835	GOT1L1
3707335	GP1BA
2711627	GP5
3306984	GPAM
3829174	GPATCH1
2402861	GPATCH3
3759186	GPATCH8
3063807	GPC2
4022447	GPC3
2615808	GPD1L
2511432	GPD2
2987038	GPER
2794584	GPM6A
2474681	GPN1
3191273	GPR107
2955932	GPR110
2955999	GPR110
2473784	GPR113
3094334	GPR124
2763278	GPR125
3581404	GPR132
3438061	GPR133
2386747	GPR137B
2505779	GPR148
2587790	GPR155
2651835	GPR160
2442911	GPR161
3978169	GPR173
3982612	GPR174
2417390	GPR177
3522644	GPR18
3417767	GPR182
3522662	GPR183
TCID	GENE_na29
3445028	GPR19
3070873	GPR37
2576988	GPR39
2506570	GPR39
3994964	GPR50
3662808	GPR56
4001654	GPR64
2553262	GPR75
3387010	GPR83
2819779	GPR98
3985260	GPRASP1
3985305	GPRASP2
3405587	GPRC5A
3683377	GPRC5B
2888385	GPRIN1
2777639	GPRIN3
2674229	GPX1
3568603	GPX2
2835715	GPX3
3815538	GPX4
2336439	GPX7
2809831	GPX8
2636626	GRAMD1C
2827057	GRAMD3
3050462	GRB10
2584712	GRB14
3770743	GRB2
2469825	GREB1
2469157	GRHL1
2325358	GRHL3
2749222	GRIA2
3989448	GRIA3
4050485	GRIN1
3770422	GRIN2C
3218151	GRIN3A
3595441	GRINL1A
3071063	GRM8
3790529	GRP
2835006	GRPEL2
2772614	GRSF1
3755903	GSDMB
3445156	GSG1
3741585	GSG2
2691014	GSK3B
3187686	GSN
3903670	GSS
2738314	GSTCD
2350922	GSTM1
2350981	GSTM1
2350952	GSTM1
2350922	GSTM2
2351004	GSTM2
2350981	GSTM2
2350952	GSTM2
2427208	GSTM3
2350922	GSTM4
2350981	GSTM4
2350922	GSTM5
2351004	GSTM5

Fig. 12 Continued

TCID	GENE_na29
3262509	GSTO1
3337168	GSTP1
2727535	GSX2
3574074	GTF2A1
2481379	GTF2A1L
2421753	GTF2B
2861183	GTF2H2
2861183	GTF2H2B
2861183	GTF2H2C
2861183	GTF2H2D
3435946	GTF2H3
2902013	GTF2H4
3008376	GTF2I
3008376	GTF2IP1
3056656	GTF2IRD2
3056656	GTF2IRD2B
3056656	GTF2IRD2P
3482888	GTF3A
2593352	GTF3C3
2954771	GTPBP2
2519480	GULP1
2787958	GYPB
2787902	GYPB
2504328	GPC
2787958	GYPE
2787902	GYPE
3446845	GYS2
2809810	GZMA
3558375	GZMB
3558375	GZMH
2809793	GZMK
2876479	H2AFY
3250602	H2AFY2
2946714	H2BFS
3770944	H3F3B
2664395	HACL1
2473735	HADHB
3466687	HAL
3757288	HAP1
2832052	HARS
2832052	HARS2
2757621	HAUS3
3360401	HBB
2878273	HBEGF
3360441	HBG1
3360441	HBG2
3018420	HPB1
2975287	HBS1L
4026956	HCFC1
3677356	HCFC1R1
3429406	HCFC2
2756309	hCG_1771830
3621728	hCG_1789710
2525852	hCG_2024410
3871459	hCG_2039146
2855963	HCN1
2907173	HCRP1
2328387	HCRT1R1
2606026	HDAC4

TCID	GENE_na29
4012204	HDAC8
2991395	HDAC9
3998444	HDHD1A
2462456	HEATR1
2386867	HEATR1
3659888	HEATR3
3559690	HEATR5A
3765059	HEATR6
3445123	HEBP1
3559570	HECTD1
3257750	HECTD2
2409970	HECTD3
2999334	HECW1
2692909	HEG1
3420497	HELB
3258910	HELLS
2776026	HELQ
3768015	HELZ
3217077	HEMGN
3628650	HERC1
3614901	HERC2
3585749	HERC2
3614901	HERC2P2
3585749	HERC2P2
3614901	HERC2P3
3585749	HERC2P3
2735459	HERC3
3292448	HERC4
2735362	HERC6
3662387	HERPUD1
2605735	HES6
3632152	HEXA
3738842	HEXDC
3774975	HEXDC
3723348	HEXIM1
2924492	HEY2
2422612	HFM1
4047070	HGD
2745899	HHIP
2457496	HHIPL2
2635184	HHLA2
3938113	HIC2
3539070	HIF1A
3836705	HIF3A
2739160	HIGD1A
3057370	HIP1
2352758	HIPK1
3075778	HIPK2
3325907	HIPK3
3952637	HIRA
3952703	HIRA
2822407	HISPPD1
3621276	HISPPD2A
2946194	HIST1H1A
2946714	HIST1H2BK
2946215	HIST1H3B
2899233	HIST1H3E
2899146	HIST1H4C
2947081	HIST1H4L

TCID	GENE_na29
2895159	HIVEP1
2977265	HIVEP2
3250278	HK1
2948926	HLA-B
2948887	HLA-B
2948926	HLA-C
2948887	HLA-C
2950263	HLA-DMB
2950329	HLA-DPA1
2903401	HLA-DPB1
2950125	HLA-DQB2
2903189	HLA-DRA
2901620	HLA-E
3727583	HLF
3881282	HM13
3091848	HMBOX1
3602873	HMG20A
3817040	HMG20B
3431483	HMGA1
3471198	HMGA1
3420316	HMGA2
3994915	HMGB3
2329386	HMGB4
2958117	HMGCCL1
2815965	HMGCR
3944046	HMGXB4
2838656	HMMR
2841802	HMP19
3770606	HN1
3754797	HNF1B
2352609	HNRNPA3
2843619	HNRNPA8
2775463	HNRNPD
3861617	HNRNPL
4037595	HNRNPM
2401275	HNRNPR
2864237	HOMER1
3448152	HOMER1
3096368	HOOK3
3042994	HOXA13
3042816	HOXA4
3761441	HOXB8
2516967	HOXD1
2516853	HOXD9
2400322	HP1BP3
2408041	HPCAL4
2794408	HPGD
3830065	HPN
3991698	HPRT1
2700244	HPS3
2647216	HPS3
3358090	HRAS
2658275	HRASLS
3867708	HRC
2656650	HRG
3145980	HRSP12
2345196	HS2ST1
4022183	HS6ST2
3671448	HSBP1

Fig. 12 Continued

TCID	GENE_na29
3328069	HSD17B12
3216195	HSD17B3
3417703	HSD17B6
2903488	HSD17B8
3120358	HSF1
2923819	HSF2
4025485	HSFX1
4025485	HSFX2
3400034	HSN2
3580179	HSP90AA1
3580179	HSP90AA4P
2908474	HSP90AB1
2908474	HSP90AB3P
3429312	HSP90B1
3429312	HSP90B3P
3874402	HSPA12B
3925439	HSPA13
2902707	HSPA1A
2902707	HSPA1B
2949450	HSPA1L
2828856	HSPA4
2742935	HSPA4L
3225398	HSPA5
2413519	HSPB11
2692136	HSPBAP1
2593733	HSPD1
2400793	HSPG2
3508330	HSPH1
2730194	HTN3
2401251	HTR1D
2916067	HTR1E
3513147	HTR2A
2655325	HTR3C
2655325	HTR3D
2715820	HTT
3049840	HUS1
4009315	HUWE1
3471327	HVCN1
3697434	HYDIN
2977621	HYMA1
3394123	HYOU1
2539821	IAH1
2380991	IARS2
2735129	IBSP
2962525	IBTK
3038065	ICA1
3820443	ICAM1
3766621	ICAM2
3820469	ICAM5
2394588	ICMT
2401493	ID3
3300350	IDE
2597010	IDH1
3603199	IDH3A
3273601	IDI1
3095223	IDO1
3095257	IDO2
4025339	IDS
2756831	IDUA

TCID	GENE_na29
3226883	IER5L
2362394	IFI16
2343511	IFI44
2343473	IFI44L
2403261	IFI6
3257268	IFIT5
3315675	IFITM1
3201242	IFNA10
3201242	IFNA16
3201242	IFNA17
3201242	IFNA21
3201242	IFNA4
3201242	IFNA7
3918447	IFNAR2
3201359	IFNE
2446198	IFRG15
3676002	IFT140
3480411	IFT88
3980455	IGBP1
3610804	IGF1R
2708922	IGF2BP2
3041409	IGF2BP3
2934308	IGF2R
3049292	IGFBP3
2598828	IGFBP5
3415744	IGFBP6
2770469	IGFBP7
2728448	IGFBP7
3836614	IGFL2
3375735	IGHG1
2772566	IGJ
2563785	IGK@
2563785	IGKC
2563785	IGKV1-5
2563785	IGKV3-15
2563785	IGKV3-20
2563785	IGKV3D-11
2563785	IGKV3D-15
3954764	IGLL3
4021777	IGSF1
2701109	IGSF10
3096092	IKBKB
2597867	IKZF2
3755862	IKZF3
2452948	IL10
3167553	IL11RA
3988538	IL13RA1
4018729	IL13RA2
2624565	IL17RB
3391255	IL18
2497119	IL18R1
2571569	IL1F8
2657831	IL1RAP
3972657	IL1RAPL1
2497082	IL1RL1
2497028	IL1RL2
3832906	IL28A
3832906	IL28B
3832906	IL29

TCID	GENE_na29
3275729	IL2RA
2660617	IL5RA
2806468	IL7R
2731332	IL8
2599303	IL8RA
2599303	IL8RB
2527580	IL8RB
2599303	IL8RBP
2527580	IL8RBP
3996971	IL9R
3367965	IMMP1L
2562685	IMMT
3142485	IMPA1
3782166	IMPACT
3136413	IMPAD1
3071700	IMPDH1
2673873	IMPDH2
3262129	INA
3554315	INF2
2536757	ING5
2596162	INO80D
2520113	INPP1
2495446	INPP4A
2787459	INPP4B
3267382	INPP5F
3942766	INPP5J
3740264	INPP5K
3339423	INPL1
2502424	INSIG2
3848243	INSR
3088405	INTS10
2738314	INTS12
3383081	INTS4
3383081	INTS4L1
3383081	INTS4L2
3514488	INTS6
2742829	INTU
2950823	IP6K3
2980449	IPCEF1
3853658	IPMK
2811812	IPO11
3319840	IPO7
2604998	IQCA1
2328767	IQCC
3472274	IQCD
2713664	IQCG
3599059	IQCH
2649824	QCJ
2816298	IQGAP2
2366028	IQWD1
2610359	IRAK2
3412296	IRAK4
2796384	IRF2
2453881	IRF6
2601995	IRS1
3525234	IRS2
4017694	IRS4
2846522	IRX2
3430776	ISCU

Fig. 12 Continued

TCID	GENE_na29
2438482	ISG20L2
2808931	ISL1
3621948	ISLR
2827709	ISOC1
3882854	ITCH
3690193	ITFG1
3401119	ITFG2
3642707	ITFG3
2809128	ITGA1
2356218	ITGA10
3630736	ITGA11
2809245	ITGA2
3759137	ITGA2B
3726154	ITGA3
2518272	ITGA4
2515627	ITGA6
3457101	ITGA7
2617188	ITGA9
3741585	ITGAE
2519229	ITGAV
3852832	ITGB1
3284188	ITGB1
2539765	ITGB1BP1
3724545	ITGB3
2416218	ITGB3BP
2692816	ITGB5
2583465	ITGB6
2991860	ITGB8
3499132	ITGBL1
3276337	ITIH5
4009751	ITIH5L
2837232	ITK
4013549	ITM2A
3874249	ITPA
2608469	ITPR1
3448152	ITPR2
2903782	ITPR3
3918779	ITSN1
3589905	IVD
2448073	IVNS1ABP
2931172	IYD
3160895	JAK2
2880361	JAKMIP2
3916527	JAM2
3248986	JMJD1C
3291682	JMJD1C
2817291	JMY
3672886	JPH3
3157060	JRK
3556990	JUB
2415084	JUN
2945440	KAAG1
3998766	KAL1
2639734	KALRN
3850676	KANK2
3699757	KARS
2978957	KATNA1
3507962	KATNAL1
3787187	KATNAL2
TCID	GENE_na29
2514413	KBTBD10
3389566	KBTBD3
3511168	KBTBD6
3511168	KBTBD7
2628260	KBTBD8
2427619	KCNA3
2648991	KCNAB1
3103062	KCNB2
3462567	KCNC2
3021009	KCND2
2428119	KCND3
2613293	KCNH8
2762944	KCNIP4
2603960	KCNJ13
3920850	KCNJ15
3733238	KCNJ16
3733275	KCNJ2
3931495	KCNJ6
2385873	KCNK1
2379974	KCNK2
3334446	KCNK4
2952834	KCNK5
3296046	KCNMA1
2653673	KCNMB2
3421985	KCNMB4
3824666	KCNN1
2436826	KCNN3
3317352	KCNQ1
3154002	KCNQ3
2471384	KCNS3
3071878	KCP
3470793	KCTD10
3687308	KCTD13
3383130	KCTD14
3944637	KCTD17
2594435	KCTD18
2380055	KCTD3
3128372	KCTD9
3128372	KCTD9P2
2827525	KDEL1
3867092	KDELR1
3945314	KDELR3
3817733	KDM4B
3161566	KDM4C
3439603	KDM5A
2451309	KDM5B
4009062	KDM5C
3709153	KDM6B
2769810	KDR
3593408	KGFLP1
3593408	KGFLP2
2960774	KHDC1
2959039	KHDRBS2
3117384	KHDRBS3
3847814	KHSRP
2399620	KIAA0090
3750872	KIAA0100
3629103	KIAA0101
2832963	KIAA0141
TCID	GENE_na29
3672368	KIAA0182
2713555	KIAA0226
3542145	KIAA0247
3554452	KIAA0284
2406139	KIAA0319L
3829638	KIAA0355
2973232	KIAA0408
3681956	KIAA0430
2338719	KIAA0485
2393654	KIAA0495
3193870	KIAA0649
3329404	KIAA0652
3456955	KIAA0748
3743119	KIAA0753
2331213	KIAA0754
3565739	KIAA0831
3301263	KIAA0894
2437753	KIAA0907
2748198	KIAA0922
3429754	KIAA1033
3255402	KIAA1128
3238962	KIAA1217
3884640	KIAA1219
2927604	KIAA1244
2432851	KIAA1245
3250093	KIAA1279
3737488	KIAA1303
3529951	KIAA1305
2350489	KIAA1324
3059942	KIAA1324L
3784999	KIAA1328
2689286	KIAA1407
3549264	KIAA1409
3161167	KIAA1432
3791168	KIAA1468
3075431	KIAA1549
2911257	KIAA1586
3308489	KIAA1598
3702499	KIAA1609
3737274	KIAA1618
3806126	KIAA1632
3961496	KIAA1659
2735815	KIAA1680
3854836	KIAA1683
2588319	KIAA1715
3164601	KIAA1797
2484457	KIAA1841
3258168	KIF11
3221822	KIF12
3898796	KIF16B
3734292	KIF19
2319661	KIF1B
3450775	KIF21A
3203935	KIF24
3212232	KIF27
2875419	KIF3A
3418298	KIF5A
3283991	KIF5B
2509900	KIF5C

Fig. 12 Continued

TCID	GENE_na29
2672629	KIF9
3841756	KIR2DL1
3841777	KIR2DL1
4053056	KIR2DL1
3841756	KIR2DL2
3841777	KIR2DL2
4053056	KIR2DL2
3841756	KIR2DL3
3841777	KIR2DL3
4053056	KIR2DL3
3841777	KIR2DL4
3841756	KIR2DS1
3841777	KIR2DS1
4053056	KIR2DS1
3841756	KIR2DS2
3841777	KIR2DS2
4053056	KIR2DS2
3841756	KIR2DS3
4053056	KIR2DS3
3841756	KIR2DS4
3841777	KIR2DS4
4053056	KIR2DS4
3841756	KIR2DS5
3841777	KIR2DS5
4053056	KIR2DS5
3841756	KIR3DL1
3841777	KIR3DL1
4053056	KIR3DL1
3841756	KIR3DL2
3841777	KIR3DL2
3841777	KIR3DL3
3841777	KIR3DP1
3841777	KIR3DP1
4053056	KIR3DP1
4053056	KIR3DP1
3841777	KIR3DS1
4053056	KIR3DS1
3841357	KIR3DX1
2727587	KIT
2724308	KLB
3553872	KLC1
3836217	KLC3
2954355	KLC4
3851840	KLF1
2723997	KLF3
3219215	KLF4
3493543	KLF5
3274361	KLF6
2596514	KLF7
3978943	KLF8
3534886	KLHDC1
2907568	KLHDC3
2452440	KLHDC8A
2622006	KLHDC8B
3721485	KLHL10
2451428	KLHL12
2750527	KLHL2
3825260	KLHL26
3562671	KLHL28
TCID	GENE_na29
2918037	KLHL32
3983324	KLHL4
2708066	KLHL6
2776998	KLHL8
3868828	KLK10
3868841	KLK11
3868857	KLK12
3868783	KLK7
3444180	KLRA1
3443804	KLRB1
3444117	KLRC1
3444117	KLRC2
3444117	KLRC3
3444086	KLRC4
3404660	KLRD1
3404030	KLRG1
3444086	KLRK1
3272566	KNDC1
3435362	KNTC1
2691982	KPNA1
4041113	KPNA2
2703217	KPNA4
2922840	KPNA5
3724782	KPNB1
3030585	KPNB1
3447863	KRAS
3422855	KRR1
3462693	KRR1
3415576	KRT18
3757108	KRT19
3415320	KRT7
3455752	KRT77
3455890	KRT79
3455516	KRT8
3455134	KRT80
3455344	KRT82
3923838	KRTAP10-12
3923838	KRTAP10-4
3923838	KRTAP10-6
3923838	KRTAP10-7
3934623	KRTAP12-1
3934623	KRTAP12-2
3923881	KRTAP12-3
3721279	KRTAP9-2
3721279	KRTAP9-3
3721279	KRTAP9-4
3721279	KRTAP9-8
3721279	KRTAP9-9
2696415	KY
2508520	KYNU
2339334	L1TD1
3797295	L3MBTL4
3597421	LACTB
3841506	LAIR1
3841357	LAIR2
3841506	LAIR2
2925237	LAMA2
3913018	LAMA5
3067302	LAMB1
TCID	GENE_na29
2674047	LAMB2
2453793	LAMB3
3067408	LAMB4
2371065	LAMC1
2371139	LAMC2
4019849	LAMP2
2707876	LAMP3
3973768	LANCL3
2836738	LARP1
2743085	LARP2
3414512	LARP4
2740005	LARP7
4010860	LAS1L
3855285	LASS1
3454296	LASS5
2514122	LASS6
3008164	LAT2
2978989	LATS1
3504526	LATS2
3348568	LAYN
2458289	LBR
2382781	LBR
3599280	LBXCOR1
2962026	LCA5
3696035	LCAT
3653619	LCMT1
3190190	LCN2
2762500	LCORL
3512874	LCP1
2886595	LCP2
2577856	LCT
3304215	LDB1
2762088	LDB2
3446868	LDHB
2400793	LDLRAD2
4024420	LDOC1
3963289	LDOC1L
2828796	LEAP2
2649182	LEKR1
2452478	LEMD1
3420079	LEMD3
2340433	LEPR
2340433	LEPROT
3414776	LETMD1
3944882	LGALS1
3833183	LGALS13
3960174	LGALS2
3536706	LGALS3
3861557	LGALS4
2462456	LGALS8
2386867	LGALS8
3258713	LGI1
3127352	LGI3
3859668	LGI4
3577078	LGMN
3367096	LGR4
3422144	LGR5
2375144	LGR6
3867573	LHB

Fig. 12 Continued

TCID	GENE_na29
2552368	LHCGR
2863885	LHFPL2
3224220	LHX6
2342475	LHX8
2854092	LIFR
3718401	LIG3
3841545	LILRA1
3841357	LILRA2
3841545	LILRB1
3454331	LIMA1
2725061	LIMCH1
3893642	LIME1
3008108	LIMK1
3942838	LIMK2
2574984	LIMS2
2838598	LIN28B
2775858	LIN54
3367183	LIN7C
3202528	LINGO2
3787855	LIPG
2708855	LIPH
2868265	LIX1
2356142	LIX1L
3460584	LLPH
2565484	LMAN2L
3081624	LMBR1
2660029	LMLN
2361279	LMNA
2985342	LMNA
3368814	LMO2
3446137	LMO3
2345286	LMO4
2451043	LMOD1
3008376	LOC100093631
3028011	LOC100124692
2366798	LOC100127910
3830216	LOC100127972
3442054	LOC100127974
3795184	LOC100127994
2321238	LOC100128068
3913018	LOC100128184
3777470	LOC100128219
3302056	LOC100128320
3872604	LOC100128398
3817733	LOC100128439
3836614	LOC100128529
2933175	LOC100128551
2606026	LOC100128563
3848243	LOC100128567
3890870	LOC100128608
3020222	LOC100128868
2321466	LOC100129042
3396107	LOC100129069
3765689	LOC100129112
3759587	LOC100129115
2601414	LOC100129171
2900269	LOC100129195
2404122	LOC100129196
2869880	LOC100129233
TCID	GENE_na29
2604998	LOC100129258
2691014	LOC100129275
2587520	LOC100129312
3696317	LOC100129324
3623031	LOC100129397
2969886	LOC100129399
3050388	LOC100129427
2889916	LOC100129453
2746693	LOC100129572
3090436	LOC100129717
3128271	LOC100129717
3290746	LOC100129721
3191273	LOC100129785
3400730	LOC100129797
3092415	LOC100129846
2409904	LOC100129897
3833183	LOC100129935
2577482	LOC100129961
3737697	LOC100130078
2563785	LOC100130100
2431595	LOC100130131
3439836	LOC100130219
2504328	LOC100130248
3742351	LOC100130311
2462693	LOC100130331
2614369	LOC100130354
2774817	LOC100130356
2942306	LOC100130357
3856720	LOC100130518
3826803	LOC100130518
3542145	LOC100130542
2662581	LOC100130542
3935016	LOC100130597
3739867	LOC100130876
3987029	LOC100130886
3946146	LOC100130899
2540317	LOC100130910
3275922	LOC100130920
2861952	LOC100130998
2602901	LOC100131015
2533670	LOC100131101
3110272	LOC100131102
3848243	LOC100131165
3814701	LOC100131178
2542795	LOC100131373
3754041	LOC100131384
3754096	LOC100131384
4040117	LOC100131384
4040849	LOC100131384
4025339	LOC100131434
3464860	LOC100131490
3868330	LOC100131519
3875642	LOC100131599
3227846	LOC100131612
3259367	LOC100131720
2994835	LOC100131724
3756193	LOC100131821
3754677	LOC100131822
2779638	LOC100131829
TCID	GENE_na29
2622026	LOC100131840
4027176	LOC100131857
2984884	LOC100131869
3155937	LOC100131910
2364677	LOC100131938
3513549	LOC100131993
3681705	LOC100131998
3684782	LOC100131998
3684548	LOC100131998
2853642	LOC100132000
2824902	LOC100132014
3583541	LOC100132025
2984275	LOC100132188
4042198	LOC100132235
3650300	LOC100132247
3652077	LOC100132247
3922793	LOC100132338
3275132	LOC100132353
2431112	LOC100132495
3552729	LOC100132532
4040117	LOC100132544
2641341	LOC100132731
3392332	LOC100132764
2321466	LOC100132942
3855538	LOC100133072
3636879	LOC100133144
3017547	LOC100133169
3298924	LOC100133190
2342220	LOC100133219
3735847	LOC100133227
3930781	LOC100133286
3279575	LOC100133308
3056656	LOC100133748
3309936	LOC100133773
3171425	LOC100133802
3841357	LOC100133875
3746881	LOC100133918
2565935	LOC100133923
3223425	LOC100133950
3976240	LOC100133957
3514879	LOC100134095
3759849	LOC100134130
3287366	LOC100134152
3227846	LOC100134189
3274758	LOC100134257
3754041	LOC100134348
3754096	LOC100134348
4040117	LOC100134348
4040849	LOC100134348
3050462	LOC100134631
3636879	LOC100134869
3325503	LOC100190939
3652077	LOC100190986
3279108	LOC100192204
3260001	LOC100270710
3683050	LOC100271836
3724591	LOC100272146
2993590	LOC100272146
4025339	LOC100272228

Fig. 12 Continued

TCID	GENE_na29	TCID	GENE_na29	TCID	GENE_na29
3454147	LOC100286844	4026560	LOC401218	2563785	LOC652493
2881607	LOC134466	2994558	LOC401317	3642162	LOC652595
3710681	LOC139201	3060994	LOC401387	2563785	LOC652694
3542847	LOC145474	2988594	LOC402509	3692701	LOC652708
3840857	LOC147804	3358950	LOC402778	2565935	LOC652726
2468105	LOC150622	3244055	LOC439911	3841777	LOC652779
2564816	LOC150759	3770944	LOC440093	2858134	LOC653198
2575897	LOC150776	3583541	LOC440243	3937967	LOC653264
2687739	LOC151657	3614901	LOC440248	3953456	LOC653264
2880361	LOC153469	3585749	LOC440248	4015838	LOC653354
2808180	LOC153684	4002011	LOC440258	3754041	LOC653382
3220846	LOC158402	2593796	LOC440258	3754096	LOC653382
3765167	LOC162632	3652077	LOC440354	4040117	LOC653382
3748400	LOC162632	3683050	LOC440354	4040849	LOC653382
3278234	LOC168474	3724698	LOC440434	3681705	LOC653390
3402978	LOC171220	4040117	LOC440434	3684782	LOC653390
3838809	LOC199800	2563785	LOC440871	3684548	LOC653390
2927604	LOC202451	3173508	LOC440897	3724698	LOC653498
3107606	LOC203107	2727226	LOC441016	3754041	LOC653498
3280787	LOC219688	3227846	LOC441089	3754096	LOC653498
3514879	LOC220115	2995076	LOC441208	4040117	LOC653498
3533499	LOC220429	3029230	LOC441294	4040849	LOC653498
3765167	LOC220594	3641597	LOC441734	3206317	LOC653501
3748400	LOC220594	2431595	LOC51152	2431886	LOC653513
2906607	LOC221442	3948259	LOC553158	3591281	LOC653566
3937967	LOC26080	3201345	LOC554202	2470470	LOC653602
3953456	LOC26080	3819016	LOC554363	2575949	LOC654264
3523499	LOC283480	3652077	LOC595101	2501317	LOC654433
3487095	LOC283508	3683050	LOC595101	3358112	LOC692247
3542847	LOC283567	3060994	LOC613126	4046876	LOC727761
3613338	LOC283683	3652077	LOC641298	2628482	LOC727775
3583541	LOC283767	3683050	LOC641298	3841756	LOC727787
3941907	LOC284685	3938817	LOC642311	3841777	LOC727787
3937967	LOC284861	2938196	LOC642335	3583541	LOC727832
3953456	LOC284861	2506335	LOC642669	4026560	LOC727895
2571075	LOC285074	3674504	LOC643224	2911903	LOC727916
2781325	LOC285456	3405207	LOC643287	3635776	LOC727963
2749380	LOC285505	2455418	LOC643454	2390518	LOC728060
2791419	LOC285505	3281068	LOC643475	3504691	LOC728099
2951567	LOC285847	3329069	LOC644172	3937967	LOC728212
3031466	LOC285972	2506335	LOC644525	3953456	LOC728212
3139882	LOC286190	3982423	LOC644732	3721279	LOC728341
3650300	LOC339047	3982423	LOC644756	3029646	LOC728377
2404958	LOC339483	3924929	LOC645159	3683050	LOC728423
3205834	LOC340508	3583541	LOC645162	3471769	LOC728543
2670903	LOC348817	2340350	LOC645195	2888103	LOC728554
3686587	LOC388242	3583541	LOC645202	2487995	LOC728731
3954729	LOC388882	3171425	LOC645961	2758602	LOC728731
2472955	LOC388931	3283613	LOC645984	3223425	LOC728779
3017080	LOC389137	3629811	LOC646358	3754041	LOC728824
3212232	LOC389765	3203311	LOC646808	3754096	LOC728824
3851651	LOC389791	3826803	LOC646864	4040117	LOC728824
3704567	LOC390748	3621276	LOC647471	4040849	LOC728824
2539869	LOC392510	3868330	LOC647678	3754070	LOC728830
3826656	LOC400680	2525533	LOC648149	2431886	LOC728989
2421883	LOC400759	2714200	LOC649851	4040117	LOC729034
3960478	LOC400927	3621276	LOC649956	2619521	LOC729085
2565935	LOC400986	2563785	LOC650405	2743085	LOC729231
2576526	LOC401010	3939914	LOC650860	2565753	LOC729234

Fig. 12 Continued

TCID	GENE_na29
2494064	LOC729234
2500275	LOC729312
3937967	LOC729461
3953456	LOC729461
3683050	LOC729513
3652077	LOC729602
2934308	LOC729603
4002011	LOC729609
2343823	LOC729828
2548274	LOC729862
3855538	LOC729991
3855538	LOC729991-MEF2B
2936857	LOC730031
3681705	LOC730092
3684548	LOC730092
3826803	LOC730110
2758602	LOC731528
3717052	LOC731788
2390518	LOC731985
3721279	LOC732428
4031834	LOC732447
3651509	LOC81691
3248986	LOC84989
2788800	LOC90826
3954764	LOC91316
3938817	LOC96610
3847356	LONP1
3659306	LONP2
3690470	LONP2
2567167	LONRF2
3988638	LONRF3
2872848	LOX
3302693	LOXL4
3220384	LPAR1
3442137	LPAR5
2845973	LPCAT1
3661718	LPCAT2
3442427	LPCAT3
2343823	LPHN2
2728938	LPHN3
3088486	LPL
3374402	LPXN
2789266	LRBA
3488602	LRCH1
2659918	LRCH3
3336576	LRFN4
3846831	LRG1
3025291	LRGUK
2680591	LRIG1
3459120	LRIG3
2739289	LRIT3
3408505	LRMP
3528895	LRP10
2979111	LRP11
2979187	LRP11
3147985	LRP12
2578790	LRP18
2586038	LRP2
TCID	GENE_na29
2796847	LRP2BP
3829242	LRP3
3158812	LRRC14
2898746	LRRC16A
2672190	LRRC2
3402984	LRRC23
3158812	LRRC24
2704763	LRRC34
3665458	LRRC36
3767169	LRRC37A
3767169	LRRC37A2
3767169	LRRC37A3
3767169	LRRC37A4
3717452	LRRC37B
3717452	LRRC37B2
2425173	LRRC39
2417737	LRRC40
3712835	LRRC48
3600212	LRRC49
3370269	LRRC4C
3671607	LRRC50
2365086	LRRC52
3358090	LRRC56
3358112	LRRC56
3315907	LRRC56
3154136	LRRC6
3106559	LRRC69
2341387	LRRC7
2811812	LRRC70
3190778	LRRC8A
2345880	LRRC8B
2345929	LRRC8C
3538403	LRRC9
3105430	LRRC1
2534456	LRRFIP1
2669184	LRRFIP2
3424705	LRRIQ1
2418339	LRRIQ3
2608309	LRRN1
3019158	LRRN3
3380996	LRTOMT
2690012	LSAMP
3846831	LSDP5
3131844	LSM1
3131205	LSM12
2746269	LSM6
3317071	LSP1
3830246	LSR
3529877	LTB4R
3529877	LTB4R2
2476510	LTBP1
3571944	LTBP2
2672140	LTF
3674960	LUC7L
3465248	LUM
3324162	LUZP2
2824902	LVRN
3119339	LY6E
2902559	LY6G5B
TCID	GENE_na29
2949230	LY6G5C
3119213	LY6K
2583254	LY75
2363248	LY9
2566689	LYG2
3098977	LYN
2576988	LYPD1
2506570	LYPD1
2510056	LYPD6
2509988	LYPD6B
3135567	LYPLA1
2380785	LYPLA1
3683845	LYRM1
2916825	LYRM2
3447863	LYRM5
2828135	LYRM7
2461999	LYST
3362826	LYVE1
3421511	LYZ
3403981	M6PR
3040518	MACC1
2331213	MACF1
2576281	MAD2L1
2396781	MAD2L2
2714672	MAEA
3905875	MAFB
3451814	MAFG
2708610	MAGEF1
3058209	MAGI2
2352609	MAGI3
3745525	MAGOHB
3444195	MAGOHB
3130823	MAK16
7385515	MALAT1
3790259	MALT1
3174121	MAMDC2
2786732	MAML3
3994710	MAMLD1
2971801	MAN1A1
2353881	MAN1A2
2326049	MAN1C1
2823551	MAN2A1
3608466	MAN2A2
3633347	MAN2C1
2779897	MANBA
3884158	MANBAL
2917767	MANEA
2330843	MANEAL
3975227	MAOA
4006210	MAOB
2814756	MAP1B
2525533	MAP2
3710681	MAP2K4
3599162	MAP2K5
3733065	MAP2K6
2574798	MAP3K2
2403027	MAP3K6
2930592	MAP3K7IP2
2672966	MAP4

Fig. 12 Continued

TCID	GENE_na29
2549260	MAP4K3
2496727	MAP4K4
2330289	MAP7D1
3954238	MAPK1
2776670	MAPK10
2904946	MAPK13
2904877	MAPK14
3788097	MAPK4
3111561	MAPK6
3713874	MAPK7
3329069	MAPK8IP1
2890605	MAPK9
2622912	MAPKAPK3
3784344	MAPRE2
2792166	MARCH1
2801608	MARCH6
3553690	MARK3
3260001	MARVELD1
3667652	MARVELD3
2396415	MASP2
2812690	MAST4
3240012	MASTL
2404122	MATN1
3108526	MATN2
3568667	MAX
3808600	MBD2
2694785	MBD4
3497586	MBNL2
2539607	MBOAT2
3870570	MBOAT7
3762625	MBTD1
3971329	MBTPS2
3674504	MC1R
4024092	MCF2
3924518	MCM3AP
3924518	MCM3APAS
3097152	MCM4
3063685	MCM7
3083778	MCPH1
3609592	MCTP2
2948564	MDC1
3019981	MDFIC
2485176	MDH1
2596386	MDH1B
3329343	MDK
3461164	MDM1
3421300	MDM2
2376037	MDM4
2964350	MDN1
2962820	ME1
3385307	ME3
2847264	MED10
2647898	MED12L
3765689	MED13
4037595	MED13L
3473083	MED13L
3228635	MED22
2974188	MED23
3838947	MED25

TCID	GENE_na29
3227846	MED27
3832964	MED29
3112713	MED30
3707990	MED31
3855538	MEF2B
2866225	MEF2C
2827299	MEGF10
3223551	MEGF9
2486178	MEIS1
3618333	MEIS2
3168508	MELK
2547332	MEMO1
2547332	MEMO1
3039485	MEOX2
3783788	MEP1B
2500550	MERTK
3020343	MET
3527745	METT11D1
3311342	METTL10
3191074	METTL11A
3375999	METTL12
3556418	METTL3
3796244	METTL4
3414739	METTL7A
3416895	METTL7B
3651955	METTL9
3808096	MEX3C
3845296	MEX3D
3748798	MFAP4
3638204	MGFE8
2653932	MFN1
3735623	MFSD11
2331679	MFSD2
2376548	MFSD4
2520138	MFSD6
2714200	MFSD7
2785035	MFSD8
3028011	MGAM
2890859	MGAT1
3534866	MGAT2
3226883	MGAT2
2566414	MGAT4A
2890239	MGAT4B
3464417	MGAT4C
2668425	MGC12488
4022690	MGC16121
3705135	MGC16385
2658275	MGC2889
2877893	MGC29506
3528172	MGC40069
3011838	MGC87042
3304012	MGEA5
3445741	MGP
3646277	MGRN1
3406589	MGST1
4042837	MIA3
3320717	MICAL2
3320819	MICALCL
3999395	MID1

TCID	GENE_na29
2685776	MINA
3742351	MINK1
3950846	MIOX
3496366	MIRHG1
3707759	MIS12
2628785	MITF
2383118	MIXL1
3312490	MKI67
3961496	MKL1
3649052	MKL2
3076178	MKRN1
3282463	MKX
3434525	MLEC
2649609	MLF1
2796510	MLF1IP
2616932	MLH1
3572235	MLH3
2734784	MLL
3847703	MLL
3453592	MLL2
3924929	MLL3
3017547	MLL5
3847703	MLLT1
3238231	MLLT10
2358693	MLLT11
3200982	MLLT3
2936857	MLLT4
3721886	MLX
2746164	MMAA
2334350	MMACHC
3763270	MMD
3388807	MMP1
3388893	MMP13
3528864	MMP14
3143643	MMP16
3143660	MMP16
3883236	MMP24
3753760	MMP28
2735759	MMRN1
3298924	MMRN2
3302360	MMS19
2748163	MND1
2362333	MNDA
3625761	MNS1
3202316	MOBK12B
3784783	MOCOS
3064501	MOGAT3
2674963	MON1A
3919952	MORC3
4017212	MORC4
2794180	MORF4
2794180	MORF4L1
3475295	MORN3
2381309	MOSC1
4022833	MOSPD1
3708874	MPDU1
3198974	MPDZ
3480013	MPHOSPH8
3476012	MPHOSPH9

Fig. 12 Continued

TCID	GENE_na29
2333107	MPL
2594987	MPP4
3541073	MPP5
2993206	MPP6
3282601	MPP7
3799167	MPPE1
3367673	MPPED2
3712363	MPRIP
2545653	MPV17
3393720	MPZL2
2370317	MR1
2915571	MRAP2
2598496	MREG
3807965	MRO
3480681	MRP63
3760945	MRPL10
2955025	MRPL14
3361116	MRPL17
3379708	MRPL21
3619479	MRPL42P5
3218067	MRPL50
2664099	MRPS25
2861952	MRPS27
2808612	MRPS30
3510925	MRPS31
3409330	MRPS35
3734760	MRPS7
3188050	MRRF
2898452	MRS2
3332403	MS4A1
3374934	MS4A4E
3374934	MS4A6A
2480992	MSH2
2817837	MSH3
2902633	MSH5
2696764	MSL2
3979659	MSN
3125571	MSR1
2674602	MST1
2835662	MST150
3992408	MST159
2674602	MSTP2
2674602	MSTP9
3662106	MT1A
3662106	MT1DP
3662139	MT1E
3662201	MT1F
3692999	MT1G
3662201	MT1H
3662236	MT1IP
3662139	MT1JP
3662130	MT1L
3662150	MT1M
3662201	MT1P2
3662150	MT1P3
3662106	MT1X
3662247	MT1X
3662236	MT1X
3662106	MT2A
TCID	GENE_na29
2952102	MTCH1
3060917	MTERF
2607020	MTERFD2
3629378	MTFMT
3101385	MTFR1
3540007	MTHFD1
2931391	MTHFD1L
2731417	MTHFD2L
3482888	MTIF3
3994795	MTM1
3994846	MTMR1
3615985	MTMR10
2434178	MTMR11
3615985	MTMR15
3764471	MTMR4
3511189	MTRF1
3151970	MTSS1
3125915	MTUS1
2437118	MUC1
3366903	MUC15
3986168	MUM1L1
2676454	MUSTN1
3259019	MUTYH
3655723	MVP
3997825	MXRA5
2512701	MXRA7
3841076	MYADM
2547751	MYADML
3428596	MYBPC1
3372129	MYBPC3
2407755	MYCBP
3518496	MYCBP2
2470838	MYCN
2931970	MYCT1
3622934	MYEF2
3593014	MYEF2
3338060	MYEOV
3745351	MYH1
3744463	MYH10
3682028	MYH11
3839206	MYH14
3745429	MYH2
3745351	MYH2
3745429	MYH3
3745429	MYH4
3745351	MYH4
3557504	MYH6
3557504	MYH7
3745429	MYH8
3745351	MYH8
3959451	MYH9
2597389	MYL1
2714200	MYL5
2651671	MYNN
2850071	MYO10
2520429	MYO1B
3752709	MYO1D
3626826	MYO1E
3431071	MYO1H
TCID	GENE_na29
3624607	MYO5A
3807595	MYO5B
2914070	MYO6
3631794	MYO9A
2443952	MYOC
3300597	MYOF
3796428	MYOM1
2835619	MYOZ3
3725779	MYST2
3133135	MYST3
2724585	N4BP2
3508696	N4BP2L2
2773872	NAAA
3344142	NAALAD2
2653114	NAALADL2
2520225	NAB1
3417809	NAB2
3695268	NAE1
3074912	NAG20
2861616	NAIP
3066818	NAMPT
4012511	NAP1L2
4014759	NAP1L3
3359469	NAP1L4
2777447	NAP1L5
3901191	NAPB
3778823	NAPG
3868400	NAPSA
3868400	NAPSB
3486883	NARG1L
3627363	NARG2
3809671	NARS
2334404	NASP
4006210	NAT13
3878934	NAT5
3323052	NAV2
2541230	NBAS
3485292	NBEA
2506335	NBEA
2523478	NBEAL1
2621122	NBEAL2
2752085	NBLA00301
2432851	NBPF1
2357217	NBPF1
2432851	NBPF10
2357217	NBPF10
2432851	NBPF11
2357217	NBPF11
2432851	NBPF12
2357217	NBPF12
2432851	NBPF14
2357217	NBPF14
2432851	NBPF15
2357217	NBPF15
2432851	NBPF16
2357217	NBPF16
2432851	NBPF20
2357217	NBPF20
2432851	NBPF8

Fig. 12 Continued

TCID	GENE_na29
2357217	NBPF8
2432851	NBPF9
2357217	NBPF9
3147173	NCALD
3349293	NCAM1
3915936	NCAM2
3402571	NCAPD2
2762500	NCAPG
3181302	NCBP1
2713074	NCBP2
2590736	NCKAP1
2673648	NCKIPSD
2473149	NCOA1
3139722	NCOA2
3887635	NCOA3
3903525	NCOA6
2924514	NCOA7
3746881	NCOR1
3839880	NCRNA00085
3656635	NCRNA00095
2762944	NCRNA00099
4028207	NCRNA00105
2390976	NCRNA00115
3879372	NCRNA00153
2363084	NCSTN
4037595	ND1
3776139	NDC80
3495076	NDFIP2
3615556	NDNL2
3555736	NDRG2
3294816	NDST2
3825838	NDUFA13
3458400	NDUFA4L2
2965674	NDUFAF4
3341497	NDUFC2
2363525	NDUFS2
2331178	NDUFS5
3778207	NDUVF2
3922921	NDUVF3
2581000	NEB
2510485	NEB
3106479	NECAB1
3625539	NEDD4
3789947	NEDD4L
3128271	NEFL
3090436	NEFM
2418078	NEGR1
3086181	NEIL2
2666884	NEK10
2642441	NEK11
3715809	NEK8
3451814	NELL2
2379068	NENF
2949471	NEU1
2343231	NEXN
3717052	NF1
3696666	NFAT5
3795184	NFATC1
3666033	NFATC3

TCID	GENE_na29
3456666	NFE2
2588827	NFE2L2
3724591	NFE2L3
2993590	NFE2L3
2338719	NFIA
3199207	NFIB
3214451	NFIL3
3561039	NFKBIA
2634091	NFKBIZ
3904189	NFS1
2558118	NFU1
3166880	NFX1
2906607	NFYA
2332013	NFYC
3529156	NGDN
3985534	NGFRAP1
2666566	NGLY1
2780060	NHEDC1
2780099	NHEDC2
2363128	NHLH1
3265047	NHLRC2
2674526	NICN1
2462160	NID1
3564620	NID2
3693083	NIP30
3613338	NIPA1
3613300	NIPA2
3146012	NIPAL2
2325410	NIPAL3
2806799	NIPBL
3956909	NIPSNAP1
3182957	NIPSNAP3A
3182957	NIPSNAP3B
3218528	NIPSNAP3B
2623859	NISCH
2363389	NIT1
2924081	NKAIN2
2549092	NKAP
2666103	NKRAS1
2619344	NKTR
3980835	NLG3
3715368	NLK
2812359	NLN
3662444	NLRC5
3319018	NLRP14
3636985	NMB
2877231	NME5
2673136	NME6
2447324	NMNAT2
2697902	NMNAT3
3723264	NMT1
3279108	NMT2
2770039	NMU
2882325	NMUR2
2576526	NOC2L
3439013	NOC4L
3044072	NOD1
3803628	NOL4
2942504	NOL7

TCID	GENE_na29
3214749	NOL8
3261492	NOLC1
3649320	NOMO1
3649320	NOMO2
3649320	NOMO3
3980887	NONO
2758076	NOP14
3442024	NOP2
2364016	NOS1AP
2514216	NOSTRIN
2431112	NOTCH2
2431112	NOTCH2NL
2949901	NOTCH4
3113202	NOV
3558745	NOVA1
4015481	NOX1
3385951	NOX4
3527514	NP
3347658	NPAT
3801411	NPC1
3801411	NPC1
3571904	NPC2
3230697	NPDC1
3891006	NPEPL1
3724698	NPEPPS
4040117	NPEPPS
2695648	NPHP3
3650300	NPIP
3652077	NPIP
3652077	NPIPL1
3652077	NPIPL3
2370926	NPL
3774029	NPLOC4
3404436	NPM1
3089049	NPM2
2738378	NPNT
2359780	NPR1
2805635	NPR3
2792127	NPY1R
3696666	NQO1
3756046	NR1D1
2614142	NR1D2
3855596	NR2C2AP
3610110	NR2F2
2879312	NR3C1
3415229	NR4A1
3181976	NR4A3
3225096	NR6A1
3067478	NRCAM
3023565	NRF1
3925639	NRIP1
2940145	NRN1
3284302	NRP1
2524301	NRP2
2898371	NRSN1
2552643	NRXN1
3545634	NRXN3
3088048	NSAP11
2842951	NSD1

Fig. 12 Continued

TCID	GENE_na29
3724197	NSF
3894637	NSFL1C
2847292	NSUN2
2334706	NSUN4
3280224	NSUN6
2724853	NSUN7
3045004	NT5C3
3757399	NT5C3L
2915828	NT5E
3143330	NTAN1
3356539	NTM
3466556	NTN4
2361761	NTRK1
3177111	NTRK2
3637818	NTRK3
2452405	NUAK2
3644162	NUBP2
3531355	NUBPL
3322251	NUCB2
2510713	NUDC
3148796	NUDCD1
4008427	NUDT10
4008427	NUDT11
3581485	NUDT14
2642543	NUDT16P
2432647	NUDT17
2951087	NUDT3
2355615	NUDT4
3235461	NUDT5
2742109	NUDT6
2364438	NUF2
3512449	NUFIP1
3571347	NUMB
2943808	NUP153
4017281	NUP62CL
3662265	NUP93
3359910	NUP98
3654699	NUPR1
2923060	NUS1
3590388	NUSAP1
3665857	NUTF2
3376155	NXF1
3739867	NXN
3305017	OBFC1
3417485	OBFC2B
2383999	OBSCN
2600237	OBSL1
3614774	OCA2
2768654	OCIAD2
2375338	OCR1
2730281	ODAM
2421121	ODF2L
4020655	ODZ1
3662041	OGFOD1
3214800	OGN
3981120	OGT
2587618	OLA1
3490892	OLFM4
3319119	OLFML1

TCID	GENE_na29
3444043	OLR1
2658346	OPA1
3301782	OPALIN
3399004	OPCML
3181976	OPHN1
2932219	OPRM1
3353914	OR10D1P
3353876	OR10G4
3353876	OR10G7
3353876	OR10G8
3353876	OR10G9
2465806	OR14A16
3188186	OR1Q1
2390180	OR2AJ1
2900832	OR2H1
2465890	OR2T12
4045946	OR2T3
2465890	OR2T33
4045946	OR2T34
2465890	OR2T8
2390180	OR2W3
3706617	OR3A4
3329983	OR4B1
3332008	OR4D6
2487995	OR7E5P
2487995	OR7E91P
3380065	ORAOV1
2594569	ORC2L
2580304	ORC4L
3658925	ORC6L
2591942	ORMDL1
3416977	ORMDL2
3418436	OS9
2693217	OSBPL11
3801621	OSBPL1A
3041875	OSBPL3
2517588	OSBPL6
3555461	OSGEP
2807359	OSMR
2542420	OSR1
2968144	OSTM1
2758602	OTOP1
2788195	OTUD4
2434233	OTUD7B
2485112	OTX1
3566176	OTX2
2331511	OXCT2
2408095	OXCT2
2612625	OXNAD1
3452664	P11
2701071	P2RY13
3513514	P2RY5
3294159	P4HA1
2875193	P4HA2
4041342	P4HB
2935311	PACRG
2720732	PACRGL
3861978	PAF1
2831209	PAIP2

TCID	GENE_na29
2659577	PAK2
3987228	PAK3
3184408	PALM2
3184408	PALM2-AKAP2
2822215	PAM
3457696	PAN2
3483159	PAN3
3874533	PANK2
3950602	PANX2
3543539	PAPLN
2484305	PAPOLG
2780999	PAPSS1
3256590	PAPSS2
2774817	PAQR3
2438016	PAQR6
2983142	PARK2
3441011	PARP11
2639054	PARP14
3527418	PARP2
3505781	PARP4
2808748	PARP8
2692060	PARP9
2811145	PART1
3320865	PARVA
3374746	PATL1
2571608	PAX8
3532793	PAX9
2976417	PBOV1
2364677	PBX1
3855818	PBX4
2829589	PCBD2
2675801	PCBP4
2696802	PCCB
2743800	PCDH10
2878987	PCDH12
3516228	PCDH20
2722823	PCDH7
3516639	PCDH9
2832423	PCDH10
2832431	PCDH11
2832439	PCDH12
2832392	PCDH13
2832447	PCDH13
2832325	PCDH14
2832459	PCDH14
2832297	PCDH14
2832392	PCDH16
2832467	PCDH18
2832325	PCDH19P
2832310	PCDH2
2832447	PCDH2
2832297	PCDH2
2832310	PCDH3
2832447	PCDH3
2832297	PCDH3
2832315	PCDH4
2832325	PCDH5
2832355	PCDH6
2832325	PCDH9

Fig. 12 Continued

TCID	GENE_na29
2832533	PCDHGA1
2832533	PCDHGA10
2832533	PCDHGA11
2832533	PCDHGA12
2832533	PCDHGA2
2832533	PCDHGA3
2832533	PCDHGA4
2832533	PCDHGA5
2832533	PCDHGA6
2832533	PCDHGA7
2832533	PCDHGA8
2832533	PCDHGA9
2832533	PCDHGB1
2832533	PCDHGB2
2832533	PCDHGB3
2832533	PCDHGB4
2832533	PCDHGB5
2832533	PCDHGB6
2832533	PCDHGB7
2832533	PCDHGB8P
2832533	PCDHGC3
2832533	PCDHGC4
2832533	PCDHGC5
3342525	PCF11
2560195	PCGF1
2714230	PCGF3
3887165	PCIF1
3059258	PCLO
3059226	PCLO
3087813	PCM1
2930863	PCMT1
3134922	PCMTD1
3542689	PCNX
2461037	PCNXL2
2698996	PCOLCE2
3921599	PCP4
2868044	PCSK1
4007550	PCSK1N
3877892	PCSK2
3642200	PCSK6
3976124	PCTK1
3466826	PCTK2
2712754	PCYT1A
3263944	PDCD4
2616317	PDCD6IP
3358361	PDDC1
2984275	PDE10A
2589017	PDE11A
2590582	PDE1A
3416651	PDE1B
3044597	PDE1C
3381150	PDE2A
2340529	PDE4B
2858134	PDE4D
2431886	PDE4DIP
2783596	PDE5A
2714132	PDE6B
3138464	PDE7A
3606034	PDE8A
TCID	GENE_na29
2816681	PDE8B
3922793	PDE9A
2791197	PDGFC
3389077	PDGFD
2727226	PDGFRA
3087703	PDGFRL
3970833	PDHA1
3369249	PDHX
2639225	PDIA5
2469529	PDIA6
2540317	PDIA6
3062082	PDK4
3301218	PDLIM1
2796951	PDLIM3
2828441	PDLIM4
2736322	PDLIM5
2766588	PDS5A
3484768	PDS5B
3650300	PDXDC2
3923257	PDXK
4011637	PDZD11
2411173	PDZK1IP1
3411810	PDZRN4
3127610	PEBP4
3766796	PECAM1
2939593	PECI
2598496	PECR
3013255	PEG10
3872053	PEG3
2556302	PELI1
3537164	PELI2
2809128	PELO
2605749	PER2
2976360	PERP
3081191	PEX1
3753690	PEX12
2440117	PEX19
2465551	PEX5
2707045	PEX5L
3709540	PFAS
2440612	PFDN2
4009811	PFKFB1
2377094	PFKFB2
3233605	PFKFB3
3413344	PFKM
2700585	PFN2
3332780	PGA3
3332780	PGA4
3332780	PGA5
2593407	PGAP1
2390518	PGBD2
3288707	PGBD3
3108226	PGCP
2871685	PGGT1B
2339786	PGM1
3381925	PGM2L1
3173508	PGM5
3173508	PGM5P1
3173508	PGM5P2
TCID	GENE_na29
3824963	PGPEP1
3988740	PGRMC1
2928930	PHACTR2
3891530	PHACTR3
3403981	PHC1
3403981	PHC1B
2405469	PHC2
2704894	PHC3
3971451	PHEX
2903673	PHF1
2986084	PHF10
3489481	PHF11
3751184	PHF12
2990043	PHF14
2829337	PHF15
3975893	PHF16
2743315	PHF17
3116535	PHF20L1
2911944	PHF3
3991650	PHF6
4009506	PHF8
4012299	PHKA1
3659156	PHKB
3359461	PHLDA2
2635906	PHLDB2
3791482	PHLPP
2428699	PHTF1
3278198	PHYH
3190796	PHYHD1
3127385	PHYHIP
3247977	PHYHIPL
3290649	PHYHIPL
2905296	PI16
3259978	PI4K2A
2721777	PI4K2B
3953724	PI4KA
3953724	PI4KAP1
3953724	PI4KAP2
3599280	PIAS1
2432647	PIAS3
3493448	PIBF1
3945180	PICK1
2602653	PID1
3428268	PIGA
4000512	PIGA
2551690	PIGF
2714025	PIGG
3569339	PIGH
2418929	PIGK
3811086	PIGN
3391214	PIH1D2
3364759	PIK3C2A
2697564	PIK3CB
2395890	PIK3CD
3018309	PIK3CG
2813060	PIK3R1
2410470	PIK3R3
2525272	PIKFYVE
3015442	PILRB

Fig. 12 Continued

TCID	GENE_na29
3621728	PIN4
2400212	PINK1
3281068	PIP4K2A
3418303	PIP4K2C
3300869	PIP5K1A
3300869	PIPSL
3957938	PISD
3732230	PITPNC1
3437801	PIWIL1
3049700	PKD1L1
3697933	PKD1L3
2735221	PKD2
3111561	PKHD1L1
3104260	PKIA
3631964	PKM2
3822723	PKN1
3922975	PKNOX1
3354535	PKNOX2
3450234	PKP2
2511820	PKP4
3681488	PLA2G10
3376529	PLA2G16
3960388	PLA2G6
2955827	PLA2G7
2583374	PLA2R1
3136178	PLAG1
2977621	PLAGL1
3133233	PLAT
3252036	PLAU
3294959	PLAU
2475042	PLB1
3875642	PLCB1
3334372	PLCB3
3875908	PLCB4
3759587	PLCD3
2527895	PLCD4
3258477	PLCE1
2316605	PLCH2
2521574	PLCL1
2612813	PLCL2
3833443	PLD3
3157901	PLEC1
2486811	PLEK
3569257	PLEK2
3268274	PLEKHA1
2517737	PLEKHA3
3867458	PLEKHA4
3407096	PLEKHA5
3451988	PLEKHA8
2995189	PLEKHA8
3451988	PLEKHA9
2995189	PLEKHA9
2505957	PLEKHB2
3107828	PLEKHF2
3832992	PLEKHG2
2798475	PLEKHG4B
3569339	PLEKHH1
2479433	PLEKHH2
3759849	PLEKHM1
TCID	GENE_na29
3759849	PLEKHM1P
3598165	PLEKHO2
2492659	PLG
2492659	PLGLA
2492659	PLGLB1
2492659	PLGLB2
3638546	PLIN
2527196	PLK1
2858023	PLK2
2923270	PLN
2699564	PLOD2
3977067	PLP2
2645906	PLS1
3987996	PLS3
2699623	PLSCR4
3907524	PLTP
3854417	PLVAP
3755510	PLXDC1
3073267	PLXNA4
2673181	PLXNB1
3426502	PLXNC1
3911217	PMEPA1
3698081	PMFBP1
3962000	PMM1
3647504	PMM2
2988459	PMS2L3
3009580	PMS2L3
2934191	PNLDC1
3571542	PNMA1
3128731	PNMA2
3316287	PNPLA2
3067592	PNPLA8
4042198	PNRC2
2336650	PODN
3073013	PODXL
2365496	POGK
3972093	POLA1
3750625	POLDIP2
3563395	POLE2
2489606	POLE4
3607537	POLG
3766861	POLG2
3788833	POLI
2757621	POLN
2562605	POLR1A
2728448	POLR2B
2655773	POLR2H
3065154	POLR2J
3065154	POLR2J2
3065154	POLR2J3
3009838	POLR2J4
3065154	POLR2J4
3109191	POLR2K
3296512	POLR3A
3430129	POLR3B
2432647	POLR3C
2432571	POLR3GL
2988459	POM121
3057755	POM121

Fig. 12 Continued

TCID	GENE_na29
2779638	PPP3CA
3294499	PPP3CB
3798291	PPP4R1
3836760	PPP5C
3471300	PPTC7
3814701	PQLC1
4007765	PRAF2
2321058	PRAMEF1
2321058	PRAMEF13
2321058	PRAMEF14
2321058	PRAMEF2
3444578	PRB1
3444578	PRB3
3444578	PRB4
3639031	PRC1
3384270	PRCP
2919669	PRDM1
3933243	PRDM15
2316953	PRDM16
2321238	PRDM2
3470037	PRDM4
3851776	PRDX2
3483885	PRDX2
2367743	PRDX6
2474265	PREB
4053085	PRELP
2550959	PREPL
3102096	PREX2
3451375	PRICKLE1
4007865	PRICKLE3
3238962	PRINS
2337716	PRKAA2
3434142	PRKAB1
2433209	PRKAB2
3852529	PRKACA
2344393	PRKACB
3453556	PRKAG1
3732885	PRKAR1A
3018375	PRKAR2B
3653123	PRKCB
2624291	PRKCD
2480168	PRKCE
3538893	PRKCH
2651916	PRKCI
3275922	PRKCQ
4054481	PRKCZ
3559192	PRKD1
2548500	PRKD3
3134034	PRKDC
3246888	PRKG1
2775214	PRKG2
3009580	PRKRIP1
3997946	PRKX
3997946	PRKY
3838809	PRMT1
3924783	PRMT2
3323443	PRMT3
2349848	PRMT6
2459173	PRO2012
TCID	GENE_na29
3883207	PROCR
2682271	PROK2
2486851	PROKR1
2761842	PROM1
2685304	PROS1
2379665	PROX1
2358171	PRPF3
2336383	PRPF38A
3534201	PRPF39
3221543	PRPF4
3414104	PRPF40B
3454006	PRPF40B
2892738	PRPF4B
3413852	PRPH
3986514	PRPS1
3986514	PRPS1L1
3771513	PRPSAP1
3838624	PRR12
3656318	PRR14
2994981	PRR15
2984543	PRR18
3948259	PRR5
3973692	PRRG1
2662473	PRRT3
2366798	PRRX1
3028766	PRSS1
2899808	PRSS16
3028766	PRSS2
3343452	PRSS23
3028766	PRSS3
3281703	PRTFDC1
2358623	PRUNE
3210616	PRUNE2
3210497	PRUNE2
4036437	PRY
4036437	PRY2
3293762	PSAP
3175971	PSAT1
3126368	PSD3
3126191	PSD3
2501238	PSD4
3543481	PSEN1
3863929	PSG1
3863761	PSG1
3863929	PSG11
3863929	PSG2
3863761	PSG2
3863761	PSG3
3863929	PSG4
3863761	PSG4
3863929	PSG5
3863761	PSG5
3863929	PSG6
3863761	PSG6
3863929	PSG7
3863761	PSG7
3863929	PSG8
3863761	PSG8
3863929	PSG9
TCID	GENE_na29
3863761	PSG9
3881282	PSIMCT-1
3960478	psiTPTE22
3935902	psiTPTE22
3363979	PSMA1
3912861	PSMA7
2358906	PSMB4
2903285	PSMB9
3372209	PSMC3
3721886	PSMC3IP
3833291	PSMC4
3730941	PSMC5
3315549	PSMD13
2512701	PSMD14
2655650	PSMD2
3300869	PSMD4
3832383	PSMD8
3722152	PSME3
3873389	PSMF1
3035408	PSMG3
2348060	PTBP2
3014742	PTCD1
3215851	PTCH1
3727787	PTEN
3727787	PTENP1
3236786	PTER
3535752	PTGDR
2353717	PTGFRN
3220673	PTGR1
2448382	PTGS2
3156307	PTK2
2907671	PTK7
3405207	PTMA
3402736	PTMS
3074857	PTN
3839006	PTOV1
2911903	PTP4A1
2404819	PTP4A2
3118818	PTP4A3
3180142	PTPDC1
3279982	PTPLA
3164601	PTPLAD2
2692411	PTPLB
3329886	PTPMT1
3888721	PTPN1
3009959	PTPN12
2734629	PTPN13
2455418	PTPN14
2505529	PTPN18
2428796	PTPN22
2621333	PTPN23
3219885	PTPN3
3874023	PTPRA
2373842	PTPRC
3270270	PTPRE
2333318	PTPRF
2626802	PTPRG
3329983	PTPRJ
2973376	PTPRK

Fig. 12 Continued

TCID	GENE_na29
3777470	PTPRM
2600089	PTPRN
2327817	PTPRU
3757917	PTRF
2838201	PTTG1
2838201	PTTG2
2404254	PUM1
2542816	PUM2
2831567	PURA
2315739	PUSL1
3015442	PVRIG
3394488	PVRL1
2688499	PVRL2
2635641	PVRL3
3134922	PXDNL
2626167	PXK
3880767	PYGB
3564210	PYGL
3625440	PYGO1
2362351	PYHIN1
3302740	PYROXD2
3443464	PZP
3443348	PZP
2762334	QDPR
2477438	QPCT
3325768	QSER1
2369950	QSOX1
3638204	QTRT1
3820727	QTRT1
3458451	R3HDM2
3598482	RAB11A
3131741	RAB11FIP1
3642875	RAB11FIP3
3223872	RAB14
3568616	RAB15
3890870	RAB22A
2958670	RAB23
2361257	RAB25
3625271	RAB27A
3788976	RAB27B
3384321	RAB30
2929699	RAB32
3751002	RAB34
2810805	RAB3C
2456849	RAB3GAP2
3421706	RAB3IP
2905196	RAB44
2641263	RAB7A
3707642	RABEP1
3686750	RABEP2
3188299	RABGAP1
2367963	RABGAP1L
2342624	RABGGTB
2989050	RAC1
3451960	RACGAP1
3454223	RACGAP1
3451960	RACGAP1P
3454223	RACGAP1P
2852989	RAD1

TCID	GENE_na29
2662020	RAD18
3183757	RAD23B
2828564	RAD50
3590086	RAD51
3728776	RAD51C
3439836	RAD52
3036476	RADIL
2930957	RAET1G
2979187	RAET1G
2930957	RAET1L
2979187	RAET1L
2663244	RAF1
3327143	RAG1
2360700	RAG1AP1
3369931	RAG2
3580234	RAGE
3712675	RAI1
4001223	RAI2
2503200	RALB
3228523	RALGDS
3105271	RALYL
3438027	RAN
2499158	RANBP2
3847538	RANBP3
2853426	RANBP3L
2942504	RANBP9
2351872	RAP1A
2663244	RAP1A
2400655	rap1GAP
2736853	RAP1GDS1
3497659	RAP2A
4022032	RAP2C
2749699	RAPGEF2
3452690	RAPGEF3
2515783	RAPGEF4
3040967	RAPGEF5
2595560	RAPH1
3720921	RARA
2614369	RARB
3456081	RARG
3333899	RARRES3
2839671	RARS
2819044	RASA1
2645579	RASA2
3065154	RASA4
3065154	RASA4B
2369110	RASAL2
3747792	RASD1
3944210	RASD2
2775259	RASGEF1B
2817941	RASGRF2
3618736	RASGRP1
2476671	RASGRP3
3867346	RASIP1
3482845	RASL11A
3286776	RASSF4
3315952	RASSF7
3464405	RASSF9
3489020	RB1

TCID	GENE_na29
2988459	RBAK
3653317	RBBP6
4000944	RBBP7
3781429	RBBP8
3144760	RBM12B
3336402	RBM14
2675628	RBM15B
3556888	RBM23
3519119	RBM26
3976519	RBM3
3033397	RBM33
2622469	RBM5
2622469	RBM6
3959203	RBM9
3417583	RBMS2
2615060	RBMS3
2697863	RBP1
2319550	RBP7
3092415	RBPMS
3224366	RC3H2
3930235	RCAN1
2325479	RCAN3
3513794	RCBTB1
3513549	RCBTB2
2398894	RCC2
3336486	RCE1
3325503	RCN1
3602723	RCN2
2378584	RCOR3
3103293	RDH10
3416921	RDH5
3753896	RDM1
3390542	RDX
3158767	RECQL4
3248999	REEP3
2871176	REEP5
2561201	REG1P
2484358	REL
2765865	RELL1
3065740	RELN
3970214	REPS2
2395245	RERE
3446297	RERGL
2728408	REST
3243846	RET
2566764	REV1
2969677	REV3L
2766359	RFC1
3433747	RFC5
2867788	RFESD
3210457	RFK
2593796	RFTN2
3625674	RFX7
3855596	RFXANK
2779335	RG9MTD2
2950590	RGL2
3639601	RGMA
2821761	RGMB
2815488	RGNEF

Fig. 12 Continued

TCID	GENE_na29
2499158	RGPD1
2499158	RGPD2
2499158	RGPD3
2499158	RGPD4
2499158	RGPD5
2499158	RGPD6
2372781	RGS1
3642707	RGS11
2372812	RGS13
2447148	RGS16
2372719	RGS18
3914307	RGS19
2372858	RGS2
2441386	RGS5
2463227	RGS7
2447192	RGS8
3731543	RGS9
2530330	RHBDD1
3009198	RHBDD2
3674848	RHBDF1
2674242	RHOA
2820925	RHOBTB3
2724671	RHOH
3717539	RHOT1
2384401	RHOU
4019700	RHOXF1
3978295	RIBC1
3430331	RIC8B
3397877	RICS
2510485	RIF1
4052378	RILPL1
3476212	RILPL1
3476212	RILPL2
3110395	RIMS2
3378191	RIN1
3781654	RIOK3
3555675	RNASE1
3527597	RNASE4
3527662	RNASE6
2851965	RNASEN
2984884	RNASET2
3453348	RND1
2580802	RND3
3434413	RNF10
3130850	RNF122
2647458	RNF13
3059393	RNF13
2857204	RNF138
2857204	RNF138P1
2468376	RNF144A
2897172	RNF144B
2884216	RNF145
2924898	RNF146
3927949	RNF160
3133479	RNF170
2790324	RNF175
3221571	RNF183
3146565	RNF19A
3036985	RNF216
TCID	GENE_na29
2988459	RNF216L
3036985	RNF216L
2924253	RNF217
3895795	RNF24
3205162	RNF38
2715440	RNF4
2963929	RNGTT
3358049	RNH1
2672629	RNU13P3
3403077	RNU7
2683763	ROBO1
3354293	ROBO3
3220977	ROD1
2640263	ROPN1
2692640	ROPN1
2640263	ROPN1B
2692640	ROPN1B
2339872	ROR1
3627422	RORA
2435261	RORC
2721959	ROS1
2971378	ROS1
3098570	RP1
2960774	RP11-257K9.7
3674504	RP11-631M21.2
3514879	RP11-64P12.3
2432851	RP11-94I2.2
2357217	RP11-94I2.2
2321466	RP1-21O18.1
3975869	RP2
3896370	RP5-1022P6.2
3045047	RP9
3044938	RP9
3045047	RP9P
3044938	RP9P
2346738	RPAP2
2597273	RPE
2525852	RPE
4004878	RPGR
2374956	RPL10
2331178	RPL10
2904683	RPL10A
3674146	RPL13
2618640	RPL14
3294159	RPL17
3309215	RPL17
3867223	RPL18
3309936	RPL21
3309936	RPL21P119
3309936	RPL21P134
3309936	RPL21P14
3309936	RPL21P18
3309936	RPL21P19
3309936	RPL21P28
3309936	RPL21P29
3309936	RPL21P37
3309936	RPL21P39
2769512	RPL21P44
3309936	RPL21P45
TCID	GENE_na29
3309936	RPL21P46
3309936	RPL21P61
2898746	RPL21P68
3309936	RPL21P69
3309936	RPL21P7
3309936	RPL21P80
3309936	RPL21P93
3309936	RPL21P97
3309936	RPL21P98
3036476	RPL22
3729569	RPL23A
2390518	RPL23AP7
2390518	RPL23AP82
3174121	RPL24
3842141	RPL28
2604998	RPL3
2694706	RPL32
2694706	RPL32P3
2527196	RPL37A
3391149	RPL37AP8
2709606	RPL39L
2696309	RPL39P5
3962587	RPL5
2423175	RPL5
3797450	RPL6
3559570	RPL9
2319340	RPL9P11
3257559	RPP30
3804143	RPRD1A
2358221	RPRD2
3900470	RPS15A
2686717	RPS18
2686717	RPS18P12
2686717	RPS18P5
3834465	RPS19
3961253	RPS19BP1
2331679	RPS2
3383046	RPS20P27
3892660	RPS21
2850071	RPS26
2850071	RPS26P53
2850071	RPS26P8
2330723	RPS27
2531779	RPS28
2326561	RPS6KA1
2984655	RPS6KA2
4002173	RPS6KA3
3576284	RPS6KA5
4014029	RPS6KA6
3460467	RPSAP52
2662397	RPUSD3
2964231	RRAGD
3867965	RRAS
3363868	RRAS2
3899173	RRBP1
3318009	RRM1
2469252	RRM2
3681705	RRN3
3684782	RRN3

Fig. 12 Continued

TCID	GENE_na29
3684548	RRN3
3962469	RRP7A
3947434	RRP7A
3947460	RRP7A
3962469	RRP7B
3947434	RRP7B
3947460	RRP7B
2675763	RRP9
2428760	RSBN1
3010030	RSBN1L
3865696	RSHL1
3037100	RSPH10B
2989316	RSPH10B
3037100	RSPH10B2
2989316	RSPH10B2
2924851	RSPO3
3662612	RSPRY1
3279575	RSU1
2348854	RTCD1
2560076	RTKN
3567050	RTN1
2967650	RTN4IP1
2657025	RTP4
2730554	RUFY3
3930360	RUNX1
3144346	RUNX1T1
2908762	RUNX2
2401994	RUNX3
3928040	RWDD2B
2749484	RXFP1
3484393	RXFP2
2805939	RXFP3
2442008	RXRG
3832457	RYR1
2387126	RYR2
4045676	S100A1
2435383	S100A10
2435410	S100A11
2435981	S100A12
4045676	S100A13
4045665	S100A14
4045643	S100A16
2436088	S100A2
4045589	S100A5
4045577	S100A6
2435989	S100A8
2359664	S100A9
2329077	S100PBP
3365249	SAAL1
2620685	SACM1L
3505319	SACS
3691326	SALL1
3150289	SAMD12
3536434	SAMD4A
3252534	SAMD8
3061456	SAMD9L
3480657	SAP18
3337618	SAPS3
3293244	SAR1A
TCID	GENE_na29
2876257	SAR1B
2829416	SAR1B
3457201	SARNP
3470253	SART3
2930243	SASH1
3564027	SAV1
3362468	SBF2
3476130	SBNO1
3844978	SBNO2
2750594	SC4MOL
3352904	SC5DL
2817053	SCAMP1
3602004	SCAMP5
3605780	SCAND2
2621333	SCAP
2672712	SCAP
3633890	SCAPER
3091475	SCARA3
3129175	SCARA5
2774049	SCARB2
3402571	SCARNA10
3442054	SCARNA11
2715076	SCARNA22
3972093	SCARNA23
2437753	SCARNA4
2703217	SCARNA7
2389789	SCCPDH
3260586	SCD
3494629	SCEL
2769182	SCFD2
2601230	SCG2
3594003	SCG3
3587495	SCG5
2898934	SCGN
2649824	SCHIP1
2990404	SCIN
2785282	SCLT1
2585400	SCN2A
2584957	SCN3A
3393622	SCN4B
2585476	SCN7A
3414969	SCN8A
2585400	SCN9A
3441885	SCNN1A
3363868	SCP2
2336585	SCP2
2794006	SCRG1
3043895	SCRN1
3760957	SCRN2
3962839	SCUBE1
3362191	SCUBE2
2904485	SCUBE3
2780734	SCYE1
3428131	SCYL2
2594812	SCYL2
2773907	SDAD1
2542795	SDC1
3108146	SDC2
2404209	SDC3
TCID	GENE_na29
3907234	SDC4
3099750	SDCBP
3563459	SDCCAG1
3229943	SDCCAG3
3092415	SDHALP2
2987843	SDK1
2592532	SDPR
3837934	SEC1
3790479	SEC11C
3735752	SEC14L1
3942350	SEC14L2
2355615	SEC22B
3561952	SEC23A
3267455	SEC23IP
2829416	SEC24A
2783316	SEC24D
3235414	SEC61A2
2651782	SEC62
3779756	SEH1L
3574207	SEL1L
2764192	SEL1L3
2443476	SELE
2435005	SELENBP1
2473784	SELI
2676901	SELK
2676927	SELK
2443450	SELL
2443417	SELP
2647792	SELT
3058759	SEMA3C
3059667	SEMA3D
3059393	SEMA3E
2361342	SEMA4A
3607927	SEMA4B
2565592	SEMA4C
4051226	SEMA4D
2692199	SEMA5B
3592755	SEMA6D
2686646	SENP7
3278234	SEPHS1
2855285	SEPP1
2764054	SEPSECS
2569908	SEPT10
2732273	SEPT11
3764527	SEPT4
3735847	SEPT9
3837504	SEPWL
2814424	SERF1A
2814424	SERF1B
3365136	SERGEF
3947434	SERHL
3947460	SERHL
3947434	SERHL2
3947460	SERHL2
2328273	SERINC2
3906942	SERINC3
2864449	SERINC5
3577612	SERPINA1
3577577	SERPINA10

Fig. 12 Continued

TCID	GENE_na29
3577612	SERPINA2
3549757	SERPINA3
3549740	SERPINA3
3549740	SERPINA5
2938972	SERPINB1
3791996	SERPINB8
2601414	SERPINE2
3331355	SERPING1
2651165	SERPINI1
3387259	SESN3
2589929	SESTD1
2672532	SETD2
3579205	SETD3
3930781	SETD4
2609608	SETD5
3489418	SETDB2
3228007	SETX
3687277	SEZ6L2
3667281	SF3B3
3942998	SFI1
2676518	SFMBT1
2326774	SFN
3132782	SFRP1
3966597	SFRS17A
3771800	SFRS2
2403740	SFRS4
3747324	SFRS6
3886050	SFRS6
2548970	SFRS7
2366156	SFT2D2
2948683	SFTA2
2562435	SFTPB
2842101	SFXN1
3260985	SFXN3
2768981	SGCB
2837029	SGCD
3061805	SGCE
2648535	SGEF
2340695	SGIP1
2975014	SGK1
3096545	SGK196
3101802	SGK3
3101765	SGK3
3289235	SGMS1
2738664	SGMS2
3250863	GPL1
3568108	SGPP1
2529421	SGPP2
3431892	SH2B3
3088213	SH2D4A
3921442	SH3BGR
3226804	SH3GLB2
3717452	SH3GLP1
3304970	SH3PXD2A
2887164	SH3PXD2B
2793137	SH3RF1
2833924	SH3RF2
3380365	SHANK2
3158114	SHARPIN
TCID	GENE_na29
3205659	SHB
3708919	SHBG
2436985	SHC1
3213847	SHC3
2374926	SHISA4
2673270	SHISA5
3418007	SHMT2
2978050	SHPRH
2682568	SHQ1
3968303	SHROOM2
3354174	SIAE
2636483	SIDT1
3839619	SIGLEC12
3869097	SIGLEC6
3839619	SIGLEC7
3839619	SIGLEC9
3934111	SIK1
2877639	SIL1
3457336	SILV
3542847	SIPA1L1
2460817	SIPA1L2
3894727	SIRPA
3894727	SIRPB1
3249587	SIRT1
3357785	SIRT3
3434308	SIRT4
2895650	SIRT5
3042610	SKAP2
2651989	SKIL
2902884	SKIV2L
3154263	SLA
2440327	SLAMF1
3570266	SLC10A1
3142485	SLC10A5
2788511	SLC10A7
3454576	SLC11A2
2827525	SLC12A2
3696057	SLC12A4
3696035	SLC12A4
3617312	SLC12A6
2638728	SLC15A2
3477917	SLC15A4
2921402	SLC16A10
3981959	SLC16A2
2427469	SLC16A4
3734648	SLC16A5
3768412	SLC16A6
3290746	SLC16A9
2960955	SLC17A5
3867842	SLC17A7
3126694	SLC18A1
3935016	SLC19A1
3160658	SLC1A1
2485636	SLC1A4
2500919	SLC20A1
2828479	SLC22A4
3333831	SLC22A9
2877861	SLC23A1
3622934	SLC24A5
TCID	GENE_na29
3593014	SLC24A5
3952543	SLC25A1
2586845	SLC25A12
3062193	SLC25A13
3486728	SLC25A15
2909167	SLC25A27
2319340	SLC25A33
2645275	SLC25A36
3090006	SLC25A37
3377569	SLC25A45
3997360	SLC25A6
2756831	SLC26A1
3418394	SLC26A10
2835300	SLC26A2
3018605	SLC26A4
3106559	SLC26A7
2452754	SLC26A9
3593575	SLC27A2
2359885	SLC27A3
3872945	SLC27A5
2827645	SLC27A6
3212420	SLC28A3
2988594	SLC29A4
2409104	SLC2A1
2974935	SLC2A12
3450899	SLC2A13
4052021	SLC2A6
3622436	SLC30A4
2476116	SLC30A6
2349043	SLC30A7
2725381	SLC30A9
2701927	SLC33A1
2721959	SLC34A2
4007588	SLC35A2
2955061	SLC35B2
2940987	SLC35B3
3074039	SLC35B4
3329018	SLC35C1
3216276	SLC35D2
3389976	SLC35F2
2385967	SLC35F3
3076076	SLC37A3
3452417	SLC38A4
3769779	SLC39A11
3158581	SLC39A4
3804195	SLC39A6
3542063	SLC39A9
3333711	SLC3A2
2591837	SLC40A1
2452691	SLC41A1
3469180	SLC41A2
3373845	SLC43A3
3183111	SLC44A1
2418570	SLC44A5
3713951	SLC47A1
2512790	SLC4A10
3895330	SLC4A11
2730746	SLC4A4
2559849	SLC4A5

Fig. 12 Continued

TCID	GENE_na29
2831664	SLC4A9
3919033	SLC5A3
3467949	SLC5A8
3688878	SLC6A10P
3439510	SLC6A12
3439549	SLC6A13
3988165	SLC6A14
3464276	SLC6A15
3867734	SLC6A16
2799030	SLC6A19
2611848	SLC6A6
3688878	SLC6A8
2786322	SLC7A11
3087659	SLC7A2
3666146	SLC7A6
3666146	SLC7A6OS
3557209	SLC7A8
2497252	SLC9A2
3695541	SLC9A5
4006841	SLC9A7
2699145	SLC9A9
2696113	SLCO2A1
2869096	SLCO4C1
3139580	SLCO5A1
2869124	SLCO6A1
3753568	SLFN11
3753500	SLFN11
3753538	SLFN12
3753538	SLFN12L
3753568	SLFN13
3753500	SLFN13
3718555	SLFN5
3302187	SLIT1
3302056	SLIT1
2720584	SLIT2
4024685	SLITRK4
3495968	SLITRK5
3519840	SLITRK6
2625793	SLMAP
3779612	SLMO1
3911814	SLMO2
3907190	SLPI
3788302	SMAD4
3788270	SMAD4
3509842	SMAD9
3454821	SMAGP
3159946	SMARCA2
3820921	SMARCA4
2745646	SMARCA5
3457455	SMARCC2
3963754	SMC1B
3182781	SMC2
3776193	SMCHD1
2553911	SMEK2
3683050	SMG1
2438042	SMG5
2371255	SMG7
2860898	SMN1
2814424	SMN1

TCID	GENE_na29
2860898	SMN2
2814424	SMN2
2937144	SMOC2
3696317	SMPD3
2575897	SMPD4
2730404	SMR3A
2730396	SMR3A
2730404	SMR3B
2730396	SMR3B
3971387	SMS
3203382	SMU1
3014411	SMURF1
3766960	SMURF2
2464909	SMYD3
3740704	SMYD4
3876245	SNAP25
2962876	SNAP91
3163136	SNAPC3
2777714	SNCA
3022465	SND1
2607020	SNED1
3230332	SNHG7
3755198	SNIP
2407163	SNIP1
3386814	SNORA1
3230332	SNORA17
3386814	SNORA18
3309215	SNORA19
3319840	SNORA23
3386814	SNORA25
3553607	SNORA28
3453177	SNORA2A
3386814	SNORA32
3453177	SNORA34
2456849	SNORA36B
2558150	SNORA36C
3808600	SNORA37
3386814	SNORA40
2437753	SNORA42
3230332	SNORA43
3693673	SNORA46
3693673	SNORA50
3359469	SNORA54
3713951	SNORA59A
3713951	SNORA59B
2841284	SNORA74B
2375706	SNORA77
3386814	SNORA8
2404254	SNORD103A
2949038	SNORD117
3851651	SNORD41
2342624	SNORD45A
2342624	SNORD45C
2444451	SNORD47
3712098	SNORD49A
3386814	SNORD5
3542847	SNORD56B
3386814	SNORD6
3371719	SNORD67

TCID	GENE_na29
2444451	SNORD74
2444451	SNORD76
2444451	SNORD78
2444451	SNORD80
2949038	SNORD84
4017961	SNORD96B
2619666	SNRK
2565262	SNRNP200
4052378	SNRNP35
3642162	SNRPA1
3903052	SNTA1
3666601	SNTB2
3597857	SNX1
3725083	SNX11
3040073	SNX13
2963313	SNX14
3334783	SNX15
3142554	SNX16
3398482	SNX19
3597914	SNX22
3628994	SNX22
2826343	SNX24
2754582	SNX25
3830864	SNX26
2359036	SNX27
2348437	SNX7
2369557	SOAT1
3772279	SOCS3
2721633	SOD3
3301263	SORBS1
2796995	SORBS2
2797202	SORBS2
3592109	SORD
3352948	SORL1
2549092	SOS1
3960302	SOX10
2897899	SOX4
3447348	SOX5
3364306	SOX6
2603051	SP110
2531310	SP140
2531310	SP140L
2587520	SP3
2587520	SP3P
2430422	SPAG17
3742415	SPAG7
2882098	SPARC
2777113	SPARCL1
2420467	SPATA1
3481543	SPATA13
2380440	SPATA17
2726910	SPATA18
3909035	SPATA2
2742134	SPATA5
2867788	SPATA9
3413950	SPATS2
2522094	SPATS2L
2585933	SPC25
3591281	SPCS2

Fig. 12 Continued

TCID	GENE_na29
2752560	SPCS3
2322103	SPEN
3621948	SPG11
2834503	SPINK5
3907320	SPINLW1
3907335	SPINLW1
3590164	SPINT1
3799461	SPIRE1
2404766	SPOCD1
2876897	SPOCK1
2735027	SPP1
3623865	SPPL2A
3623472	SPPL2A
3589141	SPRED1
2556752	SPRED2
2742224	SPRY1
3519309	SPRY2
2319252	SPSB1
3190558	SPTAN1
2482505	SPTBN1
3573152	SPTLC2
3876990	SPTLC3
3114832	SQLE
3592511	SQRDL
2890239	SQSTM1
2844479	SQSTM1
2551327	SRBD1
3884191	SRC
3057668	SRCRB4D
2907730	SRF
3060300	SRI
3677969	SRL
3619165	SRP14
3771297	SRP68
2382781	SRP9
2951674	SRPK1
3066297	SRPK2
3740998	SRR
3712098	SRrp35
3802129	SS18
2619323	SS18L2
3027915	SSBP1
2864849	SSBP2
2413685	SSBP3
2518428	SSFA2
3408831	SSPN
2940551	SSR1
2709750	SST
2391647	SSU72
3976559	SSX1
3976559	SSX4
4007415	SSX4
3976559	SSX4B
4007415	SSX4B
3976559	SSX5
4007415	SSX5
3976559	SSX7
4007415	SSX7
3976559	SSX8

TCID	GENE_na29
2699145	ST13
3356175	ST14
3135046	ST18
3635198	ST20
2562529	ST3GAL5
2633256	ST3GAL6
3361971	ST5
2656837	ST6GAL1
3771712	ST6GALNAC1
3771675	ST6GALNAC2
2342904	ST6GALNAC5
3020496	ST7
3020496	ST7OT3
2868904	ST8SIA4
3429159	STAB2
2616804	STAC
2696802	STAG1
3015338	STAG3
3015338	STAG3L1
3056705	STAG3L1
3015338	STAG3L2
3056705	STAG3L2
3015338	STAG3L3
3056705	STAG3L3
3237088	STAM
3257031	STAMBPL1
3094447	STAR
3508898	STARD13
2870828	STARD4
2565143	STARD7
3457752	STAT2
2592356	STAT4
3721658	STAT5A
3721658	STAT5B
3458337	STAT6
2731986	STBD1
3011838	STEAP1
3011861	STEAP2
2502762	STEAP3
2528386	STK16
2593159	STK17B
2607262	STK25
3146103	STK3
2993029	STK31
2834282	STK32A
3361811	STK33
3409081	STK38L
2402459	STMN1
3104489	STMN2
3914050	STMN3
3223928	STOM
2481379	STON1
2481379	STON1-GTF2A1L
3574121	STON2
3632806	STRA6
3766284	STRADA
2522789	STRADB
3224591	STRBP

TCID	GENE_na29
3621351	STRC
2548274	STRN
3531163	STRN3
3852079	STX10
2327219	STX12
3891006	STX16
3819016	STXBP2
2350339	STXBP3
2929870	STXBP5
3558418	STXBP6
3513293	SUCLA2
2648098	SUCNR1
3490741	SUGT1
3486728	SUGT1L1
3102372	SULF1
3654669	SULT1A1
3656032	SULT1A1
3654669	SULT1A2
3656032	SULT1A2
3656032	SULT1A3
3656032	SULT1A4
2498911	SULT1C2
2498951	SULT1C4
3866785	SULT2A1
2955282	SUPT3H
3764384	SUPT4H1
3715703	SUPT6H
3228652	SURF1
3228621	SURF6
3220846	SUSD1
4045426	SUSD4
3236448	SUV39H2
3379390	SUV420H1
3716664	SUZ12P
2434139	SV2A
3608638	SV2B
3319997	SWAP70
2894790	SYCP2L
3468080	SYCP3
3178952	SYK
2963407	SYNCRIP
2979871	SYNE1
3539724	SYNE2
3610982	SYNM
3294668	SYNPO2L
4007899	SYP
3423622	SYT1
2361154	SYT11
3336652	SYT12
3371114	SYT13
2378256	SYT14
3287366	SYT15
3805614	SYT4
3973891	SYTL5
3377474	SYVN1
3014065	TAC1
3292946	TACR2
2414958	TACSTD2
3203199	TAF1

Fig. 12 Continued

TCID	GENE_na29
3361021	TAF10
3671607	TAF1C
3702382	TAF1C
3386814	TAF1D
3203199	TAF1L
2907018	TAF8
2635998	TAGLN3
2411198	TAL1
3183348	TAL2
3716048	TAOK1
2761941	TAPT1
3046520	TARP
3046708	TARP
3027961	TAS2R5
3898126	TASP1
3151943	TATDN1
3422326	TBC1D15
3217194	TBC1D2
2905432	TBC1D22B
3754041	TBC1D3
3754096	TBC1D3
4040117	TBC1D3
4040849	TBC1D3
3754041	TBC1D3B
3754096	TBC1D3B
4040117	TBC1D3B
4040849	TBC1D3B
3754041	TBC1D3C
3754096	TBC1D3C
4040117	TBC1D3C
4040849	TBC1D3C
3754041	TBC1D3D
3754096	TBC1D3D
4040117	TBC1D3D
4040849	TBC1D3D
3754041	TBC1D3E
3754096	TBC1D3E
4040117	TBC1D3E
4040849	TBC1D3E
3724698	TBC1D3F
3754041	TBC1D3F
3754096	TBC1D3F
4040117	TBC1D3F
4040849	TBC1D3F
3754041	TBC1D3G
3754096	TBC1D3G
4040117	TBC1D3G
4040849	TBC1D3G
3754041	TBC1D3H
3754096	TBC1D3H
4040117	TBC1D3H
4040849	TBC1D3H
3754041	TBC1D3P2
3754096	TBC1D3P2
4040117	TBC1D3P2
4040849	TBC1D3P2
3518086	TBC1D4
2942306	TBC1D7
3986291	TBC1D8B
TCID	GENE_na29
4002011	TBC1D8B
2709414	TBCCD1
2386418	TBCE
2738378	TBCKL
2780734	TBCKL
3419849	TBK1
3354174	TBRG1
3729834	TBX2
3982689	TBX22
3472755	TBX3
3846280	TBXA2R
3027204	TBXAS1
3576704	TC2N
2989537	tcag7.903
3985644	TCEAL3
3985644	TCEAL6
3985511	TCEAL7
3969396	TCEANC
3140833	TCEB1
3590498	TCEB1
3595096	TCF12
3962338	TCF20
2926447	TCF21
3674434	TCF25
3808854	TCF4
2491386	TCF7L1
3264621	TCF7L2
3913483	TCFL5
3578152	TCL1A
3550139	TCL1B
3550139	TCL6
3374890	TCN1
3929237	TCP10L
3368520	TCP11L1
3471327	TCTN1
3429365	TDG
2909167	TDRD6
2435218	TDRKH
3320944	TEAD1
3321055	TEAD1
3401259	TEAD4
2768396	TEC
2768396	TEC
3553141	TECPR2
3165825	TEK
3415668	TENC1
3555340	TEP1
3020192	TES
3473436	TESC
3249886	TET1
2738146	TET2
3217807	TEX10
3766716	TEX2
3247784	TFAM
3247784	TFAMP1
2329920	TFAP2E
2573570	TFCP2L1
4007734	TFE3
3069082	TFEC
TCID	GENE_na29
3933559	TFF1
3933536	TFF3
3955875	TFIP11
2591421	TFPI
3061621	TFPI2
3116614	TG
2558612	TGFA
2380590	TGFB2
3181728	TGFBR1
2615360	TGFBR2
2422722	TGFBR3
3776504	TGIF1
2562198	TGOLN2
3098935	TGS1
2550542	THADA
3067644	THAP5
3589458	THBS1
3625391	THEM4
2888103	THOC3
3956854	THOC5
3677356	THOC6
3756046	THRA
2666147	THRΒ
3341497	THRSP
3514879	THSD1
3514879	THSD1P
3600283	THSD4
2558511	TIA1
3309629	TIAL1
3928668	TIAM1
2932508	TIAM2
2871821	TICAM2
2333051	TIE1
2782230	TIFA
2735598	TIGD2
2883283	TIMD4
2374956	TIMM17A
3289031	TIMM23
3833093	TIMM50
3976341	TIMP1
3772661	TIMP2
3943504	TIMP3
2663130	TIMP4
3558012	TINF2
2649113	TIPARP
3693511	TIPIN
2366132	TIPRL
3355091	TIRAP
2908008	TJAP1
3615579	TJP1
3173880	TJP2
3817116	TJP3
3695107	TK2
3751042	TLCD1
3211579	TLE1
3176209	TLE4
3211579	TLE4
2750753	TLL1
3597125	TLN2

Fig. 12 Continued

TCID	GENE_na29
2766192	TLR10
2748346	TLR2
2754937	TLR3
2766262	TLR6
3969115	TLR8
2676009	TLR9
2700365	TM4SF1
2647315	TM4SF4
3605268	TM6SF1
3448481	TM7SF3
3110608	TM7SF4
3881686	TM9SF4
3414186	TMBIM6
3635578	TMC3
3650762	TMC7
2694931	TMCC1
2376376	TMCC2
3466206	TMCC3
2442134	TMCO1
3048778	TMED4
2423264	TMED5
2871821	TMED7
3182229	TMEFF1
2592598	TMEFF2
3763390	TMEM100
3722479	TMEM106A
2990342	TMEM106B
3413278	TMEM106C
2642995	TMEM108
3471769	TMEM116
3412345	TMEM117
3388631	TMEM123
2758658	TMEM128
3478068	TMEM132D
3346147	TMEM133
3397461	TMEM135
3352485	TMEM136
3029016	TMEM139
2749380	TMEM144
3830530	TMEM147
2910364	TMEM14A
2701294	TMEM14E
2562387	TMEM150
2766289	TMEM156
2577482	TMEM163
3987029	TMEM164
3068476	TMEM168
2815220	TMEM171
2877990	TMEM173
3031624	TMEM176A
2478269	TMEM178
2537290	TMEM18
2746645	TMEM184C
3659858	TMEM188
3422231	TMEM19
3039399	TMEM195
3040465	TMEM196
2925590	TMEM200A
3072435	TMEM209

TCID	GENE_na29
3166644	TMEM215
3745525	TMEM220
2961317	TMEM30A
2725332	TMEM33
2690850	TMEM39A
2708817	TMEM41A
2611779	TMEM43
2633691	TMEM45A
4004575	TMEM47
2413423	TMEM48
3929664	TMEM50B
2409770	TMEM53
3144235	TMEM55A
2347732	TMEM56
3825244	TMEM59L
3058156	TMEM60
3591281	TMEM62
3129948	TMEM66
3136015	TMEM68
2351632	TMEM77
2427720	TMEM77
3358262	TMEM80
3316126	TMEM80
3599669	TMEM84
3620515	TMEM87A
2500615	TMEM87B
3709153	TMEM88
2450668	TMEM9
3571904	TMEM90A
3717870	TMEM98
3181240	TMOD1
3594031	TMOD2
2668205	TPPPE
2729821	TMPRSS11E
2729821	TMPRSS11E2
3393536	TMPRSS13
3351200	TMPRSS4
2491271	TMSB10
4016193	TMSB15A
4016193	TMSB15B
3449068	TMTC1
3425134	TMTC3
3812206	TMX3
3896976	TMX4
3222170	TNC
3750625	TNFAIP1
2825629	TNFAIP8
3127703	TNFRSF10B
3089816	TNFRSF10C
3150455	TNFRSF11B
3645555	TNFRSF12A
3648391	TNFRSF17
3441849	TNFRSF1A
2956052	TNFRSF21
2705706	TNFSF10
3487299	TNFSF11
3500787	TNFSF13B
3222128	TNFSF15
2444239	TNFSF18

TCID	GENE_na29
2705266	TNIK
3085270	TNKS
3373675	TNKS1BP1
3257850	TNKS2
2342176	TNNI3K
2342220	TNNI3K
2450762	TNNT2
3317117	TNNT3
3255402	TNPO1
3851651	TNPO2
3653398	TNRC6A
3946192	TNRC6B
2599153	TNS1
3049522	TNS3
2949622	TNXA
2949622	TNXB
3762473	TOB1
3907111	TOMM34
3157385	TOP1MT
3756193	TOP2A
3748262	TOP3A
2695941	TOPBP1
2446198	TOR1AIP1
2446198	TOR1AIP2
3136888	TOX
2457842	TP53BP2
2544201	TP53I3
3145149	TP53INP1
3883013	TP53INP2
3060095	TP53TG1
2317317	TP73
2915133	TPBG
3141857	TPD52
2924330	TPD52L1
3204721	TPM2
2436526	TPM3
3823511	TPM4
2944025	TPMT
2466554	TPO
3361041	TPP1
3499453	TPP2
3643679	TPSD1
3005444	TPST1
3955915	TPST2
3881443	TPX2
3528172	TRA@
3041519	TRA2A
2709062	TRA2B
3528172	TRAC
3695433	TRADD
3553337	TRAF3
2534810	TRAF3IP1
2969810	TRAF3IP2
3715839	TRAF4
2378662	TRAF5
3369890	TRAF6
3644541	TRAF7
3432267	TRAFD1
3528172	TRAJ17

Fig. 12 Continued

TCID	GENE_na29
2619120	TRAK1
2594812	TRAK2
3139882	TRAM1
2957227	TRAM2
3843188	TRAPPC2
3843188	TRAPPC2P1
3865223	TRAPPC6A
3155937	TRAPPC9
3528172	TRAV20
3528172	TRAV8-3
3528172	TRD@
2953570	TREM1
2953501	TREM2
2906720	TREML2
2906720	TREML2P
2906720	TREML3
2906720	TREML4
2954022	TRERF1
2954025	TRERF1
3046520	TRGC2
3046708	TRGV3
3046520	TRGV9
2641901	TRH
3115008	TRIB1
3489644	TRIM13
3217123	TRIM14
3746845	TRIM16
3746845	TRIM16L
3318443	TRIM22
3360622	TRIM22
2859734	TRIM23
2948259	TRIM26
3128954	TRIM35
2901503	TRIM39
2901503	TRIM39R
2430126	TRIM45
3360622	TRIM5
2845078	TRIM52
2390180	TRIM58
3372896	TRIM77
3818515	TRIP10
2602901	TRIP12
2798915	TRIP13
3597977	TRIP4
2408111	TRIT1
3896524	TRMT6
2326846	TRNP1
2660648	TRNT1
2645951	TRPC1
3903708	TRPC4AP
4018327	TRPC5
3388438	TRPC6
3359267	TRPM5
3623771	TRPM7
3149528	TRPS1
3014411	TRRAP
3028766	TRY6
3228373	TSC1
3512294	TSC22D1
TCID	GENE_na29
2647647	TSC22D2
3064039	TSC22D4
3365437	TSG101
2566586	TSGA10
3546213	TSHR
3794056	TSHZ1
3889419	TSHZ2
3889624	TSHZ2
3858285	TSHZ3
2503618	TSN
2334602	TSPAN1
3069955	TSPAN12
2991150	TSPAN13
2842707	TSPAN17
3458783	TSPAN31
3316375	TSPAN4
2778856	TSPAN5
4015397	TSPAN6
3974019	TSPAN7
3461981	TSPAN8
3978169	TSPYL2
3740998	TSR1
2538480	TSSC1
3620799	TTBK2
2654306	TTC14
3327948	TTC17
2585236	TTC21B
2806186	TTC23L
3721516	TTC25
3920385	TTC3
2589011	TTC30A
2588965	TTC30A
2588965	TTC30B
3111375	TTC35
2887693	TTC37
2412312	TTC39A
3199662	TTC39B
3781980	TTC39C
3333603	TTC9C
2500803	TTL
3761551	TTLL6
2420229	TTLL7
3906942	TTPAL
3783565	TTR
3734236	TTYH2
3453837	TUBA1A
3453732	TUBA1A
3453837	TUBA1B
3453732	TUBA1B
3413787	TUBA1B
3413787	TUBA1C
2575949	TUBA3C
2506185	TUBA3C
2575949	TUBA3D
2506185	TUBA3D
2575949	TUBA3E
2506185	TUBA3E
2528407	TUBA4A
2528407	TUBA4B
TCID	GENE_na29
3936515	TUBA8
2901913	TUBB
3891342	TUBB1
2939232	TUBB2A
2939232	TUBB2B
4050485	TUBB2C
4051521	TUBB2C
2901913	TUBB2C
3674504	TUBB3
3847959	TUBB4
3674504	TUBB4Q
3779579	TUBB6
2901913	TUBBP1
2901913	TUBBP2
3764933	TUBD1
3721926	TUBG1
3721926	TUBG2
3526151	TUBGCP3
2358993	TUFT1
3401217	TULP3
3087167	TUSC3
4027769	TWF1
3451708	TWF1
2676009	TWF2
3220156	TXN
2412668	TXNDC12
2829562	TXNDC15
3707990	TXNDC17
3778589	TXNDC2
2697372	TXNDC6
2644461	TXNDC6
3809324	TXNL1
3429460	TXNRD1
3775842	TYMS
3795850	TYMS
3795866	TYMS
3633794	TYRO3
3590498	TYRO3
3633794	TYRO3P
3590498	TYRO3P
3293215	TYSND1
3005995	TYW1
3055608	TYW1
3005995	TYW1B
3055608	TYW1B
3842345	U2AF2
3631397	UACA
2364189	UAP1
3976062	UBA1
3829768	UBA2
2345196	UBA2
2681195	UBA3
2695648	UBA5
2642911	UBA5
2771718	UBA6
2674762	UBA7
3498315	UBAC2
3203753	UBAP2
2360083	UBAP2L

Fig. 12 Continued

TCID	GENE_na29
3988874	UBE2A
2613880	UBE2E2
3072276	UBE2H
3643703	UBE2I
3771543	UBE2O
2436716	UBE2Q1
2451200	UBE2T
3097208	UBE2V2
3725481	UBE2Z
3614087	UBE3A
3431143	UBE3B
3033924	UBE3C
2319560	UBE4B
3653000	UBFD1
4027355	UBL4A
3820161	UBL5
2668351	UBP1
2437893	UBQLN4
3360553	UBQLNL
3620880	UBR1
2907190	UBR2
2399409	UBR4
3147321	UBR5
3376023	UBXN1
2402601	UBXN11
2472914	UBXN2A
2507495	UBXN4
2712858	UBXN7
2725013	UCHL1
2448971	UCHL5
3381817	UCP2
3365487	UEVLD
2754673	UFSP2
2504883	UGCGL1
3521484	UGCGL2
2766456	UGDH
2485257	UGP2
2853293	UGT3A1
2853325	UGT3A2
2740507	UGT8
2364155	UHMK1
2904270	UHRF1BP1
2930957	ULBP2
2979187	ULBP2
3749010	ULK2
2670481	ULK4
3683549	UMOD
2639874	UMPS
3167731	UNC13B
2986825	UNC84A
2758602	UNC93B3
3474697	UNQ1887
2732339	UNQ3028
3629103	UNQ353
2886535	UNQ9374
3277662	UPF2
3503224	UPF3A
3351688	UPK2
3065154	UPLP
TCID	GENE_na29
3000953	UPP1
2511712	UPP2
3652218	UQCRC2
2384705	URB2
2334279	UROD
3854311	USHBP1
3775686	USP14
3936550	USP18
3749498	USP22
3391816	USP28
3685051	USP31
3765167	USP32
3748400	USP32
2419113	USP33
2555277	USP34
2604138	USP40
3936550	USP41
3466499	USP44
2966298	USP45
3320604	USP47
2400718	USP48
3294576	USP54
3765167	USP6
3748400	USP6
3277468	USP6NL
3679564	USP7
3974708	USP9X
3974708	USP9Y
2930418	UST
3990566	UTP14A
3490504	UTP14A
3990566	UTP14C
3490504	UTP14C
3726992	UTP18
2929168	UTRN
3340697	UVRAG
3441941	VAMP1
3744217	VAMP2
2443989	VAMP4
4032755	VAMP7
2353237	VANGL1
3778601	VAPA
2902013	VARS2
3758291	VAT1
3669552	VAT1L
2426385	VAV3
3996815	VBP1
3252071	VCL
3998386	VCX
3998386	VCX2
3998386	VCX3A
3998386	VCX3B
3998386	VCY
2875954	VDAC1
3452818	VDR
2908179	VEGFA
2794792	VEGFC
2610336	VHL
2619265	VIPR1
TCID	GENE_na29
3082373	VIPR2
2477203	VIT
3005266	VKORC1L1
3160175	VLDLR
3872274	VN1R1
2974592	VNN1
2974635	VNN2
2974610	VNN3
2675628	VPRBP
3351806	VPS11
3108901	VPS13B
3627929	VPS13C
2320762	VPS13D
3874023	VPS16A
3471264	VPS29
3475545	VPS33A
3689922	VPS35
3475838	VPS37B
3620457	VPS39
7385683	VPS41
2357996	VPS45
2556215	VPS54
2434892	VPS72
3550485	VRK1
3353914	VWA5A
3441685	VWF
3240340	WAC
3298738	WAPAL
3579546	WARS
2969289	WASF1
3482572	WASF3
3445670	WBP11
3445670	WBP11P1
3771037	WBP2
3985523	WBP5
3056838	WBSCR16
3007024	WBSCR17
2601341	WDFY1
3490251	WDFY2
2776372	WDFY3
3565571	WDHD1
2760371	WDR1
2752478	WDR17
2724235	WDR19
3553017	WDR20
2725332	WDR21B
3569926	WDR22
2458082	WDR26
2354082	WDR3
2430422	WDR3
3231846	WDR37
3203855	WDR40A
2863535	WDR41
3988365	WDR44
2426840	WDR47
2704143	WDR49
3464912	WDR51B
2489228	WDR54
2673830	WDR6

Fig. 12 Continued

TCID	GENE_na29
3034449	WDR60
2344731	WDR63
2332999	WDR65
3730731	WDR68
3789442	WDR7
3625052	WDR72
2417016	WDR78
2676041	WDR82
3607766	WDR93
3593770	WDSOF1
3886938	WFDC2
3907320	WFDC6
3907348	WFDC8
2715076	WHSC1
3460198	WIF1
2587841	WIPF1
2995076	WIPF3
3768474	WIPI1
3400034	WNK1
4009604	WNK3
2528159	WNT10A
2677356	WNT5A
2830598	WNT8A
3715109	WSB1
3430620	WSCD2
2753732	WWC2
3968397	WWC3
3669650	WWOX
3105777	WWP1
3977862	XAGE1A
3977862	XAGE1B
3977862	XAGE1C
3977862	XAGE1D
3977862	XAGE1E
3956589	XBP1
3989678	XIAP
2513758	XIRP2
4015548	XKRX
3216969	XPA
2663810	XPC
3946510	XPNPEP3
3961699	XPNPEP3
2555490	XPO1
3686339	XPO6
3088983	XPO7
2370123	XPR1
2818454	XRCC4
2526980	XRCC5
3879467	XRN2
3451264	YAF2
2929870	YAP1
2405192	YARS
2329077	YARS
3639601	YBX2
2923060	YDD19
2655168	YEATS2
3832280	YIF1B
2413484	YIPF1
2879509	YIPF5

TCID	GENE_na29
3825838	YJEFN3
3240012	YME1L1
3938244	YPEL1
4024373	YTHDC2
2824483	YTHDC2
2327630	YTHDF2
3886639	YWHA8
3943207	YWHAH
2539869	YWHAQ
2437577	YY1AP1
2437645	YY1AP1
3971329	YY2
3813604	ZADH2
2704052	ZBBX
2688499	ZBED2
3349719	ZBTB16
3989089	ZBTB33
3458216	ZBTB39
3189580	ZBTB43
3398241	ZBTB44
2571217	ZC3H8
2412834	ZCCHC11
3988596	ZCCHC12
3703665	ZCCHC14
3987607	ZCCHC16
3791341	ZCCHC2
3296981	ZCCHC24
2721809	ZCCHC4
2818079	ZCCHC9
3063968	ZCWPW1
3322958	ZDHHC13
2933175	ZDHHC14
3504691	ZDHHC20
3199431	ZDHHC21
3331433	ZDHHC5
3307120	ZDHHC6
4021341	ZDHHC9
3241316	ZEB1
2579572	ZEB2
2528308	ZFAND2B
2905664	ZFAND3
3462094	ZFC3H1
3797015	ZFP161
3842794	ZFP28
3860954	ZFP30
3569754	ZFP36L1
3331730	ZFP91
3331730	ZFP91-CNTF
3110789	ZFPM2
3971923	ZFX
3971923	ZFY
2817731	ZFYVE16
3619650	ZFYVE19
3590129	ZFYVE19
3553947	ZFYVE21
3260018	ZFYVE27
2757796	ZFYVE28
3906062	ZHX3
3872053	ZIM2

TCID	GENE_na29
3015147	ZKSCAN1
2706791	ZMAT3
3132616	ZMAT4
3480129	ZMYM2
4011889	ZMYM3
2329752	ZMYM4
2408929	ZMYND12
3908149	ZMYND8
3053380	ZNF117
3849549	ZNF121
2465551	ZNF124
2808290	ZNF131
3872928	ZNF132
3843742	ZNF135
3855985	ZNF14
3319898	ZNF143
3835318	ZNF155
3976240	ZNF157
2900195	ZNF165
3819968	ZNF177
3829857	ZNF181
2900423	ZNF187
2900372	ZNF193
3359751	ZNF195
2620160	ZNF197
3717635	ZNF207
3856646	ZNF208
3856720	ZNF208
3910260	ZNF217
3244539	ZNF22
3835318	ZNF221
3835418	ZNF224
3835494	ZNF226
3835565	ZNF233
3835467	ZNF234
3285614	ZNF25
3826041	ZNF253
3872560	ZNF256
3826803	ZNF257
3392871	ZNF259
3657367	ZNF267
3004768	ZNF273
3019401	ZNF277
4021508	ZNF280C
3625823	ZNF280D
3933399	ZNF295
3063646	ZNF3
2881607	ZNF300
3829857	ZNF302
3869714	ZNF320
2946500	ZNF322A
2946500	ZNF322B
2947248	ZNF323
2745220	ZNF330
3159013	ZNF34
2842860	ZNF346
3870135	ZNF347
3818842	ZNF358
2862019	ZNF366

Fig. 12 Continued

TCID	GENE_na29
3831475	ZNF382
3129361	ZNF395
3803882	ZNF397
3803882	ZNF397OS
3793888	ZNF407
3872335	ZNF416
3843275	ZNF419
3826656	ZNF429
3856720	ZNF43
3826306	ZNF43
3677592	ZNF434
3283613	ZNF438
3851454	ZNF443
3991992	ZNF449
3078774	ZNF467
3842839	ZNF470
3244055	ZNF487
3826803	ZNF492
3216476	ZNF510
2474651	ZNF512
3840795	ZNF525
3843386	ZNF530
3866649	ZNF541
3843848	ZNF544
3843188	ZNF547
3843214	ZNF548
3872441	ZNF552
3819968	ZNF559
3849797	ZNF561
3849797	ZNF562
3849549	ZNF562
3860824	ZNF569
3114820	ZNF572
3864430	ZNF575
3842315	ZNF580
3842301	ZNF581
3871903	ZNF582
2390976	ZNF596
3213530	ZNF598
3869714	ZNF600
3872604	ZNF606
3869714	ZNF611
3747324	ZNF624
3826041	ZNF626
3826079	ZNF626
2331903	ZNF643
2422517	ZNF644
3761737	ZNF652
3014904	ZNF655
3206317	ZNF658
3206317	ZNF658B
3826306	ZNF66
2619521	ZNF662
3871935	ZNF667
2465493	ZNF670
3857171	ZNF675
3856720	ZNF676
2402691	ZNF683
3687910	ZNF689

TCID	GENE_na29
2465493	ZNF695
3849549	ZNF699
3657367	ZNF720
3657318	ZNF720
2756309	ZNF721
3856720	ZNF728
3078656	ZNF746
3843275	ZNF749
4023006	ZNF75D
3840795	ZNF761
3840795	ZNF765
3078656	ZNF767
3872274	ZNF772
3860999	ZNF781
3030585	ZNF783
3014808	ZNF789
3851454	ZNF799
3843906	ZNF8
2518889	ZNF804A
3011675	ZNF804B
3872441	ZNF814
3718382	ZNF830
3871903	ZNF835
3688381	ZNF843
3840795	ZNF845
3849549	ZNF846
3826306	ZNF85
3826079	ZNF90
3005069	ZNF92
3826079	ZNF93
3826803	ZNF98
3856720	ZNF99
3826803	ZNF99
2420958	ZNHIT6
3668834	ZNRF1
2995076	ZNRF2
3057755	ZP3
3057668	ZP3
2462693	ZP4
3050170	ZPBPP
2577700	ZRANB3
2900269	ZSCAN16
3391769	ZW10
3290210	ZWINT
2336539	ZYG11A
3741875	ZZEF1

Fig. 12 Continued

## METHODS FOR CLASSIFICATION OF TISSUE SAMPLES AS POSITIVE OR NEGATIVE FOR CANCER

### CROSS-REFERENCE

[0001] This application is a continuation-in-part of U.S. application Ser. No. 18/823,253, filed Sep. 3, 2024, which is a continuation of U.S. application Ser. No. 17/157,876, filed Jan. 25, 2021 (now U.S. Pat. No. 12,110,554, issued Oct. 8, 2024), which is a continuation of U.S. application Ser. No. 14/153,219, filed Jan. 13, 2014 (now U.S. Pat. No. 10,934,587, issued Mar. 2, 2021), which is a continuation of U.S. application Ser. No. 13/318,751, filed May 10, 2012 (now U.S. Pat. No. 8,669,057, issued Mar. 11, 2014), which is a national stage application of International Application No. PCT/US2010/034140, filed May 7, 2010, which claims the benefit of U.S. Provisional Application No. 61/176,471, filed May 7, 2009, each of which is incorporated herein by reference in its entirety.

### BACKGROUND

[0002] Thyroid cancer incidence has increased substantially in the United States in recent decades, with evidence to support both an increase in detection and a true increase in occurrence. Thyroid nodules are palpable in 5% of adults and are visualized with contemporary imaging in more than one-third of adults. Malignancy is present in only 5% to 15% of all thyroid nodules, and definitive diagnosis is achieved by surgical histopathology on resected tissue. Unfortunately, thyroid surgery is associated with discomfort, scarring, inconvenience, direct and indirect costs, potential lifelong medication, and occasional surgical complications. Efforts to exclude cancer with clinical assessment alone are admittedly imperfect, and laboratory testing of serum thyroid stimulating hormone levels and thyroid imaging with radionuclides or ultrasonography identify benignity with high confidence in only 4% to 26% of nodules. Forty years ago, the application of cytology to thyroid nodule specimens obtained by fine-needle aspiration (FNA) biopsy had a substantial effect on patient management by reducing surgery by one half and doubling the proportion of cancer among patients who underwent surgery. However, approximately one-third of thyroid nodule cytology findings today are cytologically indeterminate, with estimated risks of malignancy ranging from 5% to 30%. Consequently, approximately three quarters of patients with cytologically indeterminate thyroid nodules have been referred for surgery, even though 80% ultimately prove to have benign nodules.

### SUMMARY OF THE INVENTION

[0003] The present invention includes a method for diagnosing thyroid disease in a subject, the method comprising (a) providing a DNA sample from a subject; (b) detecting the presence of one or more polymorphisms selected from the group consisting of the polymorphisms listed in Tables 1, 3-6, 8 or lists 1-45 or their complement; and (c) determining whether said subject has or is likely to have a malignant or benign thyroid condition based on the results of step (b).

[0004] The present invention also includes a composition comprising one or more binding agents that specifically bind to the one or more polymorphisms selected from the group

consisting of the polymorphisms listed in Tables 1, 3-6, 8 or lists 1-45 or their complement.

[0005] In another embodiment, the present invention includes a kit for diagnosing thyroid disease in a subject, the kit comprising: (a) at least one binding agent that specifically binds to the one or more polymorphisms selected from the group consisting of the polymorphisms listed in Tables 1, 3-6, 8 or lists 1-45 or their complement; and (b) reagents for detecting binding of said at least one binding agent to a DNA sample from a subject.

[0006] In another embodiment, the present invention includes a business method for diagnosing thyroid disease in a subject, the business method comprising: (a) diagnosing thyroid disease from a subject using the method stated above; (b) providing the results of the diagnosis to the subject, a healthcare provider, or a third party; and (c) billing said subject, healthcare provider, or third party

[0007] The present disclosure describes enhanced technologies for characterizing genomic information, including improved methods for the measurement of RNA transcriptome expression and sequencing of nuclear and mitochondrial RNAs, measurement changes in genomic copy number, including loss of heterozygosity, and the development of enhanced bioinformatics and machine learning strategies, resulting in a more robust genomic test.

[0008] An aspect of the present disclosure provides a method for processing or analyzing a tissue sample of a subject, comprising: (a) subjecting a first portion of the tissue sample to cytological analysis that indicates that the first portion of the tissue sample is cytologically indeterminate; (b) upon identifying the first portion of the tissue sample as being cytologically indeterminate, assaying by sequencing, array hybridization, or nucleic acid amplification a plurality of gene expression products from a second portion of the tissue sample to yield a first data set; (c) in a programmed computer, using a trained algorithm that comprises one or more classifiers to process the first data set from (b) to generate a classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant, wherein the one or more classifiers comprises an ensemble classifier integrated with at least one index selected from the group consisting of: a follicular content index, a Hürthle cell index, and a Hürthle neoplasm index; and (d) outputting a report indicative of the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant.

[0009] In some embodiments, the plurality of gene expression products include two or more of sequences corresponding to mRNA transcripts, mitochondrial transcripts, and chromosomal loss of heterozygosity. In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a specificity of at least about 60%. In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a specificity of at least about 68%. In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a specificity of at least about 70%. In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a sensitivity of at least about 90%.

[0010] In some embodiments, the one or more classifiers comprises the ensemble classifier integrated with the folli-

cular content index, the Hürthle cell index, and the Hürthle neoplasm index. In some embodiments, the one or more classifiers further comprises one or more upstream classifiers, wherein the one or more upstream classifiers are selected from the group consisting of: a parathyroid classifier, a medullary thyroid cancer (MTC) classifier, a variant detection classifier, and a fusion transcript detection classifier. In some embodiments, the one or more classifiers comprises a parathyroid classifier that identifies a presence or an absence of a parathyroid tissue in the second portion of the tissue sample. In some embodiments, the upon identification of the absence of the parathyroid tissue in the second portion of the tissue sample by the parathyroid classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the one or more classifiers comprises a medullary thyroid cancer (MTC) classifier that identifies a presence or an absence of a medullary thyroid cancer (MTC) in the second portion of the tissue sample. In some embodiments, the upon identification of the absence of the MTC in the second portion of the tissue sample by the MTC classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the one or more classifiers comprises a variant detection classifier that identifies a presence or an absence of a BRAF mutation in the second portion of the tissue sample. In some embodiments, the BRAF mutation is a BRAF V600E mutation. In some embodiments, the upon identification of the absence of the BRAF mutation in the second portion of the tissue sample by the variant detection classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the one or more classifiers comprises a fusion transcript detection classifier that identifies a presence or an absence of a RET/PTC gene fusion in the second portion of the tissue sample. In some embodiments, the RET/PTC gene fusion is RET/PTC1 or RET/PTC3 gene fusion. In some embodiments, the upon identification of the absence of the RET/PTC gene fusion in the second portion of the tissue sample by the fusion transcript detection classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the follicular content index identifies follicular content in the second portion of the tissue sample.

[0011] In some embodiments, the ensemble classifier analyzes, in the first data set, sequence information corresponding to at least 500 genes of Table 3. In some embodiments, the ensemble classifier analyzes, in the first data set, sequence information corresponding to at least 1000 genes of Table 3. In some embodiments, the ensemble classifier analyzes, in the first data set, sequence information corresponding to 1115 genes of Table 3.

[0012] In some embodiments, the method further comprising (e) upon identifying the second portion of the tissue sample as being suspicious for malignancy, or malignant (i) processing the first data set to identify one or more genetic aberrations in one or more genes listed in FIG. 12; and (ii) outputting a second report indicative of a risk of malignancy, a histological subtype, and a prognosis associated with each

of one or more genetic aberration identified in the second portion of the tissue sample. In some embodiments, the one or more genetic aberrations is a DNA variant. In some embodiments, the one or more genetic aberrations is a RNA fusion. In some embodiments, the risk of malignancy characterizes the one or more genetic aberrations as (1) highly associated with malignant nodules, (2) associated with both benign and malignant nodules, or (3) has insufficient published evidence.

[0013] In some embodiments, the tissue sample is a thyroid tissue sample. In some embodiments, the tissue sample is a needle aspirate sample. In some embodiments, the needle aspirate sample is a fine needle aspirate sample. In some embodiments, the malignancy is thyroid cancer.

[0014] Another aspect of the present disclosure provides a method for processing or analyzing a tissue sample of a subject, comprising: (a) subjecting a first portion of the tissue sample to cytological analysis that indicates that the first portion of the tissue sample is cytologically indeterminate; (b) upon identifying the first portion of the tissue sample as being cytologically indeterminate, assaying by sequencing, array hybridization, or nucleic acid amplification a plurality of gene expression products from a second portion of the tissue sample to yield a first data set, wherein the plurality of gene expression products include two or more of sequences corresponding to mRNA transcripts, mitochondrial transcripts, and chromosomal loss of heterozygosity; (c) in a programmed computer, using a trained algorithm that comprises one or more classifiers to process the first data set from (b) to generate a classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant; and (d) outputting a report indicative of the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant.

[0015] In some embodiments, the one or more classifiers comprises an ensemble classifier integrated with at least one index selected from the group consisting of: a follicular content index, a Hurthle cell index, and a Hurthle neoplasm index. In some embodiments, the one or more classifiers comprises an ensemble classifier integrated with a follicular content index, a Hurthle cell index, and a Hurthle neoplasm index.

[0016] In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a specificity of at least about 60%. In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a specificity of at least about 68%. In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a specificity of at least about 70%. In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a sensitivity of at least about 90%.

[0017] In some embodiments, the one or more classifiers further comprises one or more upstream classifiers, wherein the one or more upstream classifiers are selected from the group consisting of: a parathyroid classifier, a medullary thyroid cancer (MTC) classifier, a variant detection classifier, and a fusion transcript detection classifier. In some embodiments, the one or more classifiers comprises a parathyroid classifier that identifies a presence or an absence of a parathyroid tissue in the second portion of the tissue sample. In some embodiments, the upon identification of the

absence of the parathyroid tissue in the second portion of the tissue sample by the parathyroid classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the one or more classifiers comprises a medullary thyroid cancer (MTC) classifier that identifies a presence or an absence of a medullary thyroid cancer (MTC) in the second portion of the tissue sample. In some embodiments, the upon identification of the absence of the MTC in the second portion of the tissue sample by the MTC classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the one or more classifiers comprises a variant detection classifier that identifies a presence or an absence of a BRAF mutation in the second portion of the tissue sample. In some embodiments, the BRAF mutation is a BRAF V600E mutation. In some embodiments, the upon identification of the absence of the BRAF mutation in the second portion of the tissue sample by the variant detection classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the one or more classifiers comprises a fusion transcript detection classifier that identifies a presence or an absence of a RET/PTC gene fusion in the second portion of the tissue sample. In some embodiments, the RET/PTC gene fusion is RET/PTC1 or RET/PTC3 gene fusion. In some embodiments, the upon identification of the absence of the RET/PTC gene fusion in the second portion of the tissue sample by the fusion transcript detection classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the follicular content index identifies follicular content in the second portion of the tissue sample.

**[0018]** In some embodiments, the one or more classifiers of the trained algorithm comprises an ensemble classifier, wherein the ensemble classifier analyzes, in the first data set, sequence information corresponding to at least 500 genes of Table 3. In some embodiments, the one or more classifiers of the trained algorithm comprises ensemble classifier, wherein the ensemble classifier analyzes, in the first data set, sequence information corresponding to at least 1000 genes of Table 3. In some embodiments, the one or more classifiers of the trained algorithm comprises ensemble classifier, wherein the ensemble classifier analyzes, in the first data set, sequence information corresponding to 1115 genes of Table 3.

**[0019]** In some embodiments, the method further comprising (e) upon identifying the second portion of the tissue sample as being suspicious for malignancy, or malignant (i) processing the first data set to identify one or more genetic aberrations in one or more genes listed in FIG. 12; and (ii) outputting a second report indicative of a risk of malignancy, a histological subtype, and a prognosis associated with each of one or more genetic aberration identified in the second portion of the tissue sample. In some embodiments, the one or more genetic aberrations is a DNA variant. The method of claim 53, wherein the one or more genetic aberrations is a RNA fusion. In some embodiments, the risk of malignancy characterizes the one or more genetic aberrations as (1)

highly associated with malignant nodules, (2) associated with both benign and malignant nodules, or (3) has insufficient published evidence.

**[0020]** In some embodiments, the tissue sample is a thyroid tissue sample. In some embodiments, the tissue sample is a needle aspirate sample. In some embodiments, the needle aspirate sample is a fine needle aspirate sample. In some embodiments, the malignancy is thyroid cancer.

**[0021]** Another aspect of the present disclosure provides a method for processing or analyzing a tissue sample of a subject, comprising: (a) subjecting a first portion of the tissue sample to cytological analysis that indicates that the first portion of the sample is cytologically indeterminate; (b) upon identifying the first portion of the tissue sample as being cytologically indeterminate, assaying by sequencing, array hybridization, or nucleic acid amplification a plurality of gene expression products from a second portion of the tissue sample to yield a first data set; (c) in a programmed computer, using a trained algorithm that comprises one or more classifiers to process the first data set from (b) to generate a classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant with a specificity of at least about 60%; and (d) outputting a report indicative of the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant.

**[0022]** In some embodiments, the one or more classifiers comprises an ensemble classifier integrated with at least one index selected from the group consisting of: a follicular content index, a Hürthle cell index, and a Hurthle neoplasm index. In some embodiments, the one or more classifiers comprises an ensemble classifier integrated with a follicular content index, a Hurthle cell index, and a Hurthle neoplasm index. In some embodiments, the plurality of gene expression products include two or more of sequences corresponding to mRNA transcripts, mitochondrial transcripts, and chromosomal loss of heterozygosity.

**[0023]** In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a specificity of at least about 68%. In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a specificity of at least about 70%. In some embodiments, the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant has a sensitivity of at least about 90%.

**[0024]** In some embodiments, the one or more classifiers further comprises one or more upstream classifiers, wherein the one or more upstream classifiers are selected from the group consisting of a parathyroid classifier, a medullary thyroid cancer (MTC) classifier, a variant detection classifier, and a fusion transcript detection classifier. In some embodiments, the one or more classifiers comprises a parathyroid classifier that identifies a presence or an absence of a parathyroid tissue in the second portion of the tissue sample. In some embodiments, upon identification of the absence of the parathyroid tissue in the second portion of the tissue sample by the parathyroid classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the one or more classifiers comprises a medullary thyroid cancer (MTC) classifier that identifies a presence or an absence of a medullary thyroid cancer (MTC)

in the second portion of the tissue sample. In some embodiments, the upon identification of the absence of the MTC in the second portion of the tissue sample by the MTC classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the one or more classifiers comprises a variant detection classifier that identifies a presence or an absence of a BRAF mutation in the second portion of the tissue sample. In some embodiments, the BRAF mutation is a BRAF V600E mutation. In some embodiments, the upon identification of the absence of the BRAF mutation in the second portion of the tissue sample by the variant detection classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the one or more classifiers comprises a fusion transcript detection classifier that identifies a presence or an absence of a RET/PTC gene fusion in the second portion of the tissue sample. In some embodiments, the RET/PTC gene fusion is RET/PTC1 or RET/PTC3 gene fusion. In some embodiments, the upon identification of the absence of the RET/PTC gene fusion in the second portion of the tissue sample by the fusion transcript detection classifier, the at least one classifier of the one or more classifiers generates the classification of the second portion of the tissue sample as benign, suspicious for malignancy, or malignant. In some embodiments, the follicular content index identifies follicular content in the second portion of the tissue sample.

[0025] In some embodiments, the one or more classifiers of the trained algorithm comprises an ensemble classifier, wherein the ensemble classifier analyzes, in the first data set, sequence information corresponding to at least 500 genes of Table 3. In some embodiments, the one or more classifiers of the trained algorithm comprises an ensemble classifier, wherein the ensemble classifier analyzes, in the first data set, sequence information corresponding to at least 1000 genes of Table 3. In some embodiments, the one or more classifiers of the trained algorithm comprises an ensemble classifier, wherein the ensemble classifier analyzes, in the first data set, sequence information corresponding to 1115 genes of Table 3.

[0026] In some embodiments, the method further comprising (e) upon identifying the second portion of the tissue sample as being suspicious for malignancy, or malignant (i) processing the first data set to identify one or more genetic aberrations in one or more genes listed in FIG. 12; and (ii) outputting a second report indicative of a risk of malignancy, a histological subtype, and a prognosis associated with each of one or more genetic aberration identified in the second portion of the tissue sample. In some embodiments, the one or more genetic aberrations is a DNA variant. In some embodiments, the one or more genetic aberrations is a RNA fusion. In some embodiments, the risk of malignancy characterizes the one or more genetic aberrations as (1) highly associated with malignant nodules, (2) associated with both benign and malignant nodules, or (3) has insufficient published evidence.

[0027] In some embodiments, the tissue sample is a thyroid tissue sample. In some embodiments, the tissue sample is a needle aspirate sample. In some embodiments, the needle aspirate sample is a fine needle aspirate sample. In some embodiments, the malignancy is thyroid cancer.

[0028] Another aspect of the present disclosure provides a non-transitory computer readable medium comprising machine executable code that, upon execution by one or more computer processors, implements any of the methods above or elsewhere herein.

[0029] Another aspect of the present disclosure provides a system comprising one or more computer processors and computer memory coupled thereto. The computer memory comprises machine executable code that, upon execution by the one or more computer processors, implements any of the methods above or elsewhere herein.

[0030] Additional aspects and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative embodiments of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

#### INCORPORATION BY REFERENCE

[0031] All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference. To the extent publications and patents or patent applications incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0032] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0033] FIG. 1 is an illustration of Afirma gene sequencing classifier (“GSC”) system.

[0034] FIG. 2 illustrates Standard for Reporting of Diagnostic Accuracy Studies diagram of sample flow through the study.

[0035] FIG. 3 illustrates Afirma Genomic Sequencing Classifier (“GSC”) performance across differing risk populations.

[0036] FIG. 4 illustrates that Afirma GSC significantly improves specificity and high sensitivity.

[0037] FIG. 5 illustrates that in a comparison between Afirma GEC versus Afirma GSC, Afirma GSC shows significantly more benign results.

[0038] FIG. 6 illustrates treatment recommendations based on the results of Afirma GSC.

[0039] FIG. 7 illustrates that in a performance comparison between Afirma GEC versus Afirma GSC, GSC has a higher benign rate and PPV.

[0040] FIG. 8 illustrates analytical performance of Xpression Atlas.

[0041] FIG. 9 illustrates the diagnostic overview including Afirma GSC and Xpression Atlas.

[0042] FIG. 10 illustrates an example of an Xpression Atlas result.

[0043] FIG. 11 shows a computer system that is programmed or otherwise configured to implement methods provided herein.

[0044] FIG. 12 is a table listing certain genes identified as contributing to cancer diagnosis by molecular profiling.

#### DETAILED DESCRIPTION

[0045] While various embodiments of the invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions may occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed.

[0046] The term "subject," as used herein, generally refers to any animal or living organism. Animals can be mammals, such as humans, non-human primates, rodents such as mice and rats, dogs, cats, pigs, sheep, rabbits, and others. Animals can be fish, reptiles, or others. Animals can be neonatal, infant, adolescent, or adult animals. Humans can be more than about 1, 2, 5, 10, 20, 30, 40, 50, 60, 65, 70, 75, or about 80 years of age. The subject may have or be suspected of having a disease, such as cancer. The subject may be a patient, such as a patient being treated for a disease, such as a cancer patient. The subject may be predisposed to a risk of developing a disease such as cancer. The subject may be in remission from a disease, such as a cancer patient. The subject may be healthy.

[0047] The term "disease," as used herein, generally refers to any abnormal or pathologic condition that affects a subject. Examples of a disease include cancer, such as, for example, thyroid cancer, parathyroid cancer, lung cancer, skin cancer, and others. The disease may be treatable or non-treatable. The disease may be terminal or non-terminal. The disease can be a result of inherited genes, environmental exposures, or any combination thereof. The disease can be cancer, a genetic disease, a proliferative disorder, or others as described herein.

[0048] The term "sequence variant," "sequence variation," "sequence alteration" or "allelic variant," as used herein, generally refer to a specific change or variation in relation to a reference sequence, such as a genomic deoxyribonucleic acid (DNA) reference sequence, a coding DNA reference sequence, or a protein reference sequence, or others. The reference DNA sequence can be obtained from a reference database. A sequence variant may affect function. A sequence variant may not affect function. A sequence variant can occur at the DNA level in one or more nucleotides, at the ribonucleic acid (RNA) level in one or more nucleotides, at the protein level in one or more amino acids, or any combination thereof. The reference sequence can be obtained from a database such as the NCBI Reference Sequence Database (RefSeq) database. Specific changes that can constitute a sequence variation can include a substitution, a deletion, an insertion, an inversion, or a conversion in one or more nucleotides or one or more amino acids. A sequence variant may be a point mutation. A sequence variant may be a fusion gene. A fusion pair or a fusion gene may result from a sequence variant, such as a translocation, an interstitial deletion, a chromosomal inversion, or any combination thereof. A sequence variation can constitute

variability in the number of repeated sequences, such as triplications, quadruplications, or others. For example, a sequence variation can be an increase or a decrease in a copy number associated with a given sequence (i.e., copy number variation, or CNV). A sequence variation can include two or more sequence changes in different alleles or two or more sequence changes in one allele. A sequence variation can include two different nucleotides at one position in one allele, such as a mosaic. A sequence variation can include two different nucleotides at one position in one allele, such as a chimeric. A sequence variant may be present in a malignant tissue. A sequence variant may be present in a benign tissue. Absence of a variant may indicate that a tissue or sample is benign. As an alternative, absence of a variant may not indicate that a tissue or sample is benign.

[0049] The term "disease diagnostic," as used herein, generally refers to diagnosing or screening for a disease, to stratify a risk of occurrence of a disease, to monitor progression or remission of a disease, to formulate a treatment regime for the disease, or any combination thereof. A disease diagnostic can include a) obtaining information from one or more tissue samples from a subject, b) making a determination about whether the subject has a particular disease based on the information or tissue sample obtained, c) stratifying the risk of occurrence of the disease in the subject, d) confirming whether a subject has the disease, is developing the disease, or is in disease remission, or any combination thereof. The disease diagnostic may inform a particular treatment or therapeutic intervention for the disease. The disease diagnostic may also provide a score indicating for example, the severity or grade of a disease such as cancer, or the likelihood of an accurate diagnosis, such as via a p-value, a corrected p-value, or a statistical confidence indicator. The disease diagnostic may also indicate a particular type of a disease. For example, a disease diagnostic for thyroid cancer may indicate a subtype such as follicular adenoma (FA), nodular hyperplasia (NHP), lymphocytic thyroiditis (LCT), Hürthle cell adenoma (HA), follicular carcinoma (FC), papillary thyroid carcinoma (PTC), follicular variant of papillary carcinoma (FVPTC), medullary thyroid carcinoma (MTC), Hürthle cell carcinoma (HCC), anaplastic thyroid carcinoma (ATC), renal carcinoma (RCC), breast carcinoma (BCA), melanoma (MMN), B cell lymphoma (BCL), parathyroid (PTA), or hyperplasia papillary carcinoma (HPC).

#### Introduction

[0050] Some techniques for using preoperative genomic information for thyroid nodule differential diagnosis may involve use messenger RNA ("mRNA") transcript expression levels to categorize cytologically indeterminate FNAs as either benign or suspicious. Altered messenger RNA expression can occur for several reasons, including complex upstream interactions that occur because of sequence changes in key core genes or in relevant peripheral genes, the effect of epigenetic changes that occur without DNA sequence alterations, and both internal and external modifiers, such as inflammation and lifestyle or environment. Previously, in a cohort with a 24% prevalence of malignancy, a genome expression classifier ("GEC") accurately identified 90% of malignancies (i.e., sensitivity) and 52% of benign nodules (i.e., specificity) with indeterminate Bethesda III or IV cytology. It intentionally favored high sensitivity over specificity to ensure the accuracy and safety

of a benign genomic result. In GEC, a machine learning-derived classification algorithm uses messenger RNA transcript expression levels to categorize cytologically indeterminate samples as either benign or suspicious. A test, as described in the present disclosure, that has improved specificity for identification of benign nodules and maintained high sensitivity for malignancy detection may spare even more patients from surgery with an accurate benign genomic result (negative predictive value [NPV]) and increase the cancer yield among those with a suspicious result (positive predictive value [PPV]).

[0051] The present disclosure describes enhanced technologies for characterizing genomic information, including improved methods for the measurement of RNA transcriptome expression and sequencing of nuclear and mitochondrial RNAs, measurement changes in genomic copy number, including loss of heterozygosity, and the development of enhanced bioinformatics and machine learning strategies, resulting in a more robust genomic test.

#### Methods for Generating Classification for Tissue Samples for a Disease

[0052] The present disclosure provides methods for processing or analyzing a tissue sample of a subject to generate a classification of tissue sample as benign, suspicious for malignancy, or malignant. Such methods may comprise obtaining a plurality of gene expression products from a cytologically indeterminate tissue sample and using an algorithm to analyze the gene expression products to classify the tissue samples as benign, suspicious for malignancy, or malignant. In some cases, a plurality of gene expression products comprises sequences corresponding to mRNA transcripts, mitochondrial transcripts, chromosomal loss of heterozygosity, DNA variants and/or fusion transcripts. In some examples, the method uses a trained algorithm that comprises one or more classifiers and is implemented by one or more programmed computer processors to analyze the expression gene products to generate a classification of tissue sample as benign, suspicious for malignancy, or malignant. The algorithm may be a trained algorithm (e.g., an algorithm that is trained on at least 10, 200, 100 or 500 reference samples). Reference samples may be obtained from subjects having been diagnosed with the disease or from healthy subjects. The trained algorithm may analyze the sequence information of expression gene products corresponding to about 10,000 genes. The trained algorithm may analyze the sequence information of expression gene products corresponding to at least 500 genes of Table 3. The trained algorithm may analyze the sequence information of expression gene products corresponding to at least 600 genes of Table 3. The trained algorithm may analyze the sequence information of expression gene products corresponding to at least 700 genes of Table 3. The trained algorithm may analyze the sequence information of expression gene products corresponding to at least 800 genes of Table 3. The trained algorithm may analyze the sequence information of expression gene products corresponding to at least 900 genes of Table 3. The trained algorithm may analyze the sequence information of expression gene products corresponding to at least 1000 genes of Table 3. The trained algorithm may analyze the sequence information of expression gene products corresponding to at least 1100 genes of Table 3. The trained algorithm may analyze the

sequence information of expression gene products corresponding to at least 1200 genes of Table 3.

[0053] As set forth in the present disclosure, an expression level of one or more genes of gene expression products can be obtained by assaying for an expression level. Assaying may comprise array hybridization, nucleic acid sequencing, nucleic acid amplification, or others. Assaying may comprise sequencing, such as DNA or RNA sequencing. Such sequencing may be by next generation (NextGen) sequencing, such as high throughput sequencing or whole genome sequencing (e.g., Illumina). Such sequencing may include enrichment. Assaying may comprise reverse transcription polymerase chain reaction (PCR). Assaying may utilize markers, such as primers, that are selected for each of the one or more genes of the first or second sets of genes.

[0054] Additional methods for determining gene expression levels may include but are not limited to one or more of the following: additional cytological assays, assays for specific proteins or enzyme activities, assays for specific expression products including protein or RNA or specific RNA splice variants, in situ hybridization, whole or partial genome expression analysis, microarray hybridization assays, serial analysis of gene expression (SAGE), enzyme linked immuno-absorbance assays, mass-spectrometry, immunohistochemistry, blotting, sequencing, RNA sequencing, DNA sequencing (e.g., sequencing of complementary deoxyribonucleic acid (cDNA) obtained from RNA); next generation (Next-Gen) sequencing, nanopore sequencing, pyrosequencing, or Nanostring sequencing. Gene expression product levels may be normalized to an internal standard such as total messenger ribonucleic acid (mRNA) or the expression level of a particular gene.

[0055] The methods disclosed herein may include extracting and analyzing protein or nucleic acid (RNA or DNA) from one or more samples from a subject. Nucleic acids can be extracted from the entire sample obtained or can be extracted from a portion. In some cases, the portion of the sample not subjected to nucleic acid extraction may be analyzed by cytological examination or immunohistochemistry. Methods for RNA or DNA extraction from biological samples can include for example phenol-chloroform extraction (such as guanidinium thiocyanate phenol-chloroform extraction), ethanol precipitation, spin column-based purification, or others.

[0056] The sample obtained from the subject may be cytologically ambiguous or suspicious (or indeterminate). In some cases, the sample may be suggestive of the presence of a disease. The volume of sample obtained from the subject may be small, such as about 100 microliters, 50 microliters, 10 microliters, 5 microliters, 1 microliter or less. The sample may comprise a low quantity or quality of polynucleotides, such as a tissue sample with degraded or partially degraded RNA. For example, an FNA sample may yield low quantity or quality of polynucleotides. In such examples, the RNA Integrity Number (RIN) value of the sample may be about 9.0 or less. In some examples, the RIN value may be about 6.0 or less.

#### Risk of Malignancy Using Xpression Atlas

[0057] In some cases, the methods disclosed herein further comprise processing the gene expression products using a curated panel of sequence associated with variants and/or fusions and which includes well validated variants and variants whose clinical significance is emerging (such as, for

example the Xpression Atlas to provide further genomic information on samples identified as being suspicious for malignancy, or malignant, the method comprising identifying any one of the genetic aberrations disclosed in in one or more genes listed in FIG. 12 in the sample to indicate (i) risk of malignancy, (ii) a histological subtype, and (iii) prognosis associated with each of the genetic aberration identified in the sample (FIG. 9). In some examples, this may include identifying one or more genes, genetic aberrations of the one or more genes, or other genomic information disclosed in, for example, U.S. Pat. No. 8,541,170 and U.S. Patent Publication No. 2018/0016642, each of which is entirely incorporated herein by reference. Genetic aberrations may be any one or more of the DNA variants in one or more genes listed in FIG. 12. Genetic aberrations may be any one or more of the RNA fusions in one or more genes listed in FIG. 12. FIG. 10 is an example of an Xpression Atlas result that may be provided to the patient in conjunction with the GSC results on their samples to provide further genomic information comprising genetic aberrations identified in the samples and to indicate (i) risk of malignancy, (ii) a histological subtype, and (iii) prognosis associated with each of the genetic aberration identified in the sample. FIG. 8 illustrates the analytical performance of the 761 DNA variant panel and the 130 RNA fusion panel of Xpression Atlas.

[0058] The genetic aberrations may be validated or may have emerging clinical significance. The risk of malignancy may characterize one or more genetic aberrations as (1) highly associated with malignant nodules, (2) associated with both benign and malignant nodules, or (3) as having insufficient published evidence to characterize such risk. One or more genetic aberrations in one or more genes listed in FIG. 12 may be specific for cancer (e.g., malignancy). One or more genetic aberrations in one or more genes listed in FIG. 12 may occur in both benign and malignant samples.

[0059] The methods disclosed herein provide identifying one or more genetic aberrations in a sample that are indicative of a histological subtype. Histological subtypes may include classical parathyroid cancer (cPTC), infiltrative follicular variant of papillary thyroid carcinoma (infiltrative FVPTC), noninvasive encapsulated FVPTC (EFVPTC), Follicular thyroid carcinoma (FTC), and/or follicular adenomas (FA).

[0060] The methods disclosed herein comprise identifying one or more genetic aberrations in a sample to indicate prognosis associated with the genetic aberration. Prognostic information may comprise TNM stage and American Thyroid Association (ATA) risk. The TNM Staging System is based on the extent of the tumor (T), the extent of spread to the lymph nodes (N), and the presence of metastasis (M). The T category describes the original (primary) tumor. The TNM stage may comprise stages 1-4. ATA risk of recurrence staging system may comprises risk categories 1-3 which may correspond to low, intermediate, or high risk categories. The 761 nucleotide variant panel may have a PPA rate of at least 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more. The 130 fusion panel may have a PPA rate of at least 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more. Identification of one or more genetic aberrations may increase the risk of malignancy reported by one or more classifiers as used in the methods disclosed herein. Identification of one or more genetic aberrations may not increase the risk of

malignancy reported by one or more classifiers as used in the methods disclosed herein. A reported risk of malignancy generated by one or more classifiers of the present disclosure may not be reduced in some cases where no genetic aberrations in one or more genes listed in FIG. 12 are identified.

#### Samples

[0061] A sample obtained from a subject can comprise tissue, cells, cell fragments, cell organelles, nucleic acids, genes, gene fragments, expression products, gene expression products, gene expression product fragments or any combination thereof. A sample can be heterogeneous or homogeneous. A sample can comprise blood, urine, cerebrospinal fluid, seminal fluid, saliva, sputum, stool, lymph fluid, tissue, or any combination thereof. A sample can be a tissue-specific sample such as a sample obtained from a thyroid, skin, heart, lung, kidney, breast, pancreas, liver, muscle, smooth muscle, bladder, gall bladder, colon, intestine, brain, esophagus, or prostate.

[0062] A sample of the present disclosure can be obtained by various methods, such as, for example, fine needle aspiration (FNA), core needle biopsy, vacuum assisted biopsy, incisional biopsy, excisional biopsy, punch biopsy, shave biopsy, skin biopsy, or any combination thereof.

[0063] FNA, also referred to as fine needle aspirate biopsy (FNAB), or needle aspirate biopsy (NAB), is a method of obtaining a small amount of tissue from a subject. FNA can be less invasive than a tissue biopsy, which may require surgery and hospitalization of the subject to obtain the tissue biopsy. The needle of a FNA method can be inserted into a tissue mass of a subject to obtain an amount of sample for further analysis. In some cases, two needles can be inserted into the tissue mass. The FNA sample obtained from the tissue mass may be acquired by one or more passages of the needle across the tissue mass. In some cases, the FNA sample can comprise less than about 6×10<sup>6</sup>, 5×10<sup>6</sup>, 4×10<sup>6</sup>, 3×10<sup>6</sup>, 2×10<sup>6</sup>, 1×10<sup>6</sup> cells or less. The needle can be guided to the tissue mass by ultrasound or other imaging device. The needle can be hollow to permit recovery of the FNA sample through the needle by aspiration or vacuum or other suction techniques.

[0064] Samples obtained using methods disclosed herein, such as an FNA sample, may comprise a small sample volume. A sample volume may be less than about 500 microliters (uL), 400 uL, 300 uL, 200 uL, 100 uL, 75 uL, 50 uL, 25 uL, 20 uL, 15 uL, 10 uL, 5 uL, 1 uL, 0.5 uL, 0.1 uL, 0.01 uL or less. The sample volume may be less than about 1 uL. The sample volume may be less than about 5 uL. The sample volume may be less than about 10 uL. The sample volume may be less than about 20 uL. The sample volume may be between about 1 uL and about 10 uL. The sample volume may be between about 10 uL and about 25 uL.

[0065] Samples obtained using methods disclosed herein, such as an FNA sample, may comprise small sample weights. The sample weight, such as a tissue weight, may be less than about 100 milligrams (mg), 75 mg, 50 mg, 25 mg, 20 mg, 15 mg, 10 mg, 9 mg, 8 mg, 7 mg, 6 mg, 5 mg, 4 mg, 3 mg, 2 mg, 1 mg, 0.5 mg, 0.1 mg or less. The sample weight may be less than about 20 mg. The sample weight may be less than about 10 mg. The sample weight may be less than about 5 mg. The sample weight may be between about 5 mg and about 20 mg. The sample weight may be between about 1 mg and about 5 ng.

[0066] Samples obtained using methods disclosed herein, such as FNA, may comprise small numbers of cells. The number of cells of a single sample may be less than about  $10 \times 10^6$ ,  $5.5 \times 10^6$ ,  $5 \times 10^6$ ,  $4.5 \times 10^6$ ,  $4 \times 10^6$ ,  $3.5 \times 10^6$ ,  $3 \times 10^6$ ,  $2.5 \times 10^6$ ,  $2 \times 10^6$ ,  $1.5 \times 10^6$ ,  $1 \times 10^6$ ,  $0.5 \times 10^6$ ,  $0.2 \times 10^6$ ,  $0.1 \times 10^6$  cells or less. The number of cells of a single sample may be less than about  $5 \times 10^6$  cells. The number of cells of a single sample may be less than about  $4 \times 10^6$  cells. The number of cells of a single sample may be less than about  $3 \times 10^6$  cells. The number of cells of a single sample may be less than about  $2 \times 10^6$  cells. The number of cells of a single sample may be between about  $1 \times 10^6$  and about  $5 \times 10^6$  cells. The number of cells of a single sample may be between about  $1 \times 10^6$  and about  $10 \times 10^6$  cells.

[0067] Samples obtained using methods disclosed herein, such as FNA, may comprise small amounts of deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). The amount of DNA or RNA in an individual sample may be less than about 500 nanograms (ng), 400 ng, 300 ng, 200 ng, 100 ng, 75 ng, 50 ng, 45 ng, 40 ng, 35 ng, 30 ng, 25 ng, 20 ng, 15 ng, 10 ng, 5 ng, 1 ng, 0.5 ng, 0.1 ng, or less. The amount of DNA or RNA may be less than about 40 ng. The amount of DNA or RNA may be less than about 25 ng. The amount of DNA or RNA may be less than about 15 ng. The amount of DNA or RNA may be between about 1 ng and about 25 ng. The amount of DNA or RNA may be between about 5 ng and about 50 ng.

[0068] RNA yield or RNA amount of a sample can be measured in nanogram to microgram amounts. An example of an apparatus that can be used to measure nucleic acid yield in the laboratory is a NANODROP® spectrophotometer, QUBIT® fluorometer, or QUANTUSTM fluorometer. The accuracy of a NANODROP® measurement may decrease significantly with very low RNA concentration. Quality of data obtained from the methods described herein can be dependent on RNA quantity. Meaningful gene expression or sequence variant data or others can be generated from samples having a low or un-measurable RNA concentration as measured by NANODROP®. In some cases, gene expression or sequence variant data or others can be generated from a sample having an unmeasurable RNA concentration.

[0069] The methods as described herein can be performed using samples with low quantity or quality of polynucleotides, such as DNA or RNA. A sample with low quantity or quality of RNA can be for example a degraded or partially degraded tissue sample. A sample with low quantity or quality of RNA may be a fine needle aspirate (FNA) sample. The RNA quality of a sample can be measured by a calculated RNA Integrity Number (RIN) value. The RIN value is an algorithm for assigning integrity values to RNA measurements. The algorithm can assign a 1 to 10 RIN value, where an RIN value of 10 can be completely intact RNA. A sample as described herein that comprises RNA can have an RIN value of about 9.0, 8.0, 7.0, 6.0, 5.0, 4.0, 3.0, 2.0, 1.0 or less. In some cases, a sample comprising RNA can have an RIN value equal or less than about 8.0. In some cases, a sample comprising RNA can have an RIN value equal or less than about 6.0. In some cases, a sample comprising RNA can have an RIN value equal or less than about 4.0. In some cases, a sample can have an RIN value of less than about 2.0.

[0070] A sample, such as an FNA sample, may be obtained from a subject by another individual or entity, such as a

healthcare (or medical) professional or robot. A medical professional can include a physician, nurse, medical technician or other. In some cases, a physician may be a specialist, such as an oncologist, surgeon, or endocrinologist. A medical technician may be a specialist, such as a cytologist, phlebotomist, radiologist, pulmonologist or others. A medical professional may obtain a sample from a subject for testing or refer the subject to a testing center or laboratory for the submission of the sample. The medical professional may indicate to the testing center or laboratory the appropriate test or assay to perform on the sample, such as methods of the present disclosure including determining gene sequence data, gene expression levels, sequence variant data, or any combination thereof.

[0071] In some cases, a medical professional need not be involved in the initial diagnosis of a disease or the initial sample acquisition. An individual, such as the subject, may alternatively obtain a sample through the use of an over the counter kit. The kit may contain collection unit or device for obtaining the sample as described herein, a storage unit for storing the sample ahead of sample analysis, and instructions for use of the kit.

[0072] A sample can be obtained a) pre-operatively, b) post-operatively, c) after a cancer diagnosis, d) during routine screening following remission or cure of disease, e) when a subject is suspected of having a disease, f) during a routine office visit or clinical screen, g) following the request of a medical professional, or any combination thereof. Multiple samples at separate times can be obtained from the same subject, such as before treatment for a disease commences and after treatment ends, such as monitoring a subject over a time course. Multiple samples can be obtained from a subject at separate times to monitor the absence or presence of disease progression, regression, or remission in the subject.

#### Cytological Analysis

[0073] The methods as described herein may include cytological analysis of samples. Examples of cytological analysis include cell staining techniques and/or microscope examination performed by any number of methods and suitable reagents including but not limited to: eosin-azure (EA) stains, hematoxylin stains, CYTO-STAIN™, papanicolaou stain, eosin, nissl stain, toluidine blue, silver stain, azocarmine stain, neutral red, or janus green. More than one stain can be used in combination with other stains. In some cases, cells are not stained at all. Cells can be fixed and/or permeabilized with for example methanol, ethanol, glutaraldehyde or formaldehyde prior to or during the staining procedure. In some cases, the cells may not be fixed. Staining procedures can also be utilized to measure the nucleic acid content of a sample, for example with ethidium bromide, hematoxylin, nissl stain or any other nucleic acid stain.

[0074] Microscope examination of cells in a sample can include smearing cells onto a slide by standard methods for cytological examination. Liquid based cytology (LBC) methods may be utilized. In some cases, LBC methods provide for an improved approach of cytology slide preparation, more homogenous samples, increased sensitivity and specificity, or improved efficiency of handling of samples, or any combination thereof. In LBC methods, samples can be transferred from the subject to a container or vial containing a LBC preparation solution such as for example CYTYC

THINPREP®, SUREPATH™, or MONOPREP® or any other LBC preparation solution. Additionally, the sample may be rinsed from the collection device with LBC preparation solution into the container or vial to ensure substantially quantitative transfer of the sample. The solution containing the sample in LBC preparation solution may then be stored and/or processed by a machine or by one skilled in the art to produce a layer of cells on a glass slide. The sample may further be stained and examined under the microscope in the same way as a conventional cytological preparation.

[0075] Samples can be analyzed by immuno-histochemical staining. Immuno-histochemical staining can provide analysis of the presence, location, and distribution of specific molecules or antigens by use of antibodies in a sample (e.g., cells or tissues). Antigens can be small molecules, proteins, peptides, nucleic acids or any other molecule capable of being specifically recognized by an antibody. Samples may be analyzed by immuno-histochemical methods with or without a prior fixing and/or permeabilization step. In some cases, the antigen of interest may be detected by contacting the sample with an antibody specific for the antigen and then non-specific binding may be removed by one or more washes. The specifically bound antibodies may then be detected by an antibody detection reagent such as for example a labeled secondary antibody, or a labeled avidin/streptavidin. The antigen specific antibody can be labeled directly. Suitable labels for immunohistochemistry include but are not limited to fluorophores such as fluorescein and rhodamine, enzymes such as alkaline phosphatase and horse radish peroxidase, or radionuclides such as 32P and 125I. Gene product markers that may be detected by immuno-histochemical staining include but are not limited to Her2/Neu, Ras, Rho, EGFR, VEGFR, UbcH10, RET/PTC1, cytokeratin 20, calcitonin, GAL-3, thyroid peroxidase, or thyroglobulin.

[0076] Metrics associated with classifying a tissue sample as disclosed herein, such as sequences corresponding to mRNA transcripts, mitochondrial transcripts, and/or chromosomal loss of heterozygosity, need not be a characteristic of every cell of a sample found to comprise the tissue classification. Thus, the methods disclosed herein can be useful for classifying a tissue sample, e.g. as benign, suspicious for malignancy, or malignant for cancer, within a tissue where less than all cells within the sample exhibit a complete pattern of the gene expression levels or sequence variant data, or other data indicative of tissue classification. The gene expression levels, sequence variant data, or others may be either completely present, partially present, or absent within affected cells, as well as unaffected cells of the sample. The gene expression levels, sequence variant data, or others may be present in variable amounts within affected cells. The gene expression levels, sequence variant data, or others may be present in variable amounts within unaffected cells. In some cases, the gene expression levels of a first set of genes or the presence of one or more sequence variants in a second set of genes that correlates with a risk of malignancy occurrence can be positively detected. In some instances, positive detection can occur in at least 70%, 75%, 80%, 85%, 90%, 95%, or 100% of cells drawn from a sample. In some cases, the gene expression levels of a first set of genes or the presence of one or more sequence variants in a second set of genes can be absent. In some instances, absence of detection can occur in at least 70%, 75%, 80%,

85%, 90%, 95%, or 100% of cells of a corresponding normal or benign, non-disease sample.

[0077] Routine cytological or other assays may indicate a sample as negative (without disease), diagnostic (positive diagnosis for disease, such as cancer), ambiguous or suspicious (e.g., indeterminate) (suggestive of the presence of a disease, such as cancer), or non-diagnostic (providing inadequate information concerning the presence or absence of disease). The methods as described herein may confirm results from the routine cytological assessments or may provide an original assessment similar to a routine cytological assessment in the absence of one. The methods as described herein may classify a sample as malignant or benign, including samples found to be ambiguous, suspicious, or indeterminate. The methods may further stratify samples, such as samples known to be malignant, into low risk and medium-to-high risk groups of disease occurrence, including samples found to be ambiguous, suspicious, or indeterminate.

#### Markers for Array Hybridization, Sequencing, Amplification

[0078] Suitable reagents for conducting array hybridization, nucleic acid sequencing, nucleic acid amplification or other amplification reactions include, but are not limited to, DNA polymerases, markers such as forward and reverse primers, deoxynucleotide triphosphates (dNTPs), and one or more buffers. Such reagents can include a primer that is selected for a given sequence of interest, such as the one or more genes of the first set of genes and/or second set of genes.

[0079] In such amplification reactions, one primer of a primer pair can be a forward primer complementary to a sequence of a target polynucleotide molecule (e.g. the one or more genes of the first or second sets) and one primer of a primer pair can be a reverse primer complementary to a second sequence of the target polynucleotide molecule and a target locus can reside between the first sequence and the second sequence.

[0080] The length of the forward primer and the reverse primer can depend on the sequence of the target polynucleotide (e.g. the one or more genes of the first or second sets) and the target locus. In some cases, a primer can be greater than or equal to about 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 65, 70, 75, 80, 85, 90, 95, or about 100 nucleotides in length. As an alternative, a primer can be less than about 100, 95, 90, 85, 80, 75, 70, 65, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, or about nucleotides in length. In some cases, a primer can be about 15 to about 20, about 15 to about 25, about 15 to about 30, about 15 to about 40, about 15 to about 45, about 15 to about 50, about 15 to about 55, about 15 to about 60, about 20 to about 25, about 20 to about 30, about 20 to about 35, about 20 to about 40, about 20 to about 45, about 20 to about 50, about 20 to about 55, about 20 to about 60, about 20 to about 80, or about 20 to about 100 nucleotides in length.

[0081] Primers can be designed according to known parameters for avoiding secondary structures and self-hybridization, such as primer dimer pairs. Different primer

pairs can anneal and melt at about the same temperatures, for example, within 1° C., 2° C., 3° C., 4° C., 5° C., 6° C., 7° C., 8° C., 9° C. or 10° C. of another primer pair.

[0082] The target locus can be about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 100, 150, 200, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 650, 700, 750, 800, 850, 900 or 1000 nucleotides from the 3' ends or 5' ends of the plurality of template polynucleotides.

[0083] Markers (i.e., primers) for the methods described can be one or more of the same primer. In some instances, the markers can be one or more different primers such as about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more different primers. In such examples, each primer of the one or more primers can comprise a different target or template specific region or sequence, such as the one or more genes of the first or second sets.

[0084] One or more primers can comprise a fixed panel of primers. The one or more primers can comprise at least one or more custom primers. The one or more primers can comprise at least one or more control primers. The one or more primers can comprise at least one or more housekeeping gene primers. In some instances, the one or more custom primers anneal to a target specific region or complements thereof. The one or more primers can be designed to amplify or to perform primer extension, reverse transcription, linear extension, non-exponential amplification, exponential amplification, PCR, or any other amplification method of one or more target or template polynucleotides.

[0085] Primers can incorporate additional features that allow for the detection or immobilization of the primer but do not alter a basic property of the primer (e.g., acting as a point of initiation of DNA synthesis). For example, primers can comprise a nucleic acid sequence at the 5' end which does not hybridize to a target nucleic acid, but which facilitates cloning or further amplification, or sequencing of an amplified product. For example, the sequence can comprise a primer binding site, such as a PCR priming sequence, a sample barcode sequence, or a universal primer binding site or others.

[0086] A universal primer binding site or sequence can attach a universal primer to a polynucleotide and/or amplicon. Universal primers can include -47F (M13F), alfaMF, AOX3', AOX5', BGHR, CMV-30, CMV-50, CVMF, LACrmt, lamgda gt10F, lambda gt 10R, lambda gt11F, lambda gt11R, M13 rev, M13Forward(-20), M13Reverse, male, p10SEQPpQE, pA-120, pet4, pGAP Forward, pGLRVpr3, pGLpr2R, pKLAC14, pQEFS, pQERS, pucU1, pucU2, reversA, seqIREStam, seqIRESzpet, segori, seqPCR, seqpIRES-, seqpIRES+, seqpSecTag, seqpSec-Tag+, segretro+PSI, SP6, T3-prom, T7-prom, and T7-terminalInv. As used herein, attach can refer to both or either covalent interactions and noncovalent interactions. Attachment of the universal primer to the universal primer binding site may be used for amplification, detection, and/or sequencing of the polynucleotide and/or amplicon.

#### Trained Algorithm

[0087] The trained algorithm of the present disclosure can be trained using a set of samples, such as a sample cohort. The sample cohort can comprise about 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000 or more independent samples. The sample cohort can comprise about 100 independent samples. The sample cohort can comprise about 200 independent samples. The sample cohort can comprise between about 100 and about 700 independent samples. The independent samples can be from subjects having been diagnosed with a disease, such as cancer, from healthy subjects, or any combination thereof.

[0088] The sample cohort can comprise samples from about 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 or more different individuals. The sample cohort can comprise samples from about 100 different individuals. The sample cohort can comprise samples from about 200 different individuals. The different individuals can be individuals having been diagnosed with a disease, such as cancer, health individuals, or any combination thereof.

[0089] The sample cohort can comprise samples obtained from individuals living in at least 1, 2, 3, 4, 5, 6, 67, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, or 80 different geographical locations (e.g., sites spread out across a nation, such as the United States, across a continent, or across the world). Geographical locations include, but are not limited to, test centers, medical facilities, medical offices, post office addresses, cities, counties, states, nations, or continents. In some cases, a classifier that is trained using sample cohorts from the United States may need to be re-trained for use on sample cohorts from other geographical regions (e.g., India, Asia, Europe, Africa, etc.).

[0090] The trained algorithm may comprise one or more classifiers selected from the group consisting of a parathyroid classifier, a medullary thyroid cancer (MTC) classifier, a variant detection classifier, a fusion transcript detection classifier, an ensemble classifier, a follicular content index, and one or more Hurthle classifiers (e.g., a Hurthle cell index and/or a Hurthle neoplasm index). The ensemble classifier may be integrated with one or more index selected from the group consisting of a follicular content index, a Hurthle cell index, and a Hurthle neoplasm index. A parathyroid classifier may identify a presence or an absence of a parathyroid tissue in the tissue sample. A medullary thyroid cancer (MTC) classifier may identify a presence or an absence of a medullary thyroid cancer (MTC) in the tissue sample. A variant detection classifier may identify a presence or an absence of a BRAF mutation (such as BRAF V600E) in the tissue sample. A fusion transcript detection classifier may identify a presence or an absence of a RET/PTC gene fusion (such as RET/PTC1 and/or RET/PTC3 gene fusion) in the tissue sample. A follicular content index may identify follicular content in the tissue sample. A classifier may identify one or more TRK gene fusions and one or more RET alterations (e.g., a RET gene fusion).

[0091] The ensemble classifier may comprise 10,000 or more genes with a set of 1000 or more core genes. The 10,000 or more genes may improve the ensemble classifier stability against variability. The core genes may drive the prediction behavior of the ensemble model. The ensemble classifier may comprise or consist of 12 independent classifiers. The 12 independent classifiers may comprise or

consist of 6 elastic net logistic regression models and 6 support vector machine models. The 6 elastic net logistic regression models may each differ from one another according to the gene sets disclosed in Table 2. The 6 support vector machine models may each differ from one another according to the gene sets disclosed in Table 2. The ensemble classifier may analyze the sequence information of expression gene products corresponding to about 10,000 genes. The ensemble classifier may analyze the sequence information of expression gene products corresponding to at least 500 genes of Table 3. The ensemble classifier may analyze the sequence information of expression gene products corresponding to at least 600 genes of Table 3. The ensemble classifier may analyze the sequence information of expression gene products corresponding to at least 700 genes of Table 3. The ensemble classifier may analyze the sequence information of expression gene products corresponding to at least 800 genes of Table 3. The ensemble classifier may analyze the sequence information of expression gene products corresponding to at least 900 genes of Table 3. The ensemble classifier may analyze the sequence information of expression gene products corresponding to at least 1000 genes of Table 3. The ensemble classifier may analyze the sequence information of expression gene products corresponding to at least 1100 genes of Table 3. The ensemble classifier may analyze the sequence information of expression gene products corresponding to at least 1200 genes of Table 3.

[0092] In some embodiments, the specificity of the present method is at least 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more.

[0093] In some embodiments, the sensitivity of the present method is at least 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more.

[0094] In some embodiments, the specificity is greater than or equal to 60%. The negative predictive value (NPV) is greater than or equal to 95%. In some embodiments, the NPV is at least 95%, 95.5%, 96%, 96.5%, 97%, 97.5%, 98%, 98.5%, 99%, 99.5% or more.

[0095] Sensitivity typically refers to  $TP/(TP+FN)$ , where TP is true positive and FN is false negative. Number of Continued Indeterminate results divided by the total number of malignant results based on adjudicated histopathology diagnosis. Specificity typically refers to  $TN/(TN+FP)$ , where TN is true negative and FP is false positive. The number of actual benign results is divided by the total number of benign results based on adjudicated histopathology diagnosis. Positive Predictive Value (PPV) may be determined by:  $TP/(TP+FP)$ . Negative Predictive Value (NPV) may be determined by  $TN/(TN+FN)$ .

[0096] A biological sample may be identified as cancerous with an accuracy of greater than 75%, 80%, 85%, 90%, 95%, 99% or more. In some embodiments, the biological sample is identified as cancerous with a sensitivity of greater than 90%. In some embodiments, the biological sample is identified as cancerous with a specificity of greater than 60%. In some embodiments, the biological sample is identified as cancerous or benign with a sensitivity of greater than 90% and a specificity of greater than 60%. In some embodiments, the accuracy is calculated using a trained algorithm.

[0097] Results of the expression analysis of the subject methods may provide a statistical confidence level that a given diagnosis is correct. In some embodiments, such statistical confidence level is above 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 99.5%.

[0098] A trained algorithm may produce a unique output each time it is run. For example, using a different sample or plurality of samples with the same classifier can produce a unique output each time the classifier is run. Using the same sample or plurality of samples with the same classifier can produce a unique output each time the classifier is run. Using the same samples to train a classifier more than one time, may result in unique outputs each time the classifier is run.

[0099] Characteristics of a sample (e.g., sequence information corresponding to mRNA expression, mitochondrial transcripts, genetic variants and/or fusion transcripts) can be analyzed using an algorithm that comprises one or more classifiers and which is trained using one or more annotated reference sets. The identification can be performed by the classifier. More than one characteristic of a sample can be combined to generate classification of tissue sample. For example, sequence information corresponding to mRNA expression and mitochondrial transcripts can be combined and a classification can be generated from the combined data. The combining can be performed by the classifier. In another example, sequences obtained from a sample can be compared to a reference set to determine the presence of one or more sequence variants in a sample. In some cases, gene expression levels of one or more genes from a sample can be processed relative to expression levels of a reference set of genes that are used to train one or more classifiers to determine the presence of differential gene expression of one or more genes. A reference set can comprise one or more housekeeping genes. The reference set can comprise known sequence variants or expression levels of genes known to be associated with a particular disease or known to be associated with a non-disease state.

[0100] Classifiers of a trained algorithm can perform processing, combining, statistical evaluation, or further analysis of results, or any combination thereof. Separate reference sets may be provided for different features. For example, sequence variant data may be processed relative to a sequence variant data reference set. A gene expression level data may be processed relative to a gene expression level reference set. In some cases, multiple feature spaces may be processed with respect to the same reference set.

[0101] In some cases, sequence variants of a particular gene may or may not affect the gene expression level of that same gene. A sequence variant of a particular gene may affect the gene expression level of one or more different genes that may be located adjacent to and distal from the particular gene with the sequence variant. The presence of one or more sequence variants can have downstream effects on one or more genes. A sequence variant of a particular gene may perturb one or more signaling pathways, may cause ribonucleic acid (RNA) transcriptional regulation changes, may cause amplification of deoxyribonucleic acid (DNA), may cause multiple transcript copies to be produced, may cause excessive protein to be produced, may cause single base pairs, multi-base pairs, partial genes or one or more genes to be removed from the sequence.

[0102] Data from the methods described, such as gene expression levels or sequence variant data can be further analyzed using feature selection techniques such as filters

which can assess the relevance of specific features by looking at the intrinsic properties of the data, wrappers which embed the model hypothesis within a feature subset search, or embedded protocols in which the search for an optimal set of features is built into a classifier algorithm.

[0103] Filters useful in the methods of the present disclosure can include, for example, (1) parametric methods such as the use of two sample t-tests, analysis of variance (ANOVA) analyses, Bayesian frameworks, or Gamma distribution models (2) model free methods such as the use of Wilcoxon rank sum tests, between-within class sum of squares tests, rank products methods, random permutation methods, or threshold number of misclassification (TNoM) which involves setting a threshold point for fold-change differences in expression between two datasets and then detecting the threshold point in each gene that minimizes the number of mis-classifications or (3) multivariate methods such as bivariate methods, correlation based feature selection methods (CFS), minimum redundancy maximum relevance methods (MRMR), Markov blanket filter methods, and uncorrelated shrunken centroid methods. Wrappers useful in the methods of the present disclosure can include sequential search methods, genetic algorithms, or estimation of distribution algorithms. Embedded protocols can include random forest algorithms, weight vector of support vector machine algorithms, or weights of logistic regression algorithms.

[0104] Statistical evaluation of the results obtained from the methods described herein can provide a quantitative value or values indicative of one or more of the following: the classification of the tissue sample; the likelihood of diagnostic accuracy; the likelihood of disease, such as cancer; the likelihood of a particular disease, such as a tissue-specific cancer, for example, thyroid cancer; and the likelihood of the success of a particular therapeutic intervention. Thus a medical professional, who may not be trained in genetics or molecular biology, need not understand gene expression level or sequence variant data results. Rather, data can be presented directly to the medical professional in its most useful form to guide care or treatment of the subject. Statistical evaluation, combination of separate data results, and reporting useful results can be performed by the trained algorithm. Statistical evaluation of results can be performed using a number of methods including, but not limited to: the students T test, the two sided T test, pearson rank sum analysis, hidden markov model analysis, analysis of q-q plots, principal component analysis, one way analysis of variance (ANOVA), two way ANOVA, and the like. Statistical evaluation can be performed by the trained algorithm.

#### Diseases

[0105] A disease, as disclosed herein, can include thyroid cancer. Thyroid cancer can include any subtype of thyroid cancer, including but not limited to, any malignancy of the thyroid gland such as papillary thyroid cancer (PTC), follicular thyroid cancer (FTC), follicular variant of papillary thyroid carcinoma (FVPTC), medullary thyroid carcinoma (MTC), follicular carcinoma (FC), Hurthle cell carcinoma (HC), and/or anaplastic thyroid cancer (ATC). In some cases, the thyroid cancer can be differentiated. In some cases, the thyroid cancer can be undifferentiated.

[0106] A thyroid tissue sample can be classified using the methods of the present disclosure as comprising one or more

benign or malignant tissue types (e.g. a cancer subtype), including but not limited to follicular adenoma (FA), nodular hyperplasia (NHP), lymphocytic thyroiditis (LCT), and Hurthle cell adenoma (HA), follicular carcinoma (FC), papillary thyroid carcinoma (PTC), follicular variant of papillary carcinoma (FVPTC), medullary thyroid carcinoma (MTC), Hürthle cell carcinoma (HC), and anaplastic thyroid carcinoma (ATC), renal carcinoma (RCC), breast carcinoma (BCA), melanoma (MMN), B cell lymphoma (BCL), or parathyroid (PTA).

#### Monitoring of Subjects or Therapeutic Interventions Via Molecular Profiling

[0107] In the methods of the present disclosure, a subject may be monitored. For example, a subject may be diagnosed with cancer. This initial diagnosis may or may not involve the use of methods disclosed herein. The subject may be prescribed a therapeutic intervention such as a thyroidectomy for a subject suspected of having thyroid cancer. The results of the therapeutic intervention may be monitored on an ongoing basis by methods disclosed herein to detect the efficacy of the therapeutic intervention. In another example, a subject may be diagnosed with a benign tumor or a precancerous lesion or nodule, and the tumor, nodule, or lesion may be monitored on an ongoing basis by methods disclosed herein to detect any changes in the state of the tumor or lesion.

[0108] Methods disclosed herein may also be used to ascertain the potential efficacy of a specific therapeutic intervention prior to administering to a subject. For example, a subject may be diagnosed with cancer. A genomic sequence classifier (GSC) classifier along with Xpression Atlas may indicate a presence of at least one variant associated with highly malignant tumors. In such cases, therapeutic intervention may be customized to the results obtained. A tumor sample may be obtained and cultured in vitro using methods known to the art.

#### Computer Systems

[0109] The present disclosure provides computer systems that are programmed to implement methods of the disclosure. FIG. 11 shows a computer system 1101 that is programmed or otherwise configured to implement the trained algorithm for the genomic sequencing classifier and/or the Xpression atlas. The computer system 1101 can regulate various aspects of the methods of the present disclosure, such as, for example, nucleic acid sequencing methods, interpretation of nucleic acid sequencing data and analysis of cellular nucleic acids, such as RNA (e.g., mRNA), and characterization of samples from sequencing data. The computer system 1101 can be an electronic device of a user or a computer system that is remotely located with respect to the electronic device. The electronic device can be a mobile electronic device.

[0110] The computer system 1101 includes a central processing unit (CPU, also "processor" and "computer processor" herein) 1105, which can be a single core or multi core processor, or a plurality of processors for parallel processing. The computer system 1101 also includes memory or memory location 1110 (e.g., random-access memory, read-only memory, flash memory), electronic storage unit 1115 (e.g., hard disk), communication interface 1120 (e.g., network adapter) for communicating with one or more other

systems, and peripheral devices **1125**, such as cache, other memory, data storage and/or electronic display adapters. The memory **1110**, storage unit **1115**, interface **1120** and peripheral devices **1125** are in communication with the CPU **1105** through a communication bus (solid lines), such as a motherboard. The storage unit **1115** can be a data storage unit (or data repository) for storing data. The computer system **1101** can be operatively coupled to a computer network ("network") **1130** with the aid of the communication interface **1120**. The network **1130** can be the Internet, an internet and/or extranet, or an intranet and/or extranet that is in communication with the Internet. The network **1130** in some cases is a telecommunication and/or data network. The network **1130** can include one or more computer servers, which can enable distributed computing, such as cloud computing. The network **1130**, in some cases with the aid of the computer system **1101**, can implement a peer-to-peer network, which may enable devices coupled to the computer system **1101** to behave as a client or a server.

[0111] The CPU **1105** can execute a sequence of machine-readable instructions, which can be embodied in a program or software. The instructions may be stored in a memory location, such as the memory **1110**. The instructions can be directed to the CPU **1105**, which can subsequently program or otherwise configure the CPU **1105** to implement methods of the present disclosure. Examples of operations performed by the CPU **1105** can include fetch, decode, execute, and writeback.

[0112] The CPU **1105** can be part of a circuit, such as an integrated circuit. One or more other components of the system **1101** can be included in the circuit. In some cases, the circuit is an application specific integrated circuit (ASIC).

[0113] The storage unit **1115** can store files, such as drivers, libraries and saved programs. The storage unit **1115** can store user data, e.g., user preferences and user programs. The computer system **1101** in some cases can include one or more additional data storage units that are external to the computer system **1101**, such as located on a remote server that is in communication with the computer system **1101** through an intranet or the Internet.

[0114] The computer system **1101** can communicate with one or more remote computer systems through the network **1130**. For instance, the computer system **1101** can communicate with a remote computer system of a user (e.g., medical professional, or subject). Examples of remote computer systems include personal computers (e.g., portable PC), slate or tablet PC's (e.g., Apple® iPad, Samsung® Galaxy Tab), telephones, Smart phones (e.g., Apple® iPhone, Android-enabled device, Blackberry®), or personal digital assistants. The user can access the computer system **1101** via the network **1130**.

[0115] Methods as described herein can be implemented by way of machine (e.g., computer processor) executable code stored on an electronic storage location of the computer system **1101**, such as, for example, on the memory **1110** or electronic storage unit **1115**. The machine executable or machine readable code can be provided in the form of software. During use, the code can be executed by the processor **1105**. In some cases, the code can be retrieved from the storage unit **1115** and stored on the memory **1110** for ready access by the processor **1105**. In some situations, the electronic storage unit **1115** can be precluded, and machine-executable instructions are stored on memory **1110**.

[0116] The code can be pre-compiled and configured for use with a machine having a processor adapted to execute the code, or can be compiled during runtime. The code can be supplied in a programming language that can be selected to enable the code to execute in a pre-compiled or as-compiled fashion.

[0117] Aspects of the systems and methods provided herein, such as the computer system **1101**, can be embodied in programming. Various aspects of the technology may be thought of as "products" or "articles of manufacture" typically in the form of machine (or processor) executable code and/or associated data that is carried on or embodied in a type of machine readable medium. Machine-executable code can be stored on an electronic storage unit, such as memory (e.g., read-only memory, random-access memory, flash memory) or a hard disk. "Storage" type media can include any or all of the tangible memory of the computers, processors or the like, or associated modules thereof, such as various semiconductor memories, tape drives, disk drives and the like, which may provide non-transitory storage at any time for the software programming. All or portions of the software may at times be communicated through the Internet or various other telecommunication networks. Such communications, for example, may enable loading of the software from one computer or processor into another, for example, from a management server or host computer into the computer platform of an application server. Thus, another type of media that may bear the software elements includes optical, electrical and electromagnetic waves, such as used across physical interfaces between local devices, through wired and optical landline networks and over various air-links. The physical elements that carry such waves, such as wired or wireless links, optical links or the like, also may be considered as media bearing the software. As used herein, unless restricted to non-transitory, tangible "storage" media, terms such as computer or machine "readable medium" refer to any medium that participates in providing instructions to a processor for execution.

[0118] Hence, a machine readable medium, such as computer-executable code, may take many forms, including but not limited to, a tangible storage medium, a carrier wave medium or physical transmission medium. Non-volatile storage media include, for example, optical or magnetic disks, such as any of the storage devices in any computer(s) or the like, such as may be used to implement the databases, etc. shown in the drawings. Volatile storage media include dynamic memory, such as main memory of such a computer platform. Tangible transmission media include coaxial cables; copper wire and fiber optics, including the wires that comprise a bus within a computer system. Cater-wave transmission media may take the form of electric or electromagnetic signals, or acoustic or light waves such as those generated during radio frequency (RF) and infrared (IR) data communications. Common forms of computer-readable media therefore include for example: a floppy disk, a flexible disk, hard disk, magnetic tape, any other magnetic medium, a CD-ROM, DVD or DVD-ROM, any other optical medium, punch cards paper tape, any other physical storage medium with patterns of holes, a RAM, a ROM, a PROM and EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave transporting data or instructions, cables or links transporting such a carrier wave, or any other medium from which a computer may read programming code and/or data. Many of these forms of computer readable

media may be involved in carrying one or more sequences of one or more instructions to a processor for execution. [0119] The computer system 1101 can include or be in communication with an electronic display 1135 that comprises a user interface (UI) 1140 for providing, for example, results of nucleic acid sequencing, analysis of nucleic acid sequencing data, characterization of nucleic acid sequencing samples, tissue characterizations, etc. Examples of UI's include, without limitation, a graphical user interface (GUI) and web-based user interface.

[0120] Methods and systems of the present disclosure can be implemented by way of one or more algorithms. An algorithm can be implemented by way of software upon execution by the central processing unit 1105. The algorithm can, for example, initiate nucleic acid sequencing, process nucleic acid sequencing data, interpret nucleic acid sequencing results, characterize nucleic acid samples, characterize samples, etc.

## EXAMPLES

### Example 1. Training and Validation Cohorts

[0121] This study describes the blinded clinical validation of a genomic sequence classifier (GSC), implemented in accordance with the methods described herein, on a prospective multicenter-derived set of patients with FNA samples whose referral to surgery and histopathological diagnosis were determined in the absence of genomic information.

[0122] The study was approved by institution-specific institutional review boards as well as by Liberty IRB (DeLand, Florida; now Chesapeake IRB) and Copernicus Group Independent Review Board (Cary, North Carolina). All patients provided written informed consent prior to participating in the study.

[0123] The following thyroid nodule FNA samples were included in the training set, with each sample set being independent from one another (Table 1):

#### ENHANCE Arm 1:

[0124] A dedicated molecular sample was obtained when the cytology specimen was collected from a nodule  $\geq 1$  cm during clinical care. Arm 2 samples were all unoperated, Bethesda II, or Bethesda III/IV and GEC benign, and lacked 2015 American Thyroid Association high suspicion sonographic pattern findings. Additionally, they had clinical follow-up (mean 23 months, range 17-32) and either a repeat FNA that was cytology benign, or had no growth (<50%

increase in volume or <20% increase in 2 or more dimensions) or development of high suspicion ultrasound findings after the initial FNA. Nodules were excluded from Arm 2 if repeat FNA was Bethesda V or VI, GEC suspicious, or they underwent surgery. Arm 2 nodules served as truly benign samples, recognizing that GEC benign samples were underrepresented among operated Arm 1 samples.

#### ENHANCE Arm 2:

[0125] A dedicated molecular sample was obtained when the cytology specimen was collected from a nodule  $\geq 1$  cm during clinical care. Arm 2 samples were all unoperated, Bethesda II, or Bethesda III/IV and GEC benign, and lacked 2015 American Thyroid Association high suspicion sonographic pattern findings. Additionally, they had clinical follow-up (mean 23 months, range 17-32) and either a repeat FNA that was cytology benign, or had no growth (<50% increase in volume or <20% increase in 2 or more dimensions) or development of high suspicion ultrasound findings after the initial FNA. Nodules were excluded from Aim 2 if repeat FNA was Bethesda V or VI, GEC suspicious, or they underwent surgery. Arm 2 nodules served as truly benign samples, recognizing that GEC benign samples were underrepresented among operated Arm 1 samples.

#### VERA-CVP (Non Cyto-I) Samples:

[0126] Samples described in the clinical validation of the Afirma GEC1 with sufficient materials remaining. Only Bethesda II, V, and VI samples with histopathology labels defined by an expert panel of pathologists were allowed in the training set. 60% of these samples were randomly chosen into the training set.

#### VERA-Train:

[0127] Samples used in the training set of the Afirma GEC1

#### VERA-Extra:

[0128] Collected and associated with histopathology labels identically to VERA-CVP, but these samples were not used in the training or validation of the Afirma GEC.

#### CLIA-GEC B:

[0129] Samples from the CLIA stream that are GEC Benign. These samples do not have long term follow-up or a histopathology label. Their benign GEC prediction is used as a surrogate label in algorithm training.

TABLE 1

Cohort	Composition of the core ensemble model training set.						Total
	Bethesda II	Bethesda III	Bethesda IV	Bethesda V	Bethesda VI	Bethesda NA	
ENHANCE Arm 1	8	209	76	5	10	0	308
ENHANCE Arm 2	4	50	14	0	0	0	68
VERA-CVP	23	0	0	33	29	0	85
VERA-Extra	1	4	4	6	1	0	16

TABLE 1-continued

Cohort	Composition of the core ensemble model training set.						Total
	Bethesda II	Bethesda III	Bethesda IV	Bethesda V	Bethesda VI	Bethesda NA	
VERA-Train	0	4	6	7	16	13	46
CLIA-GEC B	0	47	7	0	0	57	111
Total (Proportion)	36 (5.7%)	314 (49.5%)	107 (16.9%)	51 (8.0%)	56 (8.8%)	70 (11%)	634

#### Example 2. Validation Cohort

[0130] Dedicated thyroid nodule FNA specimens and surgical histopathology from nodules 1 cm or larger were collected using a prospective and blinded protocol at 49 academic and community centers in the United States from patients 21 years or older. These samples, stored at -80° C., were previously used to validate the GEC. The details of their enrollment and prespecified inclusion and exclusion criteria have been reported elsewhere. Histopathology diagnoses were previously established by an expert panel of thyroid surgical histopathologists that were blinded to all clinical and molecular data. BRAF V600E DNA mutational reference status was established by testing DNA from all samples with the competitive allele-specific TaqMan polymerase chain reaction, as described below. This independent validation cohort was prespecified and divided into a primary test set comprised of all patients with Bethesda III and IV samples described in the clinical validation of the Afirma GEC with sufficient RNA remaining and a secondary test set comprised of all patients with Bethesda II, V, or VI samples described in the clinical validation of the Afirma GEC with sufficient RNA remaining and not randomly assigned to the training set, as described in Example 1 above.

#### Reference Methods:

[0131] BRAF V600E status—BRAF V600E status was determined from genomic DNA using Competitive Allele Specific Taqman PCR (castPCR™, Thermo Fisher, Waltham, MA) for BRAF 1799T>A mutation, as previously described. Briefly, genomic DNA was purified with the AllPrep Micro Kit (Qiagen, Hilden, Germany) and quantified with Quanti-iT PicoGreen dsDNA Assay Kit (Thermo Fisher, Waltham, MA). Five ng of DNA was tested with wild-type and mutant assays on an ABI7900HT. Samples were labelled BRAF V600E positive if the variant allele frequency was ≥5% and wild type if the allele frequency was <5%.

[0132] Medullary Thyroid Cancer—Histopathology diagnoses, including medullary thyroid cancer, were previously established by an expert panel of thyroid histopathologists while blinded to all clinical and molecular data.

#### Example 3. Blinding of the Independent Test Set

[0133] The following steps were implemented to ensure the independent test set was securely blinded throughout algorithm development and validation.

[0134] First, each step was documented in a prespecified protocol and time-stamped on execution. Each team member was assigned a single role and allowed access only to information designated for that role. A randomly generated blinded identification number was assigned to each sample in the validation set by information technology engineers who operated independently of all other teams to ensure that all other personnel were unable to link clinical and genomic data. All historic information that may potentially reveal the clinical label on the independent test set was secured in a password-protected folder prior to the start of algorithm development. Information technology engineers conducted performance testing of the validation test set independently of all other teams.

#### Example 4. RNA Purification

[0135] RNA was purified with the AllPrep Micro kit (Qiagen, Hilden, Germany) as previously described. RNA was quantified using the QuantiFluor RNA System (Promega, Madison, WI). Fluorescence was read with a Tecan Infinite 200 Pro plate reader (Tecan, Mannedorf, Switzerland). RNA Integrity Number was determined with the Bioanalyzer 2100 (Agilent, Santa Clara, CA).

#### Example 5. Library Preparation

[0136] Samples were randomized and plated into 96 well plates according to their random order. Each plate contained Universal Human Reference RNA (Agilent, Santa Clara, CA), a benign thyroid tissue control sample, a malignant thyroid tissue control sample, a medullary thyroid carcinoma tissue control sample and 6 FNAs that were run on every plate in the study. Additionally, 3 samples from each plate were randomly selected to be included as technical replicates.

[0137] 15 ng of total RNA was transferred to a 96 well plate. The TruSeq RNA Access Library Preparation Kit (Illumina, San Diego, CA) was adapted for use on the Microlab STAR robotics platform (Hamilton, Reno, NV). During library preparation, total RNA is fragmented, reverse transcribed, end-repaired, A-tailed, and Illumina adapters with individual indexes are ligated. Following PCR and AMPure XP (Beckman Coulter, Indianapolis, IN) cleanup, library size and quantity was determined with the Fragment Analyzer (Advanced Analytical, Ankeny, IA). 250 ng of 4 libraries were combined and sequentially captured with the human exome to remove ribosomal RNA, intronic, and intergenic sequences. Following PCR and AMPure XP

(Beckman Coulter, Indianapolis, IN) cleanup, library size and quantity were determined with the Bioanalyzer 2100 (Agilent, Santa Clara, CA).

#### Example 6. Next-Generation Sequencing

**[0138]** Libraries were normalized to 2 nM, pooled to 16 samples per sequencing run, and denatured according to the manufacturer's instructions. 1% phiX library (Illumina, San Diego, CA) was spiked into each sequencing run. Denatured and diluted libraries were loaded onto NextSeq 500 machines (Illumina, San Diego, CA) and sequenced with a NextSeq v2 High Output 150 cycle kit (Illumina, San Diego, CA) for paired end 2×76 cycle sequencing. Sequencing runs were required to have >75% of bases ≥Q30 and <1% phiX error rate.

uniformity of base coverage, and base duplication and mismatch rates. All these QC metrics were generated using RNA-SeQC. Any sample that failed a QC metric was reprocessed from total RNA through library preparation and sequencing if sufficient RNA was available. Only samples passing the quality criteria were used for downstream analysis.

#### Example 8. Algorithm Development

**[0141]** Fine-needle aspiration samples (n=634) were used to build the GSC core ensemble model, as described in Example 1. The ensemble model consists of 12 independent classifiers: 6 are elastic net logistic regression models and 6 are support vector machines. The 6 models within each category differ from each other according to the gene sets used (Table 2).

TABLE 2

Feature sets used in each classifier within the final ensemble model.		
Feature set name	Description of feature set	Size
DE-significant	Top significant genes at FDR-adjusted p-value < 0.05 based on differential expression analysis using DESeq2 package	10,158
HOPACH50perc	HOPACH clustering was done on top 2,000 significant genes, then within each cluster, top 50% genes were retrieved	998
HOPACH10perc	HOPACH clustering was done on top 2,000 significant genes, then within each cluster, top 10% genes were retrieved	196
GEC	Among the 142 genes used by Afirma GEC main classifier, 140 genes <sup>②</sup> were targeted by RNA-sequencing	140
GEC-HOPACH50perc	Union of 'GEC' and 'HOPACH50perc' sets	1,115
GEC-HOPACH10perc	Union of 'GEC' and 'HOPACH10perc' sets	327

FDR—false discovery rate

<sup>②</sup> indicates text missing or illegible when filed

#### Example 7. RNA Sequencing Pipeline, Feature Extraction, and Quality Control

**[0139]** RNA-seq data was used to generate gene expression counts, identify variants, detect fusion-pairs, and calculate loss of heterozygosity (LOH) statistics. Raw sequencing data (FASTQ file) was aligned to human reference genome assembly 37 (Genome Reference Consortium) using STAR RNA-seq aligner. Expression counts were obtained by HTSeq5 and normalized using DESeq2 accounting for sequencing depth and gene-wise variability. Variants were identified using GATK variant calling pipeline, and fusion-pairs detected using STAR-Fusion. A loss of heterozygosity (LOH) statistic at chromosome and genome level was developed using variants identified genome-wide. The statistic quantifies the magnitude of LOH by calculating the proportion of variants that have a variant allele frequency (VAF; fraction of reads carrying the alternative allele) away from 0.5 (<0.2 or >0.8) after pre-filtering of variants that has a VAF exactly at zero or one, or is located in cytoband regions exhibiting abnormal excess of LOH signatures across all training samples.

**[0140]** To exclude low quality samples from downstream analysis, quality metrics were evaluated against pre-specified acceptance metrics for total numbers of sequenced and uniquely mapped reads, the overall proportion of exonic reads among mapped, the mean per-base coverage, the

**[0142]** To minimize overfitting and to accurately reflect classifier performance incorporating random noise, hyper-parameter tuning and model selections were performed using repeated nested cross-validation. Hyperparameter tuning was performed within the inner layer of the cross-validation, and the classifier performance was summarized using the outer layer of the 5-fold cross-validation repeated 40 times. For each classifier, the decision boundary was chosen to optimize specificity, with a minimum requirement of 90% sensitivity to detect malignancy.

**[0143]** The locked ensemble model uses a total of 10 196 genes, among which are 1115 core genes (Table 3). These core genes drive the prediction behavior of the model, and the remaining genes improve classifier stability against assay variability.

**[0144]** In addition to the ensemble model described above, the Afirma GSC system includes 7 other components: a parathyroid cassette, a medullary thyroid cancer (MTC) cassette, a BRAFV600E cassette, RET/PTC1 and RET/PTC3 fusion detection modules, follicular content index, Hurthle cell index, and Hurthle neoplasm index. The first 4 are upstream of the ensemble classifier, targeting specific and rare patient subgroups (FIG. 1). The last 3 (the follicular content index, Hurthle cell index, and the Hurthle neoplasm index) were developed to further improve the benign vs suspicious classification performance. They were incorporated with the ensemble classifier to form the core benign vs suspicious classifier engine.

TABLE 3

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000121270	ABCC11	16	48200821	48281479
ENSG00000173208	ABCD2	12	39943835	40013553
ENSG00000144827	ABHD10	3	111697857	111712210
ENSG00000136379	ABHD17C	15	80972025	81047962
ENSG00000166016	ABTB2	11	34172535	34379555
ENSG00000222482	AC005071.1	7	99817650	99817743
ENSG00000235978	AC018816.3	3	4855978	4928977
ENSG00000215067	AC027763.2	17	6779954	6915668
ENSG00000177076	ACER2	9	19408925	19452018
ENSG00000078124	ACER3	11	76571911	76737841
ENSG00000151726	ACSL1	4	185676749	185747972
ENSG00000184009	ACTG1	17	79476997	79490873
ENSG00000130402	ACTN4	19	39138289	39222223
ENSG00000115073	ACTR1B	2	98272431	98280570
ENSG00000115170	ACVR1	2	158592958	158732374
ENSG00000143537	ADAM15	1	155023042	155035252
ENSG00000163638	ADAMTS9	3	64501333	64673676
ENSG00000065457	ADAT1	16	75630879	75657198
ENSG00000155897	ADCY8	8	131792547	132054672
ENSG00000156110	ADK	10	75910960	76469061
ENSG00000163485	ADORA1	1	203059782	203136533
ENSG00000196526	AFAP1	4	7760441	7941653
ENSG00000144218	AFF3	2	100162323	100759201
ENSG00000038002	AGA	4	178351924	178363657
ENSG00000188157	AGRN	1	955503	991496
ENSG00000124942	AHNAK	11	62201016	62323707
ENSG00000185567	AHNAK2	14	105403581	105444694
ENSG00000173209	AHSA2	2	61404553	61418338
ENSG00000163568	AIM2	1	159032274	159116886
ENSG00000106305	AIMP2	7	6048876	6063465
ENSG00000129474	AJUBA	14	23440383	23451851
ENSG00000108599	AKAP10	17	19807615	19881656
ENSG00000214239	AL591025.1	6	159047471	159049322
ENSG00000137124	ALDH1B1	9	38392661	38398658
ENSG00000159063	ALG8	11	77811982	77850706
ENSG00000110497	AMBRA1	11	46417964	46615675
ENSG00000144233	AMMECR1L	2	128619204	128643496
ENSG00000126016	AMOT	X	112017731	112084043
ENSG00000131503	ANKHD1	5	139781399	139929163
ENSG00000144504	ANKMY1	2	241418839	241508626
ENSG00000167522	ANKRD11	1	89334038	89556969
ENSG00000174501	ANKRD36C	2	96514587	96657541
ENSG00000135299	ANKRD6	6	90142889	90343553
ENSG00000163297	ANTXR2	4	80822303	81046608
ENSG00000135046	ANXA1	9	75766673	75785309
ENSG00000103723	AP3B2	1	83328033	83378666
ENSG00000157823	AP3S2	1	90373831	90437574
ENSG00000011132	APBA3	1	3750817	3761697
ENSG00000113108	APBB3	5	139937853	139973337
ENSG00000100823	APEX1	1	20923350	20925927
ENSG00000117362	APH1A	1	150237804	150241980
ENSG00000084234	APLP2	1	129939732	130014699
ENSG00000095139	ARCN1	1	118443105	118473748
ENSG00000134884	ARGLU1	1	107194021	107220512
ENSG00000225485	ARHGAP23	1	36584662	36668628
ENSG00000177479	ARIH2	3	48956254	49023815
ENSG00000169379	ARL13B	3	93698983	93774512
ENSG00000170632	ARMC10	7	102715328	102740205
ENSG00000118690	ARMC2	6	109169619	109295186
ENSG00000169126	ARMC4	1	28064115	28287977
ENSG00000102401	ARMCX3	X	100877787	100882833
ENSG00000198960	ARMCX6	X	100870110	100872991
ENSG00000241553	ARPC4	3	9834179	9849410
ENSG00000197070	ARRDC1	9	140500106	140509812
ENSG00000151693	ASAP2	2	9346894	9545812
ENSG00000148331	ASB6	9	132399171	132404444
ENSG00000112249	ASCC3	6	100956070	101329248
ENSG00000141505	ASGR1	1	7076750	7082883
ENSG00000106819	ASPN	9	95218487	95244788
ENSG00000034533	ASTE1	3	130732719	130746493
ENSG00000119778	ATAD2B	2	23971534	24149984
ENSG00000145782	ATG12	5	115163893	115177555
ENSG00000138363	ATIC	2	216176540	216214487

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000068650	ATP11A	1	113344643	113541482
ENSG00000127249	ATP13A4	3	193119866	193310900
ENSG00000175054	ATR	3	142168077	142297668
ENSG00000224470	ATXN1L	1	71879894	71919171
ENSG00000158321	AUTS2	7	69063905	70258054
ENSG00000179913	B3GNT3	1	17905637	17923891
ENSG00000175711	B3GNTL1	1	80900031	81009686
ENSG00000105393	BABAM1	1	17378159	17392058
ENSG00000186318	BACE1	1	117156402	117186975
ENSG00000166170	BAG5	1	104022881	104029168
ENSG00000140320	BAHD1	1	40731920	40760441
ENSG00000135298	BAI3	6	69345259	70099403
ENSG00000175334	BANF1	1	65769550	65771620
ENSG00000172530	BANP	1	87982850	88110924
ENSG00000171552	BCL2L1	2	30252255	30311792
ENSG00000116128	BCL9	1	147013182	147098017
ENSG00000123095	BHLHE41	1	26272959	26278060
ENSG00000168487	BMP1	8	22022249	22069839
ENSG00000125378	BMP4	1	54416454	54425479
ENSG00000204217	BMPR2	2	203241659	203432474
ENSG00000163141	BNPL	1	151009046	151020076
ENSG00000038219	BOD1L1	4	13570362	13629347
ENSG00000133639	BTG1	1	92536286	92539673
ENSG00000186265	BTLA	3	112182815	112218408
ENSG00000155640	C10orf12	1	98741041	98745582
ENSG00000158636	C11orf30	1	76155967	76264069
ENSG00000149179	C11orf49	1	46958240	47185936
ENSG00000110696	C11orf58	1	16634679	16778428
ENSG00000166352	C11orf74	1	36616051	36694823
ENSG00000173715	C11orf80	1	66511922	66610987
ENSG00000133935	C14orf1	1	76116134	76127532
ENSG00000179933	C14orf119	1	23563974	23569665
ENSG00000133943	C14orf159	1	91526677	91691976
ENSG00000168260	C14orf183	1	50550369	50559361
ENSG00000246223	C14orf64	1	98391947	98444461
ENSG00000166780	C16orf45	1	15528152	15718885
ENSG00000185905	C16orf54	1	29753784	29757327
ENSG00000205710	C17orf107	1	4802713	4806227
ENSG00000196544	C17orf59	1	8091652	8093564
ENSG00000104979	C19orf53	1	13884982	13889276
ENSG00000162817	C1orf115	1	220863187	220872499
ENSG00000182795	C1orf116	1	207191866	207206101
ENSG00000143612	C1orf43	1	154179182	154193104
ENSG00000111731	C2CD5	1	22601517	22697480
ENSG00000119147	C2orf40	2	106679702	106694615
ENSG00000118961	C2orf43	2	20883788	21022882
ENSG00000159239	C2orf81	2	74641304	74648718
ENSG00000125730	C3	1	6677715	6730573
ENSG00000244731	C4A	6	31949801	31970458
ENSG00000224389	C4B	6	31982539	32003195
ENSG00000181751	C5orf30	5	102594403	102614361
ENSG00000205765	C5orf51	5	41904290	41921738
ENSG00000203872	C6orf163	6	88054567	88075181
ENSG00000204387	C6orf48	6	31802385	31807541
ENSG00000146963	C7orf55-	7	139025105	139108198
ENSG00000253250	C8orf88	8	91970865	91997485
ENSG00000136932	C9orf156	9	100666771	100684852
ENSG00000238227	C9orf69	9	139006427	139010731
ENSG00000063180	CA11	1	49141199	49149569
ENSG00000182985	CADM1	1	115039938	115375675
ENSG00000162545	CAMK2N1	1	20808884	20812713
ENSG00000111530	CAND1	1	67663061	67713731
ENSG00000142126	CAPN1	1	64948037	64979477
ENSG00000135387	CAPRIN1	1	34073230	34122703
ENSG00000110888	CAPRIN2	1	30862486	30907885
ENSG00000105483	CARD8	1	48684027	48759203
ENSG00000105974	CAV1	7	116164839	116201233
ENSG00000188649	CC2D2B	1	97733786	97792441
ENSG00000169193	CCDC126	7	23636998	23684327
ENSG00000244607	CCDC13	3	42734155	42814745
ENSG00000004766	CCDC132	7	92861653	92988338
ENSG000000135205	CCDC146	7	76751751	76958850
ENSG000000153237	CCDC148	2	159027593	159313265

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000159588	CCDC17	1	46085716	46089729
ENSG00000216937	CCDC7	1	32735068	32863492
ENSG00000091986	CCDC80	3	112323407	112368377
ENSG00000149231	CCDC82	1	96085933	96123087
ENSG00000172724	CCL19	9	34689564	34691274
ENSG00000110092	CCND1	1	69455855	69469242
ENSG00000118971	CCND2	1	4382938	4414516
ENSG00000134480	CCNH	5	86687311	86708836
ENSG00000163660	CCNL1	3	156864297	156878549
ENSG00000260916	CCPG1	1	55632230	55700708
ENSG00000115484	CCT4	2	62095224	62115939
ENSG00000135624	CCT7	2	73460548	73480149
ENSG00000177697	CD151	1	832843	839831
ENSG00000198087	CD2AP	6	47445525	47594999
ENSG00000169217	CD2BP2	1	30362087	30366682
ENSG00000135218	CD36	7	79998891	80308593
ENSG00000117877	CD3EAP	1	45909467	45914024
ENSG00000026508	CD44	1	35160417	35253949
ENSG00000169442	CD52	1	26644448	26647014
ENSG00000153283	CD96	3	111011566	111384597
ENSG00000105401	CDC37	1	10501810	10530797
ENSG00000171219	CDC42BPG	1	64590859	64612041
ENSG00000128283	CDC42EP1	2	37956454	37965412
ENSG00000179604	CDC42EP4	1	71279763	71308314
ENSG00000140937	CDH11	1	64977656	65160015
ENSG00000166589	CDH16	1	66942025	66952887
ENSG00000124215	CDH26	2	58533471	58609066
ENSG00000062038	CDH3	1	68670092	68756519
ENSG00000179242	CDH4	2	59827482	60515673
ENSG00000065883	CDK13	7	39989636	40136733
ENSG00000136861	CDK5RAP2	9	123151147	123342448
ENSG00000134058	CDK7	5	68530668	68573250
ENSG00000100490	CDKL1	1	50796310	50883179
ENSG0000006837	CDKL3	5	133541305	133706738
ENSG0000007129	CEACAM21	1	42055886	42093197
ENSG00000102901	CENPT	1	67862060	67881714
ENSG00000174799	CEP135	4	56815037	56899529
ENSG00000126001	CEP250	2	34042985	34099804
ENSG00000198707	CEP290	1	88442793	88535993
ENSG00000183137	CEP57L1	6	109416313	109485135
ENSG00000111860	CEP85L	6	118781935	119031238
ENSG00000000971	CFH	1	196621008	196716634
ENSG00000205403	CFI	4	110661852	110723335
ENSG00000163320	CGGBP1	3	88101094	88199035
ENSG00000111642	CHD4	1	6679249	6716642
ENSG00000072609	CHFR	1	133398773	133532890
ENSG00000109220	CHIC2	4	54875956	54930857
ENSG00000115526	CHST10	2	101008327	101034118
ENSG00000175040	CHST2	3	142838173	142841800
ENSG00000138615	CILP	1	65488337	65503826
ENSG00000141076	CIRH1A	1	69165194	69265033
ENSG00000125931	CITED1	X	71521488	71527037
ENSG00000273192	CITF22-1A6.3	2	50295876	50298224
ENSG00000104859	CLASRP	1	45542298	45574214
ENSG00000163347	CLDN1	3	190023490	190040264
ENSG00000113946	CLDN16	3	190040330	190129932
ENSG00000189143	CLDN4	7	73213872	73247014
ENSG00000105270	CLIP3	1	36505562	36524245
ENSG00000179335	CLK3	1	74890841	74932057
ENSG00000188603	CLN3	1	28477983	28506896
ENSG00000049656	CLPTM1L	5	1317859	1345214
ENSG00000171603	CLSTN1	1	9789084	9884584
ENSG00000120885	CLU	8	27454434	27472548
ENSG00000170293	CMTM8	3	32280171	32411817
ENSG00000117519	CNN3	1	95362507	95392834
ENSG00000080802	CNOT4	7	135046547	135194875
ENSG00000173786	CNP	1	40118759	40129749
ENSG00000144810	COL8A1	3	99357319	99518070
ENSG00000171812	COL8A2	1	36560837	36590821
ENSG00000169019	COMMID8	4	47452885	47465736
ENSG00000129083	COPB1	1	14464986	14521573
ENSG00000184432	COPB2	3	139074442	139108574
ENSG00000115520	COQ10B	2	198318147	198340032

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000109472	CPE	4	166282346	166419472
ENSG00000117322	CR2	1	207627575	207663240
ENSG00000166426	CRABP1	1	78632666	78640572
ENSG00000169372	CRADD	1	94071151	94288616
ENSG00000095794	CREM	1	35415719	35501886
ENSG00000006016	CRLF1	1	18683030	18718551
ENSG00000175315	CST6	1	65779312	65780976
ENSG00000102974	CTCF	1	67596310	67673086
ENSG00000183248	CTD-	1	7933605	7939326
ENSG00000044115	CTNNA1	5	137946656	138270723
ENSG00000066032	CTNNA2	2	79412357	80875905
ENSG00000119326	CTNNAL1	9	111704851	111775809
ENSG00000168036	CTNNB1	3	41236328	41301587
ENSG00000085733	CTTN	1	70244510	70282690
ENSG00000044090	CUL7	6	43005355	43021683
ENSG00000108296	CWC25	1	36956687	36981734
ENSG00000168329	CX3CR1	3	39304985	39323226
ENSG00000156234	CXCL13	4	78432907	78532988
ENSG00000145824	CXCL14	5	134906373	134914969
ENSG00000103018	CYB5B	1	69458428	69500169
ENSG00000166394	CYB5R2	1	7686331	7698453
ENSG00000172115	CYCS	7	25159710	25164980
ENSG00000142973	CYP4B1	1	47223510	47285085
ENSG00000152207	CYSLTR2	1	49280951	49283498
ENSG00000108669	CYTH1	1	76670130	76778379
ENSG00000153071	DAB2	5	39371780	39462402
ENSG00000136848	DAB2IP	9	124329336	124547809
ENSG00000115827	DCAF17	2	172290727	172341562
ENSG00000057019	DCBLD2	3	98514785	98620533
ENSG00000164935	DCSTAMP	8	105351315	105368917
ENSG00000150401	DCUN1D2	1	114110134	114145267
ENSG00000178404	DDC8	1	76866992	76899299
ENSG00000197312	DDI2	1	15943995	15995539
ENSG00000089737	DDX24	1	94517266	94547591
ENSG00000145833	DDX46	5	134094469	134190823
ENSG00000118197	DDX59	1	200593024	200639097
ENSG00000160570	DEDD2	1	42702750	42724292
ENSG00000164825	DEFB1	8	6728097	6735544
ENSG00000105339	DENNND3	8	142127377	142205907
ENSG00000174839	DENNND6A	3	57611184	57678816
ENSG00000023697	DERA	1	16064106	16190220
ENSG00000183628	DGCR6	2	18893541	18901751
ENSG00000157680	DGKI	7	137065783	137531838
ENSG00000172893	DHCR7	1	71139239	71163914
ENSG00000167536	DHRS13	1	27224799	27230089
ENSG00000162496	DHRS3	1	12627939	12677737
ENSG00000160305	DIP2A	2	47878812	47989926
ENSG00000162595	DIRAS3	1	68511645	68517314
ENSG00000164741	DLC1	8	12940870	13373167
ENSG00000198947	DMD	X	31115794	33357558
ENSG00000114841	DNAH1	3	52350335	52434507
ENSG00000138246	DNAJC13	3	132136370	132257876
ENSG00000179532	DNHD1	1	6518490	6614988
ENSG00000088387	DOCK9	1	99445741	99738879
ENSG00000125170	DOK4	1	57505863	57521239
ENSG00000197635	DPP4	2	162848751	162931052
ENSG00000130226	DPP6	7	153584182	154685995
ENSG00000162961	DPY30	2	32092878	32264881
ENSG00000113657	DPYSL3	5	146770374	146889619
ENSG00000175550	DRAP1	1	65686728	65689032
ENSG00000096696	DSP	6	7541808	7586950
ENSG00000110042	DTX4	1	58938903	58976060
ENSG00000120875	DUSP4	8	29190581	29208185
ENSG00000138166	DUSP5	1	112257596	112271302
ENSG00000107404	DVL1	1	1270656	1284730
ENSG00000077380	DYNC1I2	2	172543919	172604930
ENSG00000146425	DYNLT1	6	159057506	159065771
ENSG00000145088	EAF2	3	121554030	121605373
ENSG00000255423	EBLN2	3	73110810	73112488
ENSG00000117298	ECE1	1	21543740	21671997
ENSG00000143369	ECM1	1	150480538	150486265
ENSG00000203734	ECT2L	6	139117063	139225207
ENSG00000151617	EDNRA	4	148402069	148466106

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000156508	EEF1A1	6	74225473	74233520
ENSG00000178852	EFCAB13	1	45400656	45518678
ENSG00000215529	EFCAB8	2	31446729	31549006
ENSG00000172638	EFEMP2	1	65633912	65641063
ENSG00000142634	EFHD2	1	15736391	15756839
ENSG00000169242	EFNA1	1	155099936	155107333
ENSG00000090776	EFNB1	X	68048840	68061990
ENSG00000138798	EGF	4	110834040	110933422
ENSG00000120738	EGR1	5	137801179	137805004
ENSG00000115504	EHBP1	2	62900986	63273622
ENSG00000024422	EHD2	1	48216600	48246391
ENSG00000204371	EHMT2	6	31847536	31865464
ENSG00000084623	EIF3I	1	32687529	32697205
ENSG00000156976	EIF4A2	3	186500994	186507689
ENSG00000109381	ELF2	4	139949266	140098372
ENSG00000163435	ELF3	1	201977073	201986316
ENSG000001267767	ELK1	X	47494920	47510003
ENSG00000155849	ELMO1	7	36893961	37488852
ENSG00000102890	ELMO3	1	67233014	67237932
ENSG00000213853	EMP2	1	10622279	10674555
ENSG00000131355	EMR3	1	14729929	14800839
ENSG00000149218	ENDOD1	1	94822974	94865809
ENSG00000167280	ENGASE	1	77071021	77084681
ENSG00000167302	ENTHD2	1	79202077	79212891
ENSG00000183317	EPHA10	1	38179552	38230805
ENSG00000142627	EPHA2	1	16450832	16482582
ENSG00000116106	EPHA4	2	222282747	222438922
ENSG00000182580	EPHB3	3	184279572	184300197
ENSG00000227184	EPPK1	8	144939497	144952632
ENSG00000151491	EPS8	1	15773092	16035263
ENSG00000065361	ERBB3	1	56473641	56497289
ENSG00000104714	ERICH1	8	564746	688106
ENSG00000107566	ERLIN1	1	101909851	101948091
ENSG00000116285	ERRF1	1	8064464	8086368
ENSG00000091831	ESR1	6	151977826	152450754
ENSG00000105755	ETHE1	1	44010871	44031396
ENSG00000143845	ETNK2	1	204100190	204121307
ENSG00000175832	ETV4	1	41605212	41656988
ENSG00000167880	EVPL	1	74000583	74023533
ENSG00000170323	FABP4	8	82390654	82395498
ENSG00000103876	FAH	1	80444832	80479288
ENSG00000183688	FAM101B	1	289769	295730
ENSG00000136830	FAM129B	9	130267618	130341268
ENSG00000152380	FAM151B	5	79783788	79838382
ENSG00000146067	FAM193B	5	176946789	176981542
ENSG00000198673	FAM19A2	1	62102040	62672931
ENSG00000108950	FAM20A	1	66531254	66597530
ENSG00000205085	FAM71F2	7	128312342	128326929
ENSG00000126882	FAM78A	9	134133463	134151934
ENSG00000162981	FAM84A	2	14772810	14790933
ENSG00000171262	FAM98B	1	38746328	38779911
ENSG00000197601	FAR1	1	13690217	13753893
ENSG00000146267	FAXC	6	99719045	99797938
ENSG00000170271	FAXDC2	5	154198051	154238812
ENSG00000142449	FBN3	1	8130286	8214730
ENSG00000116661	FBXO2	1	11708424	11715842
ENSG00000135108	FBXO21	1	117581146	117628336
ENSG00000181617	FDCSP	4	71091788	71100969
ENSG00000214814	FER1L6	8	124864227	125132302
ENSG00000113578	FGF1	5	141971743	142077617
ENSG00000138685	FGF2	4	123747863	123819391
ENSG00000127951	FGL2	7	76822688	76829143
ENSG00000125848	FLRT3	2	14303634	14318262
ENSG00000115414	FN1	2	216225163	216300895
ENSG00000115226	FNDC4	2	27714750	27718112
ENSG00000137166	FOXP4	6	41514164	41570122
ENSG00000171049	FPR2	1	52255279	52273779
ENSG00000150893	FREM2	1	39261266	39460074
ENSG00000111816	FRK	6	116252312	116381921
ENSG00000172159	FRMD3	9	85857905	86153461
ENSG00000139926	FRMD6	1	51955818	52197445
ENSG00000075539	FRYL	4	48499378	48782339
ENSG00000070404	FSTL3	1	676392	683385

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000137726	FXYD6	1	117707693	117748201
ENSG00000157240	FZD1	7	90893783	90898123
ENSG00000164930	FZD6	8	104310661	104345094
ENSG00000155760	FZD7	2	202899310	202903160
ENSG00000123689	G0S2	1	209848765	209849733
ENSG00000136928	GABBR2	9	101050391	101471479
ENSG00000145864	GABRB2	5	160715436	160976050
ENSG00000182256	GABRG3	1	27216429	27778373
ENSG00000116717	GADD45A	1	68150744	68154021
ENSG00000197093	GAL3ST4	7	99756867	99766373
ENSG00000117308	GALE	1	24122089	24127271
ENSG00000119514	GALNT12	9	101569981	101612363
ENSG00000109586	GALNT7	4	174089904	174245118
ENSG00000114480	GBE1	3	81538850	81811312
ENSG0000006625	GGCT	7	30536237	30591095
ENSG00000146830	GIGYF1	7	100277130	100287071
ENSG00000213203	GIMAP1	7	150413645	150421372
ENSG00000106560	GIMAP2	7	150382785	150390729
ENSG00000133574	GIMAP4	7	150264365	150271041
ENSG00000145723	GIN1	5	102421704	102455855
ENSG00000139436	GIT2	1	110367607	110434194
ENSG00000187513	GJA4	1	35258599	35261348
ENSG00000188910	GJB3	1	35246790	35251970
ENSG00000166105	GLB1L3	1	134144139	134189458
ENSG00000186417	GLDN	1	51633826	51700210
ENSG00000250571	GLI4	8	144349603	144359101
ENSG00000135423	GLS2	1	56864736	56882198
ENSG00000063169	GLTSCR1	1	48111453	48206533
ENSG00000168237	GLYCTK	3	52321105	52329272
ENSG00000130755	GMFG	1	39818993	39833012
ENSG00000204590	GNL1	6	30509154	30524951
ENSG00000130119	GNL3L	X	54556644	54587504
ENSG00000136935	GOLGA1	9	127640646	127710771
ENSG00000174567	GOLT1A	1	204167288	204183220
ENSG00000115806	GORASP2	2	171784974	171823639
ENSG00000120053	GOT1	1	101156627	101190381
ENSG00000204438	GPANK1	6	31629006	31634060
ENSG00000089916	GPATCH2L	1	76618259	76720685
ENSG00000183484	GPR132	1	105515728	105531782
ENSG00000163328	GPR155	2	175296966	175351822
ENSG00000143147	GPR161	1	168053997	168106821
ENSG00000147138	GPR174	X	78426469	78427726
ENSG00000166073	GPR176	1	40091233	40213093
ENSG00000188394	GPR21	9	125796806	125797975
ENSG00000167191	GPRC5B	1	19868616	19897489
ENSG00000141738	GRB7	1	37894180	37903544
ENSG00000158055	GRHL3	1	24645812	24690972
ENSG00000148180	GSN	9	123970072	124095121
ENSG00000172986	GXYLT2	3	72937224	73047289
ENSG00000113088	GZMK	5	54320081	54330398
ENSG00000214367	HAUS3	4	2229191	2243891
ENSG00000068024	HDAC4	2	239969864	240323348
ENSG00000173064	HECTD4	1	112597992	112819896
ENSG00000198265	HELZ	1	65066554	65242105
ENSG00000103657	HERC1	1	63900817	64126141
ENSG00000135547	HEY2	6	126068810	126082415
ENSG00000163909	HEYL	1	40089825	40105617
ENSG00000165102	HGSNAT	8	42995556	43057998
ENSG00000196312	HIATL2	9	99660348	99775862
ENSG00000169567	HINT1	5	130494720	130507428
ENSG00000204632	HLA-G	6	29794744	29798902
ENSG00000149948	HMGAA2	1	66217911	66360075
ENSG00000189403	HMBG1	1	31032884	31191734
ENSG00000198830	HMGN2	1	26798941	26802463
ENSG00000177733	HNRNPA0	5	137087075	137090039
ENSG00000127483	HP1BP3	1	21069154	21113816
ENSG00000116983	HPCAL4	1	40144320	40157361
ENSG00000105707	HPN	1	35531410	35557475
ENSG00000025423	HSD17B6	1	57145945	57181574
ENSG00000096384	HSP90AB1	6	44214824	44221620
ENSG00000113013	HSPA9	5	137890571	137911133
ENSG00000068001	HYAL2	3	50355221	50360337
ENSG00000242028	HYPK	1	44088340	44095241

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000105376	ICAM5	1	10400657	10407454
ENSG00000116237	ICMT	1	6281253	6296032
ENSG00000115738	ID2	2	8818975	8824583
ENSG00000188483	IER5L	9	131937835	131940540
ENSG0000010295	IFFO1	1	6647541	6665239
ENSG00000114446	IFT57	3	107879659	107941417
ENSG00000073792	IGF2BP2	3	185361527	185542844
ENSG00000115461	IGFBP5	2	217536828	217560248
ENSG00000167779	IGFBP6	1	53491220	53496129
ENSG00000182700	IGIP	5	139505521	139508391
ENSG00000147255	IGSF1	X	130407480	130533677
ENSG00000162729	IGSF8	1	160061130	160068733
ENSG00000104365	IKBKB	8	42128820	42189973
ENSG0000030419	IKZF2	2	213864429	214017151
ENSG00000144730	IL17RD	3	57124010	57204334
ENSG00000115602	IL1RL1	2	102927962	102968497
ENSG00000134352	IL6ST	5	55230923	55290821
ENSG00000168685	IL7R	5	35852797	35879705
ENSG00000143621	ILF2	1	153634512	153643524
ENSG00000178035	IMPDH2	3	49061758	49066841
ENSG00000163083	INHBB	2	121103719	121109384
ENSG00000241644	INMT	7	30737601	30797218
ENSG00000185085	INTS5	1	62414320	62420774
ENSG00000164941	INTS8	8	95825539	95893974
ENSG0000074706	IPCEF1	6	154475631	154677926
ENSG00000205339	IPO7	1	9406169	9469673
ENSG00000132321	IQCA1	2	237232794	237416185
ENSG00000145703	IQGAP2	5	75699074	76003957
ENSG00000665883	ISOC1	5	128430444	128449721
ENSG00000105655	ISYNA1	1	18545198	18549111
ENSG00000164171	ITGA2	5	52285156	52390609
ENSG0000005884	ITGA3	1	48133332	48167845
ENSG00000135424	ITGA7	1	56078352	56109827
ENSG00000144668	ITGA9	3	37493606	37865005
ENSG00000132470	ITGB4	1	73717408	73753899
ENSG00000105855	ITGB8	7	20370325	20455377
ENSG00000135916	ITM2C	2	231729354	231743963
ENSG00000086544	ITPKC	1	41223008	41246765
ENSG00000096433	ITPR3	6	33588142	33664351
ENSG00000205730	ITPR1PL2	1	19125254	19132946
ENSG0000077684	JADE1	4	129730779	129796379
ENSG00000102221	JADE3	X	46771711	46920641
ENSG00000171135	JAGN1	3	9932238	9936033
ENSG00000171988	JMD1C	1	64926981	65225722
ENSG00000130522	JUND	1	18390563	18392432
ENSG00000197256	KANK2	1	11274943	11308467
ENSG00000114982	KANSL3	2	97258907	97308524
ENSG00000177272	KCNA3	1	111214310	111217655
ENSG00000151704	KCNJ1	1	128706210	128737268
ENSG00000124249	KCNK15	2	43374421	43379675
ENSG00000164626	KCNK5	6	39156749	39197226
ENSG00000184156	KCNQ3	8	133133108	133493200
ENSG00000174943	KCTD13	1	29916333	29938356
ENSG00000100196	KDELR3	2	38864067	38879452
ENSG00000004487	KDM1A	1	23345941	23410182
ENSG00000127663	KDM4B	1	4969125	5153606
ENSG00000117139	KDM5B	1	202696526	202778598
ENSG00000165757	KIAA1462	1	30301729	30404423
ENSG00000134444	KIAA1468	1	59854491	59974355
ENSG00000166004	KIAA1731	1	93394805	93463522
ENSG00000173214	KIAA1919	6	111580551	111592370
ENSG00000157404	KIT	4	55524085	55606881
ENSG00000102554	KLF5	1	73629114	73651676
ENSG00000162873	KLHD8A	1	205305220	205326218
ENSG00000129451	KLK10	1	51515995	51523431
ENSG00000169035	KLK7	1	51479729	51487355
ENSG00000139187	KLRG1	1	9102640	9163356
ENSG0000025800	KPNA6	1	32573639	32642169
ENSG00000111057	KRT18	1	53342655	53346685
ENSG00000171345	KRT19	1	39679869	39684560
ENSG00000157992	KRTCAP3	2	27665233	27669348
ENSG00000141068	KSR1	1	25783670	25953461
ENSG00000159166	LAD1	1	201342372	201368736

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000196878	LAMB3	1	209788215	209825811
ENSG00000135862	LAMC1	1	182992595	183114727
ENSG00000058085	LAMC2	1	183155373	183214035
ENSG00000068697	LAPTM4A	2	20232411	20251789
ENSG00000107929	LARP4B	1	855484	977564
ENSG00000135338	LCA5	6	80194708	80247175
ENSG00000205629	LCMT1	1	25123050	25189552
ENSG00000136167	LCP1	1	46700055	46786006
ENSG00000182195	LDOC1	X	140269934	140271310
ENSG00000225880	LINC00115	1	761586	762902
ENSG00000260032	LINC00657	2	34633544	34638882
ENSG00000163898	LIPH	3	185224050	185270401
ENSG00000131899	LLGL1	1	18128901	18148189
ENSG00000168216	LMBRD1	6	70385694	70507003
ENSG00000160789	LMNA	1	156052364	156109880
ENSG00000048540	LMO3	1	16701307	16763528
ENSG00000143013	LMO4	1	87794151	87814606
ENSG00000170500	LONRF2	2	100889753	100939195
ENSG00000167210	LOXHD1	1	44056935	44236996
ENSG00000186001	LRCH3	3	197518097	197615307
ENSG00000077454	LRCH4	7	100169855	100183776
ENSG00000147650	LRP12	8	105501459	105601252
ENSG00000168702	LRP1B	2	140988992	142889270
ENSG00000134569	LRP4	1	46878419	46940193
ENSG00000214954	LRRC69	8	92114060	92231464
ENSG00000093167	LRRFIP2	3	37094117	37225180
ENSG00000105699	LSR	1	35739233	35758867
ENSG00000119681	LTBP2	1	74964873	75079306
ENSG00000168056	LTBP3	1	65306276	65326401
ENSG00000198862	LTN1	2	30300466	30365270
ENSG00000176018	LYSMD3	5	89811428	89825401
ENSG00000183742	MACC1	7	20174278	20257027
ENSG00000172264	MACROD2	2	13976015	16033842
ENSG00000198517	MAFK	7	1570350	1582679
ENSG00000081026	MAGI3	1	113933371	114228545
ENSG00000161021	MAML1	5	179159851	179223512
ENSG00000013619	MAML1D1	X	149529689	149682448
ENSG00000078018	MAP2	2	210288782	210598842
ENSG00000107968	MAP3K8	1	30722866	30750762
ENSG00000156711	MAPK13	6	36095586	36107842
ENSG00000138834	MAPK8IP3	1	1756184	1820318
ENSG00000075413	MARK3	1	103851729	103970168
ENSG00000132561	MATN2	8	98881068	99048944
ENSG00000015479	MATR3	5	138609441	138667360
ENSG00000146701	MDH2	7	75677369	75696826
ENSG00000110492	MDK	1	46402306	46405375
ENSG00000111554	MDM1	1	68666223	68726161
ENSG00000198625	MDM4	1	204485511	204542871
ENSG00000124733	MEA1	6	42979832	42981706
ENSG00000163875	MEAF6	1	37958176	37980375
ENSG00000085276	MECOM	3	168801287	169381406
ENSG00000144893	MED12L	3	150803484	151154860
ENSG00000108510	MED13	1	60019966	60142643
ENSG00000102802	MEDAG	1	31480328	31499709
ENSG00000105976	MET	7	116312444	116438440
ENSG00000165792	METTL17	1	21457929	21465189
ENSG00000123427	METTL21B	1	58165275	58176324
ENSG00000170439	METTL7B	1	56075330	56078395
ENSG00000181588	MEX3D	1	1554668	1568057
ENSG00000140545	MFGE8	1	89441916	89456642
ENSG00000174514	MFSD4	1	205538013	205572046
ENSG00000151690	MFSD6	2	191273081	191373931
ENSG00000128268	MGAT3	2	39853349	39888199
ENSG00000161013	MGAT4B	5	179224597	179233952
ENSG00000008394	MGST1	1	16500076	16762193
ENSG00000177427	MIEF2	1	18163848	18169866
ENSG00000100253	MIOX	2	50925213	50929077
ENSG00000207939	MIR223	X	65238712	65238821
ENSG00000202566	MIR421	X	73438212	73438296
ENSG00000207652	MIR621	1	41384902	41384997
ENSG00000207997	MIR644A	2	33054130	33054223
ENSG00000167842	MIS12	1	5389605	5394134
ENSG000000196588	MKL1	2	40806285	41032706

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000130396	MLLT4	6	168227602	168372703
ENSG00000175727	MLXIP	1	122516628	122631894
ENSG00000133131	MORC4	X	106057101	106243474
ENSG00000185787	MORF4L1	1	79102829	79190475
ENSG00000060762	MPC1	6	166778407	166796486
ENSG00000197629	MPEG1	1	58975983	58980424
ENSG00000103152	MPG	1	127006	135852
ENSG00000051825	MPHOSPH9	1	123636867	123728561
ENSG00000130830	MPP1	X	154006959	154049282
ENSG00000066382	MPPED2	1	30406040	30608419
ENSG00000149573	MPZL2	1	118124118	118135251
ENSG0000011028	MRC2	1	60704762	60770958
ENSG00000173141	MRP63	1	21750784	21753223
ENSG00000180992	MRPL14	6	44081194	44095194
ENSG00000143436	MRPL9	1	151732119	151736040
ENSG00000102738	MRPS31	1	41303432	41345309
ENSG00000166928	MS4A14	1	60146003	60185161
ENSG00000052802	MSMO1	4	166248775	166264312
ENSG00000164078	MST1R	3	49924435	49941299
ENSG00000198417	MT1F	1	56691606	56694610
ENSG00000125144	MT1G	1	56700643	56701977
ENSG00000205358	MT1H	1	56703726	56705041
ENSG00000177000	MTHFR	1	11845780	11866977
ENSG00000108389	MTMR4	1	56566898	56595266
ENSG0000003987	MTMR7	8	17155539	17271037
ENSG00000120662	MTRF1	1	41790505	41837742
ENSG00000132613	MTSS1L	1	70695107	70719969
ENSG00000129422	MTUS1	8	17501304	17658426
ENSG00000185499	MUC1	1	155158300	155162707
ENSG00000204544	MUC21	6	30951495	30957680
ENSG00000162576	MXRA8	1	1288069	1297157
ENSG00000104177	MYEF2	1	48431625	48470714
ENSG00000133026	MYH10	1	8377523	8534079
ENSG00000101335	MYL9	2	35169887	35178228
ENSG00000196535	MYO18A	1	27400528	27507430
ENSG00000196586	MYO6	6	76458909	76629254
ENSG00000172766	NAA16	1	41885341	41951166
ENSG00000138386	NAB1	2	191511472	191557492
ENSG00000166886	NAB2	1	57482677	57489259
ENSG00000131400	NAPSA	1	50861734	50869087
ENSG00000185818	NAT8L	4	2061239	2070816
ENSG00000166833	NAV2	1	19372271	20143144
ENSG00000114503	NCBP2	3	196662273	196669468
ENSG00000020129	NCDN	1	36023074	36032875
ENSG00000178127	NDUFV2	1	9102628	9134343
ENSG00000188986	NELFB	9	140149625	140167998
ENSG00000184613	NELL2	1	44902058	45315631
ENSG00000173848	NET1	1	5454514	5500426
ENSG00000050344	NFE2L3	7	26191860	26226745
ENSG00000147862	NFIB	9	14081842	14398982
ENSG00000066248	NGEF	2	233743396	233877982
ENSG00000064300	NGFR	1	47572655	47592379
ENSG00000145912	NHP2	5	177576461	177580968
ENSG00000001461	NIPAL3	1	24742284	24799466
ENSG00000101882	NKAP	X	119059014	119077735
ENSG00000169992	NLGN2	1	7308193	7323179
ENSG00000169251	NMD3	3	160822484	160971320
ENSG00000106100	NOD1	7	30464143	30518400
ENSG00000225921	NOL7	6	13615559	13632971
ENSG00000147140	NONO	X	70503042	70521018
ENSG00000198929	NOS1AP	1	162039564	162353321
ENSG00000213240	NOTCH2NL	1	145209119	145291972
ENSG00000074181	NOTCH3	1	15270444	15311792
ENSG00000139910	NOVA1	1	26912299	27066960
ENSG00000086991	NOX4	1	89057524	89322779
ENSG00000119655	NPC2	1	74942895	74960880
ENSG00000107281	NPDC1	9	139933922	139940655
ENSG00000185864	NPIPBA	1	21845890	21892148
ENSG00000221890	NPTXR	2	39214457	39239987
ENSG00000091129	NRCAM	7	107788068	108097161
ENSG00000180530	NRIP1	2	16333556	16437321
ENSG00000241058	NSUN6	1	18834490	18940551
ENSG00000168268	NT5DC2	3	52558386	52569070

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000135318	NTSE	6	86159809	86205500
ENSG00000140538	NTRK3	1	88418230	88799999
ENSG00000198585	NUDT16	3	131100515	131107674
ENSG00000186364	NUDT17	1	145586115	145589439
ENSG00000069248	NUP133	1	229577045	229644103
ENSG00000176046	NUPR1	1	28548606	28550495
ENSG00000167693	NXN	1	702553	883010
ENSG00000145247	OCLAD2	4	48887036	48908954
ENSG00000197822	OCLN	5	68788119	68853931
ENSG00000145623	OSMR	5	38845960	38945698
ENSG00000155100	OTUD6B	8	92082424	92099323
ENSG00000162881	OXER1	2	42989642	42991401
ENSG00000154814	OXNAD1	3	16306706	16391806
ENSG00000078589	P2RY10	X	78200829	78217451
ENSG00000181631	P2RY13	3	151044100	151047336
ENSG00000079462	PAFAH1B3	1	42801185	42807698
ENSG00000099864	PALM	1	708953	748329
ENSG00000145730	PAM	5	102089685	102366809
ENSG00000138964	PARVG	2	44568836	44615413
ENSG00000115687	PASK	2	242045514	242089679
ENSG00000229474	PATL2	1	44957930	45003514
ENSG00000173599	PC	1	66615704	66725847
ENSG00000156453	PCDH1	5	141232938	141258811
ENSG00000189184	PCDH18	4	138440072	138453648
ENSG00000243232	PCDHAC2	5	140345820	140391936
ENSG00000240184	PCDHGC3	5	140855580	140892542
ENSG00000102109	PCSK1N	X	48689504	48694035
ENSG00000154678	PDE1C	7	31790793	32338941
ENSG00000138735	PDE5A	4	120415550	120550146
ENSG00000073417	PDE8A	1	85523671	85682376
ENSG00000160191	PDE9A	2	44073746	44195619
ENSG00000131828	PDHAI	X	19362011	19379823
ENSG00000107438	PDLIM1	1	96997329	97050781
ENSG00000131435	PDLIM4	5	131593364	131609147
ENSG00000162734	PEA15	1	160175127	160185166
ENSG00000133027	PEMT	1	17408877	17495022
ENSG00000112378	PERP	6	138409642	138428648
ENSG00000143256	PFDN2	1	161070346	161087901
ENSG00000158571	PKFB1	X	54959394	55024967
ENSG00000123836	PKFB2	1	207222801	207254369
ENSG00000164219	PGGT1B	5	114546527	114598569
ENSG00000101856	PGRMC1	X	118370216	118378429
ENSG00000116273	PHF13	1	6673745	6684093
ENSG00000116793	PHTF1	1	114239453	114302111
ENSG00000107537	PHYH	1	13319796	13344412
ENSG00000168490	PHYHIP	8	22077222	22089854
ENSG00000175309	PHYKPL	5	177635498	177659792
ENSG00000131788	PIAS3	1	145575233	145586546
ENSG00000105229	PIAS4	1	4007644	4039384
ENSG00000197563	PIGN	1	59710800	59854351
ENSG00000141506	PIK3R5	1	8782233	8869029
ENSG00000102096	PIM2	X	48770459	48776301
ENSG00000254093	PINX1	8	10622473	10697394
ENSG00000241878	PISD	2	32014477	32058418
ENSG00000205038	PKHD1L1	8	110374706	110542559
ENSG00000057294	PKP2	1	32943679	33049774
ENSG00000144283	PKP4	2	159313476	159539391
ENSG00000176485	PLA2G16	1	63340667	63384355
ENSG00000181690	PLAG1	8	57073463	57123883
ENSG00000182621	PLCB1	2	8112824	8949003
ENSG00000161714	PLCD3	1	43186335	43210721
ENSG00000115896	PLCL1	2	198669426	199437305
ENSG00000115956	PLEK	2	68592305	68624585
ENSG00000105559	PLEKHA4	1	49340354	49371889
ENSG00000052126	PLEKHA5	1	19282648	19529334
ENSG00000143850	PLEKHA6	1	204187979	204346793
ENSG00000187583	PLEKH1N	1	901877	911245
ENSG00000145632	PLK2	5	57749809	57756087
ENSG00000171566	PLRG1	4	155456158	155471587
ENSG00000120756	PLS1	3	142315229	142432506
ENSG00000102024	PLS3	X	114795501	114885181
ENSG00000130827	PLXNA3	X	153686621	153701989
ENSG00000196576	PLXNB2	2	50713408	50746056

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000176903	PNMA1	1	74178494	74181128
ENSG00000146278	PNRC1	6	89790470	89794879
ENSG00000102978	POLR2C	1	57496299	57505922
ENSG00000185900	POMK	8	42948658	42978577
ENSG00000105854	PON2	7	95034175	95064510
ENSG00000137709	POU2F3	1	120107349	120190653
ENSG00000180817	PPA1	1	71962586	71993667
ENSG00000141934	PPAP2C	1	281040	291393
ENSG00000171497	PPID	4	159630286	159644548
ENSG00000145725	PPIP5K2	5	102455853	102548500
ENSG00000118898	PPL	1	4932508	5010742
ENSG00000100034	PPM1F	2	22273793	22307209
ENSG00000077157	PPP1R12B	1	202317827	202561834
ENSG00000115685	PPP1R7	2	242088991	242123067
ENSG00000105568	PPP2R1A	1	52693292	52730687
ENSG00000156475	PPP2R2B	5	145967936	146464347
ENSG00000114485	PPP5C	1	46850251	46896238
ENSG00000196850	PPTC7	1	110969120	111021125
ENSG00000139174	PRICKLE1	1	42852140	42984157
ENSG00000106617	PRKAG2	7	151253197	151574210
ENSG00000154229	PRKCA	1	64298754	64806861
ENSG00000065675	PRKCQ	1	6469105	6622263
ENSG00000185532	PRKG1	1	52750945	54058110
ENSG00000126457	PRMT1	1	50179043	50192286
ENSG00000171867	PRNP	2	4666882	4682236
ENSG00000184500	PROS1	3	93591881	93692910
ENSG00000112739	PRPF4B	6	4021501	4065217
ENSG00000205352	PRR13	1	53835389	53840429
ENSG00000183530	PRR14L	2	32072242	32146126
ENSG00000176532	PRR15	7	29603427	29606911
ENSG00000204469	PRRC2A	6	31588497	31605548
ENSG00000005001	PRSS22	1	2902728	2908171
ENSG00000150687	PRSS23	1	86502101	86663952
ENSG00000105227	PRX	1	40899675	40919273
ENSG00000156011	PSD3	8	18384811	18942240
ENSG00000112655	PTK7	6	43044006	43129457
ENSG00000188921	PTPLAD2	9	20995306	21031635
ENSG00000088179	PTPN4	2	120517207	120741394
ENSG00000081237	PTPRC	1	198607801	198726545
ENSG00000132334	PTPRE	1	129705325	129884119
ENSG00000142949	PTPRF	1	43990858	44089343
ENSG00000144724	PTPRG	3	61547243	62283288
ENSG00000152894	PTPRK	6	128289924	128841870
ENSG00000139304	PTPRQ	1	80799774	81072802
ENSG00000060656	PTPRU	1	29563028	29653325
ENSG00000177469	PTRF	1	40554470	40575535
ENSG00000091127	PUS7	7	105080108	105162714
ENSG00000100362	PVALB	2	37196728	37215523
ENSG00000143217	PVRL4	1	161040785	161059389
ENSG00000100504	PYGL	1	51324609	51411454
ENSG00000163564	PYHIN1	1	158900586	158946844
ENSG00000126838	PZP	1	9301436	9360966
ENSG00000157869	RAB28	4	13362978	13485989
ENSG00000109113	RAB34	1	27041299	27045447
ENSG00000119318	RAD23B	9	110045418	110094475
ENSG00000203722	RAET1G	6	150238014	150244257
ENSG00000175097	RAG2	1	36597124	36619829
ENSG00000131831	RAI2	X	17818169	17879457
ENSG00000158987	RAPGEF6	5	130759614	130970929
ENSG00000165917	RAPSIN	1	47459308	47470730
ENSG00000172819	RARG	1	53604354	53626764
ENSG00000145715	RASA1	5	86563705	86687748
ENSG00000100302	RASD2	2	35936915	35950048
ENSG00000068028	RASSF1	3	50367219	50378411
ENSG00000146587	RBAK	7	5085452	5109119
ENSG00000102054	RBBP7	X	16857406	16888537
ENSG00000127993	RBM48	7	92158087	92167319
ENSG0000003756	RBM5	3	50126341	50156454
ENSG00000076067	RBMS2	1	56915713	56984745
ENSG00000117906	RCN2	1	77223960	77242601
ENSG00000079313	REXO1	1	1815248	1848452
ENSG00000127074	RGS13	1	192605275	192629390
ENSG00000155366	RHOC	1	113243728	113250056

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000116574	RHOU	1	228870824	228882416
ENSG00000176406	RIMS2	8	104512976	105268322
ENSG00000170881	RNF139	8	125486979	125500155
ENSG00000141576	RNF157	1	74138534	74236454
ENSG00000101236	RNF24	2	3907956	3996229
ENSG00000149489	ROM1	1	62379194	62382592
ENSG00000221817	RP11-137L10.6	1	75255283	75279828
ENSG00000271141	RP11-17112.4	2	179481308	179481850
ENSG00000132383	RPA1	1	1732996	1803376
ENSG00000156313	RPGR	X	38128416	38186817
ENSG00000198755	RPL10A	6	35436185	35438562
ENSG00000174748	RPL15	3	23958036	23965183
ENSG00000114391	RPL24	3	101399935	101405626
ENSG00000122406	RPL5	1	93297582	93307481
ENSG00000148303	RPL7A	9	136215069	136218281
ENSG00000141425	RPRD1A	1	33564350	33647539
ENSG00000163125	RPRD2	1	150335567	150449042
ENSG00000100784	RPS6KA5	1	91336799	91526980
ENSG00000170889	RPS9	1	54704610	54752862
ENSG00000155876	RRAGA	9	19049372	19051019
ENSG0000025039	RRAGD	6	90074355	90121989
ENSG00000126458	RRAS	1	50138549	50143458
ENSG00000048392	RRM2B	8	103216730	103251346
ENSG00000101282	RSPO4	2	939095	982907
ENSG00000143171	RXRG	1	165370159	165414433
ENSG00000188643	S100A16	1	153579362	153585621
ENSG00000197956	S100A6	1	153507075	153508720
ENSG00000109929	SCSD	1	121163162	121179403
ENSG00000139218	SCAF11	1	46312914	46385903
ENSG00000168077	SCARA3	8	27491385	27534293
ENSG00000136155	SCEL	1	78109809	78219398
ENSG00000166922	SCG5	1	32933877	32989299
ENSG00000146285	SCML4	6	108025308	108145521
ENSG00000159307	SCUBE1	2	43593289	43739394
ENSG00000146197	SCUBE3	6	35182190	35220856
ENSG00000124145	SDC4	2	43953928	43977064
ENSG00000073578	SDHA	5	218356	256815
ENSG00000146555	SDK1	7	3341080	4308632
ENSG00000100445	SDR39U1	1	24908972	24912111
ENSG00000075826	SEC31B	1	102246399	102289628
ENSG00000085415	SEH1L	1	12947132	12987535
ENSG00000186838	SELV	1	40005753	40011326
ENSG00000153993	SEMA3D	7	84624869	84816171
ENSG00000001617	SEMA3F	3	50192478	50226508
ENSG00000138468	SENP7	3	101043049	101232085
ENSG00000183291	SEP15	1	87328132	87380107
ENSG00000109618	SEPSECS	4	25121627	25162204
ENSG00000168385	SEPT2	2	242254515	242293442
ENSG00000178980	SEPW1	1	48281829	48287943
ENSG00000129158	SERGEF	1	17809595	18034709
ENSG00000197249	SERPINA1	1	94843084	94857030
ENSG00000197019	SERTAD1	1	40927499	40931932
ENSG00000139718	SETD1B	1	122242086	122270562
ENSG00000168066	SF1	1	64532078	64546258
ENSG000000115128	SF3B14	2	24290454	24299313
ENSG00000087365	SF3B2	1	65818200	65836779
ENSG00000189091	SF3B3	1	70557691	70608820
ENSG00000061936	SFWAP	1	132195626	132284282
ENSG00000163069	SGCB	4	52886872	52904648
ENSG00000127990	SGCE	7	94214542	94285521
ENSG00000164023	SGMS2	4	108745719	108836203
ENSG00000104611	SH2D4A	8	19171128	19253729
ENSG00000160691	SHC1	1	154934774	154946871
ENSG00000169291	SHE	1	154442248	154474589
ENSG00000138606	SHF	1	45459412	45493373
ENSG00000158352	SHROOM4	X	50334647	50557302
ENSG00000181788	SIAH2	3	150458914	150481264
ENSG00000147955	SIGMAR1	9	34634719	34637806
ENSG00000162739	SLAMF6	1	160454820	160493052
ENSG00000120519	SLC10A7	4	147175127	147443123
ENSG00000064651	SLC12A2	5	127419458	127525380
ENSG000000155380	SLC16A1	1	113454469	113499635
ENSG000000168679	SLC16A4	1	110905470	110933704

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000119899	SLC17A5	6	74303102	74363878
ENSG00000259803	SLC22A31	1	89262406	89268072
ENSG00000102743	SLC25A15	1	41363548	41384247
ENSG00000155287	SLC25A28	1	101370282	101380366
ENSG00000125434	SLC25A35	1	8191081	8198661
ENSG00000140284	SLC27A2	1	50474393	50528592
ENSG00000113396	SLC27A6	5	127873706	128369335
ENSG00000160326	SLC2A6	9	136336217	136344259
ENSG00000152683	SLC30A6	2	32390933	32449448
ENSG00000136868	SLC31A1	9	115983808	116028674
ENSG00000136867	SLC31A2	9	115913222	115926417
ENSG00000157765	SLC34A2	4	25656923	25680370
ENSG00000121073	SLC35B1	1	47778305	47786376
ENSG00000110660	SLC35F2	1	107661717	107799019
ENSG00000183780	SLC35F3	1	234040679	234460262
ENSG00000141424	SLC39A6	1	33688495	33709348
ENSG00000134802	SLC43A3	1	57174427	57195053
ENSG00000004939	SLC4A1	1	42325753	42345509
ENSG00000080493	SLC4A4	4	72053003	72437804
ENSG00000169241	SLC50A1	1	155107820	155111329
ENSG00000140675	SLC5A2	1	31494323	31502181
ENSG00000103064	SLC7A6	1	68298433	68335722
ENSG00000145147	SLT2	4	20254883	20622184
ENSG00000163681	SLMAP	3	57741177	57914895
ENSG00000124107	SLPI	2	43880880	43883205
ENSG00000137776	SLTM	1	59171244	59225852
ENSG00000157106	SMG1	1	18816175	18937776
ENSG00000163683	SMIM14	4	39547950	39640710
ENSG00000130768	SMPDL3B	1	28261504	28285668
ENSG00000122692	SMU1	9	33041762	33076665
ENSG00000145335	SNCA	4	90645250	90759466
ENSG00000173267	SNCG	1	88718375	88723017
ENSG00000212443	SNORA53	1	98993413	98993661
ENSG00000163788	SNRK	3	43328004	43466256
ENSG00000028528	SNX1	1	64386322	64438289
ENSG0000002919	SNX11	1	46180719	46200436
ENSG00000147164	SNX12	X	70279094	70288273
ENSG00000167208	SNX20	1	50700211	50715264
ENSG00000157734	SNX22	1	64443914	64449680
ENSG00000109762	SNX25	4	186125391	186291339
ENSG00000173548	SNX33	1	75940247	75954642
ENSG00000089006	SNX5	2	17922241	17949623
ENSG00000198944	SOWAHA	5	132149033	132152488
ENSG00000124776	SOX4	6	21593972	21598847
ENSG00000172845	SP3	2	174771187	174830430
ENSG00000196141	SPATS2L	2	201170604	201346986
ENSG00000166145	SPINT1	1	41136216	41150405
ENSG00000198369	SPRED2	2	65537985	65659771
ENSG00000164056	SPRY1	4	124317950	124324910
ENSG00000187678	SPRY4	5	141689992	141706020
ENSG00000197694	SPTAN1	9	131314866	131395941
ENSG00000090054	SPTLC1	9	94794281	94877666
ENSG00000075142	SRI	7	87834433	87856308
ENSG00000167881	SRP68	1	74035184	74068734
ENSG00000135250	SRPK2	7	104751151	105039755
ENSG00000116350	SRSF4	1	29474255	29508499
ENSG00000145687	SSBP2	5	80708840	81047616
ENSG00000149136	SSRP1	1	57093459	57103351
ENSG00000160075	SSU72	1	1477053	1510249
ENSG00000157350	ST3GAL2	1	70413338	70473140
ENSG00000115525	ST3GAL5	2	86066267	86116137
ENSG00000167323	STIM1	1	3875757	4114439
ENSG00000169302	STK32A	5	146614526	146767415
ENSG00000165283	STOML2	9	35099888	35103154
ENSG00000137868	STRA6	1	74471807	74504608
ENSG00000104915	STX10	1	13254872	13261197
ENSG00000124222	STX16	2	57226328	57254582
ENSG00000111450	STX2	1	131274145	131323811
ENSG00000177688	SUMO4	6	149721495	149722177
ENSG00000102710	SUPT20H	1	37583449	37633850
ENSG00000196235	SUPT5H	1	39926796	39967310
ENSG00000148291	SURF2	9	136223428	136228045
ENSG00000099994	SUSD2	2	24577227	24585078

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000159164	SV2A	1	149874870	149889434
ENSG00000173928	SWSAP1	1	11485361	11487627
ENSG00000171992	SYNPO	5	149980642	150038782
ENSG00000061114	SYNRG	1	35874900	35969544
ENSG00000147041	SYTL5	X	37865835	37988072
ENSG00000184292	TACSTD2	1	59041099	59043166
ENSG00000064995	TAF11	6	34845555	34855866
ENSG00000103168	TAF1C	1	84211458	84220669
ENSG00000165632	TAF3	1	7860467	8058590
ENSG00000144559	TAMM41	3	11831916	11888393
ENSG00000183597	TANGO2	2	20004537	20053449
ENSG00000113838	TBCCD1	3	186263862	186288332
ENSG00000176896	TCEANC	X	13671225	13700083
ENSG00000116205	TCEANC2	1	54519260	54578192
ENSG00000139437	TCHP	1	110338069	110421646
ENSG00000182134	TDRKH	1	151742583	151763892
ENSG00000205356	TECPRI	7	97843936	97881563
ENSG0000009694	TENM1	X	123509753	124097666
ENSG00000115112	TFCP2L1	2	121974163	122042783
ENSG00000163235	TGFA	2	70674412	70781325
ENSG00000140682	TGFB1I1	1	31482906	31489281
ENSG00000092969	TGFB2	1	218519577	218617961
ENSG00000092295	TGM1	1	24718320	24733638
ENSG00000169231	THBS3	1	155165379	155178842
ENSG00000151365	THRSP	1	77774907	77779397
ENSG00000102265	TIMP1	X	47441712	47446188
ENSG00000035862	TIMP2	1	76849059	76921469
ENSG00000163659	TIPARP	3	156391024	156424559
ENSG00000119139	TJP2	9	71736209	71870124
ENSG00000169908	TM4SF1	3	149086809	149095652
ENSG00000169903	TM4SF4	3	149191761	149221068
ENSG00000144868	TMEM108	3	132757235	133116636
ENSG00000116368	TMEM159	1	21169698	21191937
ENSG00000164180	TMEM161B	5	87485450	87565293
ENSG00000152128	TMEM163	2	135213330	135476570
ENSG00000157600	TMEM164	X	109245859	109425962
ENSG00000187713	TMEM203	9	140098534	140100090
ENSG00000131634	TMEM204	1	1578689	1605581
ENSG00000186501	TMEM222	1	27648651	27662891
ENSG00000106609	TMEM248	7	66386212	66423538
ENSG00000112697	TMEM30A	6	75962640	75994684
ENSG00000163900	TMEM41A	3	185194284	185216845
ENSG00000145014	TMEM44	3	194308402	194354418
ENSG00000180694	TMEM64	8	91634223	91803860
ENSG00000163472	TMEM79	1	156252726	156262976
ENSG00000103978	TMEM87A	1	42502730	42565861
ENSG00000153214	TMEM87B	2	112812800	112876895
ENSG0000006042	TMEM98	1	31254928	31272124
ENSG00000137648	TMPRSS4	1	117947753	117992605
ENSG00000187045	TMPRSS6	2	37461476	37505603
ENSG00000034510	TMSB10	2	85132749	85133795
ENSG00000041982	TNC	9	117782806	117880536
ENSG0000006327	TNFRSF12A	1	3068446	3072384
ENSG00000048462	TNFRSF17	1	12058964	12061925
ENSG00000067182	TNFRSF1A	1	6437923	6451280
ENSG00000173273	TNKS	8	9413424	9639856
ENSG00000183864	TOB2	2	41829496	41843027
ENSG00000132773	TOE1	1	45805342	45809647
ENSG00000173726	TOMM20	1	235272651	235292251
ENSG00000177302	TOP3A	1	18174742	18218321
ENSG00000169905	TOR1AIP2	1	179809102	179846938
ENSG00000160404	TOR2A	9	130493803	130497604
ENSG00000143514	TP53BP2	1	223967601	224033674
ENSG00000170638	TRABD	2	50624344	50638027
ENSG00000056972	TRAF3IP2	6	111877657	111927481
ENSG00000175104	TRAF6	1	36508577	36531822
ENSG00000160218	TRAPP10	2	45432200	45526433
ENSG00000171853	TRAPP12	2	3383446	3488865
ENSG00000196655	TRAPP4	1	118889142	118896164
ENSG00000204599	TRIM39	6	30294256	30311506
ENSG00000183718	TRIM52	5	180681417	180688119
ENSG00000166436	TRIM66	1	8633584	8693413
ENSG00000173113	TRMT112	1	64083932	64085556

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000072315	TRPC5	X	111017543	111326004
ENSG00000102804	TSC22D1	1	45007655	45151283
ENSG00000157514	TSC22D3	X	106956451	107020572
ENSG00000179981	TSHZ1	1	72922710	73001905
ENSG00000187189	TSPYL4	6	116571151	116575261
ENSG00000182670	TTC3	2	38445526	38575413
ENSG00000214021	TTLL3	3	9849770	9896822
ENSG00000188229	TUBB4B	9	140135665	140138159
ENSG00000104723	TUSC3	8	15274724	15624158
ENSG0000017862	TXND12	1	52485803	52521843
ENSG0000092445	TYRO3	1	41849873	41871536
ENSG00000117143	UAP1	1	162531323	162569627
ENSG00000184787	UBE2G2	2	46188955	46221934
ENSG00000103275	UBE2I	1	1355548	1377019
ENSG00000215218	UBE2QL1	5	6448736	6495022
ENSG00000162543	UBXN10	1	20512578	20522541
ENSG00000158062	UBXN11	1	26607819	26644854
ENSG00000116750	UCHL5	1	192981380	193029237
ENSG00000143222	UFC1	1	161122566	161128646
ENSG00000109814	UGDH	4	39500375	39529931
ENSG00000131015	ULBP2	6	150263136	150270371
ENSG00000177169	ULK1	1	132379196	132407712
ENSG00000151461	UPF2	1	11962021	12085169
ENSG00000125351	UPF3B	X	118967985	118986961
ENSG00000077254	USP33	1	78161672	78225537
ENSG00000132952	USPL1	1	31191830	31233686
ENSG00000156697	UTP14A	X	129040097	129063737
ENSG00000163945	UVSSA	4	1341054	1381837
ENSG00000168140	VASN	1	4421849	4433529
ENSG00000100483	VCPKMT	1	50575350	50583318
ENSG00000187650	VMAC	1	5904869	5910864
ENSG00000139722	VPS37B	1	123349882	123380991
ENSG00000156931	VPS8	3	184529931	184770402
ENSG00000165633	VSTM4	1	50222290	50323554
ENSG00000151532	VTI1A	1	114206756	114578503
ENSG00000179403	VWA1	1	1370241	1378262
ENSG00000110002	VWA5A	1	123986069	124018428
ENSG00000204396	VWA7	6	31733367	31745108
ENSG00000015285	WAS	X	48534985	48549818
ENSG00000196998	WDR45	X	48929385	48958108
ENSG00000070540	WIP1I	1	66417089	66453654
ENSG00000142279	WTIP	1	34971874	34997258
ENSG00000182489	XKRX	X	100168431	100184422
ENSG00000143324	XPR1	1	180601140	180859387
ENSG00000079246	XRCC5	2	216972187	217071026
ENSG00000177494	ZBED2	3	111311747	111314290
ENSG00000126804	ZBTB1	1	64970430	65000408
ENSG00000205189	ZBTB10	8	81397854	81438500
ENSG00000177485	ZBTB33	X	119384607	119392253
ENSG00000168826	ZBTB49	4	4291924	4323513
ENSG00000104427	ZC2H1CA	8	79578282	79632000
ENSG00000122299	ZC3H7A	1	11844442	11891123
ENSG00000144161	ZC3H8	2	112969102	113012713
ENSG00000174460	ZCCHC12	X	117957753	117960931
ENSG00000186908	ZDHHC17	1	77157368	77247476
ENSG00000156599	ZDHHC5	1	57435219	57468659
ENSG00000153786	ZDHHC7	1	85007787	85045141
ENSG00000133858	ZFC3H1	1	72003252	72061505
ENSG00000152518	ZFP36L2	2	43449541	43453748
ENSG00000039319	ZFYVE16	5	79703832	79775169
ENSG00000172667	ZMAT3	3	178735011	178790067
ENSG00000165061	ZMAT4	8	40388109	40753532
ENSG00000163867	ZMYM6	1	35449523	35497569
ENSG00000172262	ZNF131	5	43065278	43192123
ENSG00000256294	ZNF225	1	44616334	44637027
ENSG00000159917	ZNF235	1	44732882	44809199
ENSG00000158805	ZNF276	1	89786808	89807311
ENSG00000160961	ZNF333	1	14800613	14844558
ENSG00000130684	ZNF337	2	25654851	25677477
ENSG00000189180	ZNF33A	1	38299578	38354016
ENSG00000113761	ZNF346	5	176449697	176508190
ENSG00000256683	ZNF350	1	52467596	52490109
ENSG00000197024	ZNF398	7	1488803508	148880116

TABLE 3-continued

List of 1115 core genes deriving the ensemble model prediction.

Gene_id	Gene_name	Chromosome	Start	End
ENSG00000215421	ZNF407	1	72265106	72777627
ENSG00000133250	ZNF414	1	8575462	8579048
ENSG00000173480	ZNF417	1	58411664	58427978
ENSG00000183621	ZNF438	1	31109136	31320866
ENSG00000185219	ZNF445	3	44481262	44519162
ENSG00000197016	ZNF470	1	57078880	57100279
ENSG00000101493	ZNF516	1	74069644	74207146
ENSG0000074657	ZNF532	1	56529832	56653712
ENSG00000258405	ZNF578	1	52956829	53015407
ENSG00000198466	ZNF587	1	58361225	58376480
ENSG00000197343	ZNF655	7	99156029	99174076
ENSG00000196757	ZNF700	1	12035883	12061588
ENSG00000181135	ZNF707	8	144766622	144796068
ENSG00000196456	ZNF775	7	150065879	150109558
ENSG00000198556	ZNF789	7	99070464	99101273
ENSG00000204524	ZNF805	1	57751973	57766503
ENSG00000178917	ZNF852	3	44540462	44552128
ENSG00000106479	ZNF862	7	149535456	149564568
ENSG0000070476	ZXDC	3	126156444	126194762
ENSG0000074755	ZZEF1	1	3907739	4046314

## Example 9. Statistical Analysis

[0145] Statistical analyses were performed using R statistical software version 3.2.3. Continuous variables were compared using t test, and categorical variables were compared using Fisher exact test. Test performance was evaluated using sensitivity, specificity, and NPV and PPV based on established methods. All confidence intervals are 2-sided 95% CIs and were computed using the exact binomial test. Test performance comparison between the GSC and GEC was done using McNemar  $\chi^2$  test on the matched data set. Significance level in differential gene expression analysis is reported using a false discovery rate-adjusted P value. Two-sided P values less than 0.05 were used to declare significance.

## RESULTS

[0146] FNA samples that previously validated the GEC were used to independently validate the GSC. The earlier

GEC validation samples were derived from 4812 nodule aspirations prospectively collected from 3789 patients at 49 clinical sites in the United States over a 2-year period. Of the 210 validation samples with corresponding Bethesda III or IV cytology and blinded postoperative consensus histopathology diagnoses, 191 (91.0%) had sufficient residual RNA for GSC testing. These samples from cytologically indeterminate nodules constituted the blinded primary test set.

[0147] The previously established thyroid nodule cytological diagnosis was used again. Patient demographic characteristics and baseline data are shown in Table 4. Age, sex, clinical risk factors, nodule size, histology subtype (Table 5), number of FNA passes, prevalence of malignancy (Table 6), and proportion of samples collected at community centers did not differ significantly between the primary study population (n=191) and the GEC clinical validation cohort of samples (n=210), consistent with unbiased drop out.

TABLE 4

Baseline demographic and clinical characteristics of the study cohort <sup>a</sup> .		
Variable	GEC Validation	GSC Validation
Total, No.		
Samples	210	191
Patients	199	183
Type of study site, No. (%) of samples		
Academic	76 (36.2)	65 (34.0)
Community	134 (63.8)	126 (66.0)
No. of fine-needle aspiration passes, No. (%) of samples		
1	88 (41.9)	73 (38.2)
2	122 (58.1)	118 (61.8)
Age of patients, mean (range), y	51.2 (22.0-85.0)	51.7 (22.0-85.0)
Male	46 (23.1)	41 (22.4)
Female	153 (76.9)	142 (77.6)
Risk factors, No. (%) of patients		
Radiation exposure to head, neck, or both	7 (3.5)	5 (2.7)
Family history of thyroid cancer	14 (7.0)	13 (7.1)

TABLE 4-continued

Baseline demographic and clinical characteristics of the study cohort <sup>a</sup> .		
Variable	GEC Validation	GSC Validation
<b>Nodule</b>		
Size of ultrasonography, median (range), cm	2.5 (1.0-9.1)	2.6 (1.0-9.1)
Size group, No. (%) of nodules, cm		
1.00-1.99	69 (32.9)	60 (31.4)
2.00-2.99	62 (29.5)	60 (31.4)
3.00-3.99	42 (20.0)	37 (19.4)
≥4.00	37 (17.6)	34 (17.8)

Abbreviations:

GEC, gene expression classifier;

GSC, genomic sequencing classifier

<sup>a</sup>Statistical tests were performed to compare the 19 nodules in the GEC validation that were excluded in the GSC validation because of insufficient RNA quantity. The 2 groups differ only on the number of fine-needle aspiration passes, which is not unexpected, as only samples with sufficient remaining RNA were included in the GSC evaluation.

TABLE 5

Histology subtype comparison between validation cohorts.			
Histology Subtype Group	GEC (N = 210)	GSC (N = 191)	P-value
BNF, HN	63	54	0.47
FA	56	54	
FT-UMP, WDT-UMP	18	17	
HCA	19	17	
CLT, HT	2	2	
HTA	1	1	
PTC, PTC-TCV	18	17	
FVPTC	12	11	
HCC-c, HCC-v	9	19	
FC-c, FC-v, WDC-NOS	9	7	
PDC, ML, MTC	3	2	

[0148] P-value is from a test comparing the 191 GSC nodules with the 19 nodules in the GEC validation that were excluded in the GSC validation due to insufficient RNA quantity. Histology subtype abbreviations: BFN-benign follicular nodule, HN-hyperplastic nodule, FA follicular adenoma, FT-UMP-follicular tumor of uncertain malignant potential, WDT-UMP well differentiated tumor of uncertain malignant potential, HCA-Hürthle cell adenoma, CLT chronic lymphocytic thyroiditis, HT-Hashimoto's thyroiditis, HTA-hyalinizing trabecular adenoma, PTC-papillary thyroid cancer, PTC-TCV-papillary thyroid cancer tall cell variant, FVPTC-papillary thyroid cancer follicular variant, HCC-c-Hürthle cell carcinoma capsular invasion, HCC-v-Hürthle cell carcinoma vascular invasion, FC-c-follicular carcinoma capsular invasion, FC-v-follicular carcinoma vascular invasion, WDC-NOS-well differentiated carcinoma not otherwise specified, PDC-poorly differentiated carcinoma, ML malignant lymphoma, MTC-medullary thyroid cancer

TABLE 6

Prevalence of malignancy between validation cohorts.			
Histologic Label	GEC (N = 210)	GSC (N = 191)	P-value
Benign	159	145	1.00
Malignant	51	46	
Cancer prevalence	24.3%	24.1%	

P-value is from a test comparing the 191 GSC nodules with the 19 nodules in the GEC validation that were excluded in the GSC validation due to insufficient RNA quantity.

[0149] The Standards for Reporting of Diagnostic Accuracy Studies was developed to improve the quality of reporting diagnostic accuracy studies. FIG. 2 shows the flow of samples through the study in a Standards for Reporting of Diagnostic Accuracy Studies diagram. Of these 191 indeterminate FNAs, 46 (24.1%) were diagnosed as malignant by an expert surgical histopathology panel who were blinded to all cytologic and genomic results and to the local histopathology diagnosis. Results are reported in the order of testing through the GSC test system (FIG. 1). Initially, all GSC samples are tested for RNA quantity and quality. None of the 191 samples failed. Subsequently, the GSC aimed to identify nodules composed of parathyroid tissue, those with MTC, and those with a BRAF V600E mutation or RET/PTC1 or RET/PTC3 fusion. Samples testing positive for these are included in performance calculations described below, except for samples testing positive for parathyroid tissue, as this result does not indicate a benign or malignant etiology. Among the 191 samples, positive results for parathyroid, MTC, BRAF, and RET/PTC occurred in 0, 1, 3, and 0 samples, respectively. All MTC and BRAF V600E results were concordant with reference methods. After this testing, samples were evaluated for follicular cell content by the follicular content index classifier. One sample, negative for the above results, was deemed to have inadequate follicular content and therefore was assigned no result. This sample was excluded from subsequent analyses, leaving 190 samples. Table 7 summarizes clinical performance characteristics for Bethesda III and IV nodules.

TABLE 7

GSC Result	Reference Standard, % (95% CI)	
	Malignant	Benign
<u>Performance across the primary test set of Bethesda III and IV indeterminate nodules (n = 190)</u>		
Suspicious, No./total No.	41/45	46/145
Benign, No./total No.	4/45	99/145
Sensitivity	91.1 (79-98)	
Specificity	68.3 (60-76)	
NPV	96.1 (90-99)	
PPV	47.1 (36-58)	
Prevalence of malignant lesions, %	23.7	
Bethesda III: atypia of undermined significance/follicular lesion of undetermined significance (n = 114 [60.0%])		
Suspicious, No./total No.	26/28	25/86
Benign, No./total No.	2/28	61/86
Sensitivity	92.9 (76-99)	
Specificity	70.9 (60-80)	
NPV	96.8 (89-100)	
PPV	51.0 (37-65)	
Prevalence of malignant lesions, %	24.6	
Bethesda IV: follicular or Hürthle cell neoplasm or suspicious for follicular neoplasm (n = 76 [40.0%])		
Suspicious, No./total No.	15/17	21/59
Benign, No./total No.	2/17	38/59
Sensitivity	88.2 (64-99)	
Specificity	64.4 (51-76)	
NPV	95.0 (83-99)	
PPV	41.7 (26-59)	
Prevalence of malignant lesions, %	22.4	
Performance across the secondary test set of Bethesda II, V, and VI nodules (n = 61) <sup>a</sup>		
Suspicious, No./total No.	34/34	7/26
Benign, No./total No.	0/34	19/26
Sensitivity	100 (90-100)	
Specificity	73.1 (52-88)	
NPV	100 (82-100)	
PPV	82.9 (68-93)	
Prevalence of malignant lesions, %	56.7	
Bethesda II: cytopathologically benign (n = 19 [31.1%]) <sup>a</sup>		
Suspicious, No./total No.	2.2	2/16
Benign, No./total No.	2/0	14/16
Sensitivity	100 (16-100)	
Specificity	87.5 (62-98)	
NPV	100 (77-100)	
PPV	50.0 (7-93)	
Prevalence of malignant lesions, %	11.1	
Bethesda V: suspicious for malignancy (n = 23 [37.7%])		
Suspicious, No./total No.	13/13	5/10
Benign, No./total No.	0/13	5/10
Sensitivity	100 (75-100)	
Specificity	50.0 (19-81)	
NPV	100 (48-100)	
PPV	72.2 (47-90)	
Prevalence of malignant lesions, %	56.5	
Bethesda VI: cytopathologically malignant (n = 19 [31.1%])		
Suspicious, No./total No.	19/19	0/0
Benign, No./total No.	0/19	0/0

TABLE 7-continued

Performance of the Genomic Sequencing Classifier (GSC) According to the Final Histopathological Diagnoses and Cytopathological Category.

GSC Result	Reference Standard, % (95% CI)	
	Malignant	Benign
Sensitivity	100 (82-100)	
PPV	100 (82-100)	
Prevalence of malignant lesions, %	100	

Abbreviations:

NPV, negative predictive value;

PPV, positive predictive value

<sup>a</sup>One sample has no result because of low follicular content that is not summarized in the table.

[0150] The GSC correctly identified 41 of N45 malignant samples as suspicious, yielding a sensitivity of 91.1% (95% CI, 79-98), and 99 of 145 nonmalignant samples were correctly identified as benign by the GSC, yielding a specificity of 68.3% (95% CI, 60-76). Among Bethesda III and IV samples, the NPV was 96.1% (95% CI, 90-99) and the PPV was 47.1% (95% CI, 36-58). Performance of the GSC was similar between Bethesda III and IV categories (Table 7). [0151] Among the 190 Bethesda III and IV samples, 17 (8.9%) were histologically Hürthle cell adenomas and 9 (4.7%) were Hürthle cell carcinomas, while 164 samples (86.3%) were histologically non-Hürthle. For samples with Hürthle histology, the sensitivity was 88.9% (95% CI, 52-100) and the specificity was 58.8% (95% CI, 33-82). For

samples with non-Hürthle histology, the sensitivity was 91.7% (95% CI, 78-98) and the specificity was 69.5% (95% CI, 61-77).

[0152] A wide variety of malignant subtypes were correctly classified as suspicious (Table 8). Four false-negative cases occurred (Table 9). Patient age or sex, malignancy subtype, or nodule size by ultrasonography or on histopathology were assessed to determine whether they associated with false-negative cases, and none were. The performance of the GSC in secondary analyses of nodules with Bethesda 11, V, or VI cytopathology are reported in Table 7. Among the entire secondary analysis group, the GSC sensitivity was 100% (95% CI, 90-100) and the specificity was 73.1% (95% CI, 52-88).

TABLE 8

Performance of Genomic Sequencing Classifier (GSC) According to Histopathological Subtype.		
Histopathological Subtype	Nodules, No. (%)	Result with GSC, Benign, No./Suspicious, No.
<b>Benign</b>		
Total, No.	145	NA
Benign follicular nodule	49 (33.8)	33/11
Hyperplastic nodule	5 (3.4)	5/0
Follicular adenoma	54 (37.2)	37/17
Follicular tumor of uncertain malignant potential	9 (6.2)	4/5
Well-differentiated tumor of uncertain malignant potential	8 (5.5)	4/4
Hürthle cell adenoma	17 (11.7)	10/7
Chronic lymphocytic thyroiditis	2 (1.4)	1/1
Hyalinizing trabecular adenoma	1 (0.7)	0/1
<b>Malignant</b>		
Total, No.	45	NA
Papillary thyroid carcinoma	15 (33.3)	2/13
Tall-cell variant	1 (2.2)	0/1
Follicular carcinoma	11 (24.4)	1/10
Hürthle cell carcinoma <sup>a</sup>	9 (20.0)	1/8
Follicular carcinoma <sup>b</sup>	7 (15.6)	0/7
Poorly differentiated carcinoma	1 (2.2)	0/1
Medullary thyroid cancer	1 (2.2)	0/1

Abbreviation:

NA, not applicable

<sup>a</sup>Among the Hürthle cell carcinomas, 7 showed capsular invasion and 2 showed vascular invasion. The false-negative case was previously false-negative on the gene expression classifier.<sup>20</sup>

<sup>b</sup>Among the follicular carcinomas, 3 showed capsular invasion and 4 were well-differentiated carcinomas not otherwise specified.

TABLE 9

Cytologic Findings and Histopathological Diagnosis in 4 False-Negative Results on Genomic Sequencing Classification

Patient No./Sex	Nodule Size, cm Ultrasougraphic Imaging	Bethesda Pathological Examination	Cytologic Diagnosis	Final Histologic Diagnosis
1/M	1.1	1.2	III	PTC
2/F	2.5	1.5	III	PTC
3/F	3.2	3.0	IV	FVPTC
4/F	2.9	3.5	IV	HCC-v

Abbreviations: FVPTC, papillary thyroid cancer follicular variant; HCC-v, Hurthle cell carcinoma, vascular invasion; PTC, papillary thyroid cancer.

[0153] Genomic sequence classifier to gene expression classifier comparison on a per-samples basis: 190 Bethesda III/IV primary validation samples yielded both GSC and GEC results (FIG. 5, Table 10). GSC had 99 true negative results; 67 of which were also benign per the GEC, and 32 were GEC suspicious (false positive). GSC had 46 false positive results; 40 of which were also suspicious per the GEC, and 6 were GEC benign (true negative). Of all benign samples (145), GSC reclassified as benign 32 of the GEC's 72 false positive results. Conversely, only 6 of the GEC's 73 true negative results were incorrectly classified as GSC suspicious. The net reclassification of 26 benign nodules to a GSC benign result accounts for the rise in GSC specificity compared to the GEC. GSC had 41 true positive results; 39 of which were also suspicious per the GEC, and 2 were GEC benign (false negative). GSC had 4 false negative results; 3 of which were also benign per the GEC, and 1 was GEC suspicious (true positive). Of all malignant samples (45), GSC reclassified as suspicious 2 of the GEC's 5 false negative results. Conversely, only 1 of the GEC's 40 true positive results were incorrectly classified as GSC benign. The net reclassification of 1 malignant nodules to a GSC suspicious result accounts for the maintained sensitivity of the GSC compared to the GEC.

clinical validation of the GSC in a prospectively collected, representative, universally operated, and histopathologically diagnosed cohort demonstrates the required high NPV across these ranges of cancer prevalence encountered in Bethesda III and IV nodules in clinical practice (FIG. 3). To independently validate the GSC a set of strict blinding and de-identification protocols were implemented that enabled the use of the same FNA samples previously used to validate the GEC. Use of these samples allowed testing of complete and representative sets of nodules with corresponding surgical histology unaffected by the current widespread use of molecular testing to avoid or encourage surgery.

[0155] Test sensitivity of the GSC (91%; 95% CI, 79-98) compared with the GEC (89%; 95% CI, 76-96) was maintained, with the point estimate within the counterpart's 95% CI, and the McNemar  $\chi^2$  test ( $df=1$ ) on the matched sample set renders a test statistic of 0 ( $P>0.99$ ). On the other hand, test specificity of the GSC (68%; 95% CI, 60-76) was significantly improved from the GEC (50%; 95% CI, 42-59), with the point estimate outside the counterpart's 95% CI, and the McNemar  $\chi^2$  test ( $df=1$ ) on the matched sample set renders a test statistic of 16.447 ( $P<0.001$ ) (Table 10). In practice, this enhanced performance indicates that among Bethesda III and IV nodules that are histopathologically benign, at least one-third more will receive a benign result using the GSC compared with the GEC (FIG. 5, and FIG. 7). At a cancer prevalence of 24%, more than half of tested patients are projected to receive a GSC benign result, and among GSC suspicious nodules, nearly half are anticipated to have cancer on surgical histology. This increased benign call rate is expected to result in more patients being assigned to active observation as opposed to diagnostic surgery. FIG. 6, for example, illustrates the treatment recommendations to the patients based on the results from Afirma GSC. Given the high cost of surgery in the United States among Medicare and private payers, the increased avoidance of diag-

TABLE 10

Performance comparison between the genomic sequence classifier and gene expression classifier

		GEC				
		Histo B		Histo M		
		True Negative	False Positive	True Positive	False Negative	
GSC	Histo B	True Negative (TN)	67	32		99
		False Positive (FP)	6	40		46
	Histo M	True Positive (TP)		39	2	41
		False Negative (FN)		1	3	4
		73	72	40	5	190

[0154] A 2016 meta-analysis reported the risks of malignancy among Bethesda III and IV thyroid nodules to be 17% (95% CI, 11-23) and 25% (95% CI, 20-29), respectively. To safely avoid unnecessary diagnostic surgery among these cytologically indeterminate nodules, a test with a high sensitivity and NPV for malignancy is required. This blinded

nostic surgery because of GSC benign results is expected to further improve cost-effectiveness and reduce surgical complications.

[0156] While genomic data has been incorporated in clinical management decisions of multiple medical conditions for more than a decade, progress continues toward under-

standing the complexities of genomic and non-genomic pathways in the development and behavior of disease. Current evidence suggests that most common diseases are associated with small effects from a large number of genes and that most of these contributions are derived from transcriptionally active portions of the genome. This implies that diseases such as thyroid cancer are unlikely to be accounted for by the effects of a small number of genes. The fact that few genomic variants are associated with 100% penetrance toward malignant histology suggests that a complex interaction of multiple factors ultimately determines the benign or malignant nature of thyroid nodules. As the number of these factors expands, it becomes critical to use machine learning and statistical models to interpret their signals in a trained model to derive an accurate diagnosis. [0157] Hürthle lesions exemplify the challenges inherent in complex biology and the opportunity to harness high dimensional genomic data for predictive model training and subsequent validation. Most Hurthle cell-dominant Bethesda III and IV thyroid nodules have historically undergone surgery given the potential for Hurthle cell carcinoma, yet most have proven to be histologically benign. The GEC identified these samples at a high NPV, but most were categorized as GEC suspicious. Current methods sought to maintain a high NPV while providing more benign results by including 2 dedicated classifiers to work with the core GSC classifier. Among the 26 Hurthle cell adenomas or Hurthle cell carcinomas reported here, the final GSC sensitivity was 88.9% and the specificity was 58.8%; the GEC sensitivity was 88.9% and the specificity was 11.8% among these same neoplasms. Thus, while the overall GSC sensitivity of 91.1% reported here is comparable with that of the GEC (by design), the improved overall GSC specificity of 68.3% results from significantly improved performances among both Hurthle and non-Hurthle specimen types. Given that most histologically benign Hurthle and non-Hurthle specimens are now both identified as GSC benign, GSC testing may further safely reduce unnecessary surgery among both specimen types.

[0158] A secondary analysis of 61 Bethesda II, V, or VI samples that also were included in the GEC validation study is included in Table 7. The consistency of these performance metrics within the Bethesda III and IV categories is reassuring and supportive of the findings in the primary analysis.

[0159] Methods and systems of the present disclosure may be combined with or modified by other methods or systems, such as, for example, those described in U.S. Pat. No. 8,541,170, U.S. Patent Publication No. 2018/0157789, and U.S. Patent Publication No. 2018/0016642, each of which is entirely incorporated herein by reference.

[0160] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. It is not intended that the invention be limited by the specific examples provided within the specification. While the invention has been described with reference to the aforementioned specification, the descriptions and illustrations of the embodiments herein are not meant to be construed in a limiting sense. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. Furthermore, it shall be understood that all aspects of the invention are not limited to the specific depictions, configurations or relative proportions set forth

herein which depend upon a variety of conditions and variables. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is therefore contemplated that the invention shall also cover any such alternatives, modifications, variations or equivalents. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

What is claimed is:

1. A method for diagnosing thyroid disease in a subject, the method comprising:
  - (a) providing a DNA sample from a subject;
  - (b) detecting the presence of one or more polymorphisms selected from the polymorphisms listed in Tables 3-6 or their complement; and
  - (c) determining whether said subject has or is likely to have a malignant or benign thyroid condition based on the results of step (b).
2. The method of claim 1, wherein the malignant condition is selected from the group consisting of follicular carcinoma, follicular variant of papillary carcinoma, and papillary thyroid carcinoma.
3. The method of claim 1, wherein the benign thyroid condition is selected from the group consisting of follicular adenoma, and nodular hyperplasia.
4. The method of claim 1, wherein the DNA sample provided from said subject is obtained from a sample comprising thyroid tissue.
5. The method of claim 1, wherein the polymorphism comprises a variation in copy number as compared to a normal sample.
6. The method of claim 5, wherein the variation in copy number as compared to a normal sample comprises a deletion.
7. The method of claim 5, wherein the variation in copy number as compared to a normal sample comprises an increase in the copy number.
8. The method of claim 5, wherein the normal sample comprises a sample of DNA from the same subject.
9. The method of claim 5, wherein the normal sample comprises a sample of DNA from a different subject.
10. The method of claim 5, wherein the normal sample comprises a known or generally accepted value.
11. The method of claim 1, wherein the detecting step (b) comprises:
  - (a) contacting said DNA sample with one or more binding agents that specifically bind to the one or more polymorphisms listed in Tables 3-6, or their complement; and
  - (b) determining whether said DNA sample specifically binds to said one or more binding agents, wherein binding of said DNA sample to said one or more binding agents indicates the presence of the polymorphism in said subject.
12. The method of claim 1, wherein the detecting step (b) comprises sequencing of one or more nucleic acid regions comprising the one or more marker regions listed in Tables 3-6 or their complement.
13. The method of claim 1, wherein the detecting step (b) comprises quantifying the amount of DNA comprising the one or more marker regions listed in Tables 3-6 or their complement.

**14.** The method of claim **13**, wherein the quantifying comprises PCR.

**15.** The method of claim **14**, wherein the PCR comprises real-time PCR.

**16.** The method of claim **13**, wherein the quantifying comprises hybridization.

**17.** The method of claim **1**, wherein the method further comprises determining the expression level of one or more genes correlated with follicular adenoma, follicular carcinoma, nodular hyperplasia, follicular variant of papillary carcinoma, or papillary thyroid carcinoma.

**18.** A composition comprising one or more binding agents that specifically bind to the one or more polymorphisms listed in Tables 3-6, or their complement.

**19.** A kit for diagnosing thyroid disease in a subject, the kit comprising:

- (a) at least one binding agent that specifically binds to the one or more polymorphisms selected from the group consisting of the polymorphisms listed in Tables 3-6, or their complement; and
- (b) reagents for detecting binding of said at least one binding agent to a DNA sample from a subject.

\* \* \* \* \*